7.1 Homologieklonierung und Proteinanalyse: Ergebnisse der Datenbanksuche

blastx-searches mit DNA-Fragmenten:

hy-1_r43a18006 (s. Abschn. 7.1.1)
hy-2_r43a30007 (s. Abschn. 7.1.2)
hy-3_11a245004 (s. Abschn. 7.1.3)
hy-4_11b250001 (s. Abschn. 7.1.4)
hy-5_11b250002 (s. Abschn. 7.1.5)
hy-6_11b250005 (s. Abschn. 7.1.6)
hy-7_11b235002 (s. Abschn. 7.1.7)
hy-8_13a250005 (s. Abschn. 7.1.8)
hy-9_13b236005 (s. Abschn. 7.1.9)
hy-10_15b145025 (s. Abschn. 7.1.10)
hy-11_16b50013 (s. Abschn. 7.1.11)
hy-12_17a150017 (s. Abschn. 7.1.12)
hy-13_20235003 (s. Abschn. 7.1.13)
hy-14_20240016 (s. Abschn. 7.1.14)
hy-15_20240017 (s. Abschn. 7.1.15)
hy-16_12169804 (s. Abschn. 7.1.16)
hy-17_lf1-1a (s. Abschn. 7.1.17)
04199901 (*upstream-message*) (s. Abschn. 7.1.18)

Proteinanalysen mit der pAC:

balstp (s. Abschn. 7.1.1)
Pfam-search (s. Abschn. 7.1.20)
Smith/Waterman-Suche (s. Abschn. 7.1.1)
TopPred2 (s. Abschn. 7.1.1)

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020950350-010982-11963
```

7.1.1 Query= hy-1_r43a18006

```
         (151 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           933,889 sequences; 292,803,047 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 147 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                        (bits) Value

gi|17554726|ref|NP_497972.1|  (NM_065571) Ras-like protein [...     41   0.003
gi|6912698|ref|NP_036382.1|  (NM_012250) oncogene TC21 [Homo...     37   0.054
gi|11436135|ref|XP_006102.1|  (XM_006102) oncogene TC21 [Hom...     37   0.054
gi|13399308|ref|NP_080122.1|  (NM_025846) related RAS viral ...     37   0.054
gi|4150910|emb|CAA77070.1|  (Y18167) ras protein [Suberites ...     37   0.054
gi|12834582|dbj|BAB22967.1|  (AK003733) data source:SPTR, so...     37   0.054
gi|13537429|dbj|BAB40669.1|  (AB054578) small GTPase Rab1 [E...     36   0.091
gi|13646284|ref|XP_016531.1|  (XM_016531) similar to Ras-rel...     36   0.091
gi|6919951|sp|Q12526|RAS_EMENI  RAS-LIKE PROTEIN >gi|1078631...     35   0.16
gi|5738166|gb|AAD50280.1|AF169950_1  (AF169950) putative int...     35   0.16
gi|17737457|ref|NP_523687.1|  (NM_078963) Rab-protein 3 [Dro...     35   0.20
gi|19114491|ref|NP_593579.1|  (NC_003424) ras-like protein 1...     35   0.20
gi|68941|pir||TVBYPR  transforming protein ras - fission yea...     35   0.20
gi|6919949|sp|O93856|RAS_LACBI  RAS-LIKE PROTEIN >gi|4104252...     35   0.20
gi|20139581|sp|Q96AX2|RB37_HUMAN  Ras-related protein Rab-37...     35   0.20
gi|17457950|ref|XP_058828.1|  (XM_058828) similar to RAB37, ...     35   0.20
gi|7657492|ref|NP_055168.1|  (NM_014353) RAB26, member RAS o...     35   0.20
gi|13537437|dbj|BAB40673.1|  (AB054582) small GTPase Rab5 [E...     35   0.27
gi|6678930|ref|NP_032650.1|  (NM_008624) muscle and microspi...     35   0.27
gi|6981222|ref|NP_037113.1|  (NM_012981) muscle and microspi...     35   0.27
gi|1083775|pir||JC2528  GTP-binding protein Rab26 - rat             35   0.27
gi|10946770|ref|NP_067386.1|  (NM_021411) RAB37, member of R...     35   0.27
gi|19424272|ref|NP_598264.1|  (NM_133580) RAB26, member RAS ...     35   0.27
gi|825440|gb|AAB07703.1|  (L42299) RAS [Aspergillus fumigatus]      35   0.27
gi|1845598|gb|AAB47925.1|  (L49400) Rab3 [Loligo pealei]            34   0.35
```

```
gi|392973|gb|AAA03315.1|  (U00986) Rab3 [Aplysia californica]      34   0.35
gi|13375179|emb|CAC34553.1|  (AJ409198) putative GTPase [Pla...    34   0.35
gi|3334326|sp|O14462|SEC4_CANAL  Ras-related protein SEC4 >g...    34   0.35
gi|17137220|ref|NP_477172.1|  (NM_057824) Rab6-P1; warthog [...    34   0.35
gi|6912514|ref|NP_036351.1|  (NM_012219) muscle RAS oncogene...    34   0.35
gi|131881|sp|P22126|RAS1_NEUCR  RAS-1 PROTEIN >gi|101876|pir...    34   0.35
gi|14423577|gb|AAK62471.1|AF389109_1  (AF389109) small GTP-b...    34   0.45
gi|11558647|emb|CAC17832.1|  (AJ278658) secretion related GT...    33   0.59
gi|12084563|pdb|1G16|A  Chain A, Crystal Structure Of Sec4-G...    33   0.59
gi|12084567|pdb|1G17|A  Chain A, Crystal Structure Of Sec4-G...    33   0.59
gi|7496249|pir||T15546  hypothetical protein C18A3.6 - Caeno...    33   0.59
gi|20071543|gb|AAH26915.1|  (BC026915) Similar to RAB6, memb...    33   0.59
gi|17535675|ref|NP_495129.1|  (NM_062728) synaptic vesicle-a...    33   0.59
gi|17535677|ref|NP_495128.1|  (NM_062727) synaptic vesicle-a...    33   0.59
gi|12381851|emb|CAC24717.1|  (AJ304407) Sec4p [Pichia pastoris]    33   0.59
gi|13195674|ref|NP_077249.1|  (NM_024287) RAB6, member RAS o...    33   0.59
gi|17458313|ref|XP_035122.2|  (XM_035122) RAB6A, member RAS ...    33   0.59
gi|17512290|gb|AAH19118.1|AAH19118  (BC019118) Similar to RA...    33   0.59
gi|19923231|ref|NP_002860.2|  (NM_002869) RAB6A, member RAS ...    33   0.59
gi|2959746|emb|CAA12071.1|  (AJ224685) putative SEC4 protein...    33   0.59
gi|6685834|sp|Q9WVB1|RB6A_RAT  Ras-related protein Rab-6A >g...    33   0.59
gi|14318517|ref|NP_116650.1|  (NC_001138) Involved in transp...    33   0.59
gi|17553774|ref|NP_498993.1|  (NM_066592) Rab6 (ras protein)...    33   0.77
gi|2342660|gb|AAB67800.1|  (AF014120) GTP-binding protein sp...    33   0.77
gi|68942|pir||TVBYSR  transforming protein ras - fission yea...    33   1.0
gi|11274531|pir||T45545  transforming protein ras - fission ...    33   1.0
gi|19115492|ref|NP_594580.1|  (NC_003424) ypt1-related prote...    33   1.0
gi|17509233|ref|NP_491857.1|  (NM_059456) RAS [Caenorhabditi...    33   1.0
gi|7767071|pdb|1EK0|A  Chain A, Gppnhp-Bound Ypt51 At 1.48 A...    32   1.3
gi|17570073|ref|NP_510790.1|  (NM_078389) ras-related protei...    32   1.3
gi|7497401|pir||T32953  hypothetical protein C44C11.1 - Caen...    32   1.3
gi|6324663|ref|NP_014732.1|  (NC_001147) Rab5-like GTPase in...    32   1.3
gi|14475537|emb|CAC41973.1|  (AJ272025) putative Rab/GTPase ...    32   1.3
gi|19114161|ref|NP_593249.1|  (NC_003424) gtp-binding protei...    32   1.3
gi|417591|sp|P32254|RASS_DICDI  RAS-LIKE PROTEIN RASS >gi|32...    32   1.3
gi|17535679|ref|NP_496623.1|  (NM_064222) RAS-related protei...    32   1.3
gi|68952|pir||TVFFR3  transforming protein ras3 - fruit fly ...    32   1.7
gi|5729997|ref|NP_004154.2|  (NM_004163) RAB27B, member RAS ...    32   1.7
gi|17136706|ref|NP_476857.1|  (NM_057509) R-P1; Enhancer of ...    32   1.7
gi|17066210|emb|CAD12439.1|  (AJ420321) Rab5c GTPase [Plasmo...    32   1.7
gi|8133108|gb|AAF73473.1|AF268471_1  (AF268471) Ras1p [Schiz...    32   1.7
gi|5738170|gb|AAD50282.1|AF169952_1  (AF169952) putative int...    32   2.3
gi|223831|prf||1001202A  protein c-ras sc1 [Saccharomyces ce...    32   2.3
gi|405066|gb|AAA21444.1|  (U01051) rap homologue 1 [Entamoeb...    32   2.3
gi|6324675|ref|NP_014744.1|  (NC_001147) ras proto-oncogene ...    32   2.3
gi|172361|gb|AAA34958.1|  (K01970) RAS1 protein [Saccharomyc...    32   2.3
gi|13470090|ref|NP_076341.1|  (NM_023852) RAB3C, member RAS ...    31   2.9
gi|6808528|gb|AAF28422.1|AF124200_1  (AF124200) Rab6-like pr...    31   2.9
gi|18844918|dbj|BAB85387.1|  (AP003924) putative RIC1_ORYSA ...    31   2.9
gi|6174943|sp|Q63482|RB3C_RAT  RAS-RELATED PROTEIN RAB-3C          31   2.9
gi|1045640|gb|AAC52704.1|  (U37099) rab3c [Rattus norvegicus]      31   2.9
gi|13537433|dbj|BAB40671.1|  (AB054580) small GTPase RabD [E...    31   2.9
gi|730510|sp|P40392|RIC1_ORYSA  RAS-RELATED PROTEIN RIC1 >gi...    31   2.9
gi|18598963|ref|XP_086934.1|  (XM_086934) similar to RAB6C, ...    31   2.9
gi|14721222|ref|XP_038274.1|  (XM_038274) RAB6C, member RAS ...    31   2.9
gi|14149799|ref|NP_115520.1|  (NM_032144) RAB6C, member RAS ...    31   2.9
gi|7438392|pir||T14391  GTP-binding protein homolog - turnip...    31   3.8
gi|16159420|ref|XP_056817.1|  (XM_056817) hypothetical prote...    31   3.8
gi|89584|pir||C29224  GTP-binding protein smg-25C - bovine         31   3.8
gi|17737663|ref|NP_524172.1|  (NM_079448) Rab-protein 8 [Dro...    31   3.8
gi|16758202|ref|NP_445911.1|  (NM_053459) RAB27B, member RAS...    31   3.8
```

```
gi|13385282|ref|NP_085031.1|  (NM_030554) RAB27b, member RAS...   31   3.8
gi|7706675|ref|NP_057661.1|  (NM_016577) RAB6B, member RAS o...   31   3.8
gi|6919950|sp|P87018|RAS_BOTCI  RAS-LIKE PROTEIN >gi|1916796...   31   3.8
gi|16157957|ref|XP_055543.1|  (XM_055543) similar to RAB6C, ...   31   3.8
gi|14275890|dbj|BAB58891.1|  (AB028165) rab-like protein E [...   30   5.0
gi|730474|sp|P38976|RAS2_HYDMA  RAS-LIKE PROTEIN RAS2 >gi|32...   30   5.0
gi|19424194|ref|NP_598220.1|  (NM_133536) RAB3C, member RAS ...   30   5.0
gi|5764095|gb|AAD51132.1|AF107762_1  (AF107762) small GTP-bi...   30   5.0
gi|10047433|gb|AAG12240.1|AF183937_1  (AF183937) guanine nuc...   30   5.0
gi|131848|sp|P22128|RAB8_DISOM  RAS-RELATED PROTEIN RAB-8 (O...   30   5.0
gi|15224916|ref|NP_181989.1|  (NM_130025) putative small GTP...   30   5.0
gi|464532|sp|P34143|RABC_DICDI  RAS-RELATED PROTEIN RABC >gi...   30   5.0
gi|7438385|pir||T03627  GTP-binding protein Rab6 - common to...   30   5.0
gi|8394142|ref|NP_059013.1|  (NM_017317) low Mr GTP-binding ...   30   5.0
```

Alignments

```
>gi|17554726|ref|NP_497972.1|  (NM_065571) Ras-like protein [Caenorhabditis elegans]
 gi|7499318|pir||T21066  hypothetical protein F17C8.4 - Caenorhabditis elegans
 gi|1737463|gb|AAB38969.1|  (U80675) GTP binding protein [Caenorhabditis elegans]
 gi|3876002|emb|CAA84796.1|  (Z35719) contains similarity to Pfam domain: PF00071 (Ras
family), Score=295.6, E-value=2e-85, N=1 [Caenorhabditis elegans]
          Length = 211

 Score = 41.2 bits (95), Expect = 0.003
 Identities = 23/33 (69%), Positives = 26/33 (78%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNV 100
           QRVVS+Q+G ELA QL + YIETSAK  P VNV
Sbjct: 137 QRVVSEQEGRELAAQLKLMYIETSAK-EPPVNV 168



>gi|6912698|ref|NP_036382.1|  (NM_012250) oncogene TC21 [Homo sapiens]
 gi|68947|pir||TVHUC2  GTP-binding protein RAS/TC21 - human
 gi|190877|gb|AAA36545.1|  (M31468) ras-like protein [Homo sapiens]
 gi|20147741|gb|AAM12638.1|AF493924_1  (AF493924) Ras family small GTP binding protein TC21
[Homo sapiens]
          Length = 203

 Score = 37.0 bits (84), Expect = 0.054
 Identities = 17/35 (48%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR V+Q++G++LA QL V Y+E SAK+   V+ A+
Sbjct: 133 QRQVTQEEGQQLARQLKVTYMEASAKIRMNVDQAF 167



>gi|11436135|ref|XP_006102.1|  (XM_006102) oncogene TC21 [Homo sapiens]
 gi|2507282|sp|P17082|RRA2_HUMAN  Ras-related protein R-Ras2 (Ras-like protein TC21)
(Teratocarcinoma oncogene)
 gi|13278028|gb|AAH03871.1|AAH03871  (BC003871) RIKEN cDNA 2610016H24 gene [Mus musculus]
```

```
 gi|15341857|gb|AAH13106.1|AAH13106  (BC013106) Unknown (protein for MGC:21928) [Homo
sapiens]
          Length = 204

 Score = 37.0 bits (84), Expect = 0.054
 Identities = 17/35 (48%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR V+Q++G++LA QL V Y+E SAK+   V+ A+
Sbjct: 134 QRQVTQEEGQQLARQLKVTYMEASAKIRMNVDQAF 168
```

```
>gi|13399308|ref|NP_080122.1|  (NM_025846) related RAS viral (r-ras) oncogene homolog 2;
RIKEN cDNA 2610016H24 gene [Mus musculus]
 gi|12847531|dbj|BAB27607.1|  (AK011419) data source:SPTR, source key:P17082,
evidence:ISS~homolog to RAS-RELATED PROTEIN R-RAS2 (RAS-LIKE PROTEIN TC21) (TERATOCARCINOMA
ONCOGENE)~putative [Mus musculus]
          Length = 204

 Score = 37.0 bits (84), Expect = 0.054
 Identities = 17/35 (48%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR V+Q++G++LA QL V Y+E SAK+   V+ A+
Sbjct: 134 QRQVTQEEGQQLARQLKVTYMEASAKIRMNVDQAF 168
```

```
>gi|4150910|emb|CAA77070.1|  (Y18167) ras protein [Suberites domuncula]
          Length = 191

 Score = 37.0 bits (84), Expect = 0.054
 Identities = 16/26 (61%), Positives = 21/26 (80%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVV+  +GEELA QL + Y+ETSAK
Sbjct: 129 ERVVTLSEGEELAQQLKIKYVETSAK 154
```

```
>gi|12834582|dbj|BAB22967.1|  (AK003733) data source:SPTR, source key:P17082,
evidence:ISS~homolog to RAS-RELATED PROTEIN R-RAS2 (RAS-LIKE PROTEIN TC21) (TERATOCARCINOMA
ONCOGENE)~putative [Mus musculus]
          Length = 92

 Score = 37.0 bits (84), Expect = 0.054
 Identities = 17/35 (48%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR V+Q++G++LA QL V Y+E SAK+   V+ A+
Sbjct: 22  QRQVTQEEGQQLARQLKVTYMEASAKIRMNVDQAF 56
```

```
>gi|13537429|dbj|BAB40669.1|  (AB054578) small GTPase Rab1 [Entamoeba histolytica]
          Length = 205

 Score = 36.2 bits (82), Expect = 0.091
 Identities = 16/37 (43%), Positives = 26/37 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAWVQ 112
           +R V+ +QG+ELAN LNV ++E SAK    V+  +++
Sbjct: 128 ERAVTVEQGQELANSLNVSFMEASAKKSLNVDELFIR 164



>gi|13646284|ref|XP_016531.1|  (XM_016531) similar to Ras-related protein R-Ras2 (Ras-
like protein TC21) (Teratocarcinoma oncogene) [Homo sapiens]
          Length = 127

 Score = 36.2 bits (82), Expect = 0.091
 Identities = 17/35 (48%), Positives = 25/35 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR V+Q++G +LA QL V Y+E SAK+   V+ A+
Sbjct: 57  QRQVTQEEGRQLARQLKVTYMEASAKIRMNVDQAF 91



>gi|6919951|sp|Q12526|RAS_EMENI  RAS-LIKE PROTEIN
 gi|1078631|pir||A53778  GTP-binding protein A-ras - Emericella nidulans
 gi|458024|gb|AAA20964.1|  (U03023) Ras-like protein [Aspergillus nidulans]
 gi|531818|gb|AAA20965.1|  (U03025) Ras-like protein [Aspergillus nidulans]
          Length = 212

 Score = 35.4 bits (80), Expect = 0.16
 Identities = 16/26 (61%), Positives = 20/26 (76%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVVS+Q+GE LA Q    +IETSAK
Sbjct: 128 ERVVSEQEGESLARQFGCKFIETSAK 153



>gi|5738166|gb|AAD50280.1|AF169950_1  (AF169950) putative intermediate compartment protein
[Tetrahymena thermophila]
          Length = 205

 Score = 35.4 bits (80), Expect = 0.16
 Identities = 17/35 (48%), Positives = 25/35 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           QR VS ++GEELA++  V + ETSAK G  ++ A+
Sbjct: 130 QRKVSTREGEELADRYGVKFFETSAKEGVNISEAF 164
```

```
>gi|17737457|ref|NP_523687.1|   (NM_078963) Rab-protein 3 [Drosophila melanogaster]
 gi|131800|sp|P25228|RAB3_DROME  Ras-related protein Rab-3
 gi|84987|pir||JH0425  GTP-binding protein rab3A homolog - fruit fly (Drosophila
melanogaster)
 gi|158193|gb|AAA28843.1|   (M64621) rab3 [Drosophila melanogaster]
 gi|7303712|gb|AAF58762.1|   (AE003828) Rab3 gene product [Drosophila melanogaster]
 gi|16648446|gb|AAL25488.1|   (AY060449) LP05860p [Drosophila melanogaster]
          Length = 220

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 15/26 (57%), Positives = 22/26 (83%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
            QRV+S ++G +LA+QL V++ ETSAK
Sbjct: 141 QRVISFERGRQLADQLGVEFFETSAK 166



>gi|19114491|ref|NP_593579.1|  (NC_003424) ras-like protein 1. [Schizosaccharomyces pombe]
 gi|3123259|sp|P08647|RAS_SCHPO  Ras-like protein 1
 gi|7493264|pir||T37875  ras-like protein 1 - fission yeast (Schizosaccharomyces pombe)
 gi|2330716|emb|CAB11218.1|  (Z98597) ras-like protein 1. [Schizosaccharomyces pombe]
          Length = 219

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 17/35 (48%), Positives = 25/35 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
            +RVVS+ +GE+LA   ++  Y+ETSAKL   V  A+
Sbjct: 128 ERVVSRAEGEQLAKSMHCLYVETSAKLRLNVEEAF 162



>gi|68941|pir||TVBYPR  transforming protein ras - fission yeast  (Schizosaccharomyces pombe)
 gi|5036|emb|CAA27399.1|  (X03771) put. ras protein [Schizosaccharomyces pombe]
          Length = 214

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 17/35 (48%), Positives = 25/35 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
            +RVVS+ +GE+LA   ++  Y+ETSAKL   V  A+
Sbjct: 123 ERVVSRAEGEQLAKSMHCLYVETSAKLRLNVEEAF 157



>gi|6919949|sp|O93856|RAS_LACBI  RAS-LIKE PROTEIN
 gi|4104252|gb|AAD01987.1|  (AF034098) ras protein [Laccaria bicolor]
          Length = 209

 Score = 35.0 bits (79), Expect = 0.20
```

```
 Identities = 16/26 (61%), Positives = 20/26 (76%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVVS+Q+GE LA Q     +IETSAK
Sbjct: 128 ERVVSKQEGESLARQFGCKFIETSAK 153



>gi|20139581|sp|Q96AX2|RB37_HUMAN  Ras-related protein Rab-37
 gi|16741621|gb|AAH16615.1|AAH16615  (BC016615) Similar to RAB37, member of RAS oncogene
family [Homo sapiens]
          Length = 223

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 15/36 (41%), Positives = 24/36 (66%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAWV 109
           +RV+  + GE LA +  V ++ETSAK G  V +A++
Sbjct: 150 ERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFL 185



>gi|17457950|ref|XP_058828.1|  (XM_058828) similar to RAB37, member of RAS oncogene
family [Homo sapiens]
          Length = 180

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 15/36 (41%), Positives = 24/36 (66%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAWV 109
           +RV+  + GE LA +  V ++ETSAK G  V +A++
Sbjct: 107 ERVIRSEDGETLAREYGVPFLETSAKTGMNVELAFL 142



>gi|7657492|ref|NP_055168.1|  (NM_014353) RAB26, member RAS oncogene family [Homo
sapiens]
 gi|12230537|sp|Q9ULW5|RB26_HUMAN  Ras-related protein Rab-26
 gi|5931612|dbj|BAA84707.1|  (AB027137) RAB-26 [Homo sapiens]
 gi|14043379|gb|AAH07681.1|AAH07681  (BC007681) RAB26, member RAS oncogene family [Homo
sapiens]
 gi|20379080|gb|AAM21100.1|AF498952_1  (AF498952) small GTP binding protein RAB26 [Homo
sapiens]
          Length = 190

 Score = 35.0 bits (79), Expect = 0.20
 Identities = 15/35 (42%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           +RVV ++ GE+LA +  + ++ETSAK G  V++A+
Sbjct: 118 ERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAF 152
```

```
>gi|13537437|dbj|BAB40673.1|  (AB054582) small GTPase Rab5 [Entamoeba histolytica]
          Length = 195

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 15/31 (48%), Positives = 22/31 (70%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           RV+++++ E  A  L++DYIETSAK    VN
Sbjct: 128 RVITKEEAEGYARSLSIDYIETSAKANINVN 158



>gi|6678930|ref|NP_032650.1| ☐ (NM_008624) muscle and microspikes RAS [Mus musculus]
 gi|13124463|sp|O08989|RASM_MOUSE  Ras-related protein M-Ras (Ras-related protein R-Ras3)
(Muscle and microspikes Ras) (X-Ras)
 gi|2209065|dbj|BAA20538.1| ☐ (AB004879) M-Ras [Mus musculus]
 gi|4104046|gb|AAD01926.1| ☐ (AF031159) X-ras [Mus musculus]
 gi|4105151|gb|AAD02277.1| ☐ (AF043581) M-Ras [Mus musculus]
 gi|19353897|gb|AAH24389.1| ☐ (BC024389) muscle and microspikes RAS [Mus musculus]
          Length = 208

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 17/32 (53%), Positives = 24/32 (74%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVNV 100
           R V+++ QG+E+A + N+ YIETSAK  P +NV
Sbjct: 134 RKVTRDQGKEMATKYNIPYIETSAK-DPPLNV 164



>gi|6981222|ref|NP_037113.1| ☐ (NM_012981) muscle and microspikes RAS [Rattus norvegicus]
 gi|2851409|sp|P97538|RASM_RAT  Ras-related protein M-Ras (Ras-related protein R-Ras3)
 gi|7438441|pir||T10774  M-Ras GTP-phosphohydrolase (EC 3.6.1.-) - rat
 gi|2209075|dbj|BAA20531.1| ☐ (D89863) M-Ras [Rattus norvegicus]
          Length = 208

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 17/32 (53%), Positives = 24/32 (74%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVNV 100
           R V+++ QG+E+A + N+ YIETSAK  P +NV
Sbjct: 134 RKVTRDQGKEMATKYNIPYIETSAK-DPPLNV 164



>gi|1083775|pir||JC2528  GTP-binding protein Rab26 - rat
          Length = 190

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 15/35 (42%), Positives = 26/35 (73%)
 Frame = +2
```

```
Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           +RVV ++ GE+LA +  + ++ETSAK G  V++A+
Sbjct: 118 ERVVKREDGEKLAKEYGLPFMETSAKSGLNVDLAF 152




>gi|10946770|ref|NP_067386.1|  (NM_021411) RAB37, member of RAS oncogene family; GTPase
Rab37 [Mus musculus]
 gi|20139748|sp|Q9JKM7|RB37_MOUSE  Ras-related protein Rab-37
 gi|7677422|gb|AAF67162.1|AF233582_1  (AF233582) GTPase Rab37 [Mus musculus]
          Length = 223

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 15/36 (41%), Positives = 24/36 (66%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAWV 109
           +RV+  + GE LA +  V ++ETSAK G  V +A++
Sbjct: 150 ERVIRSEDGETLAREYGVPFMETSAKTGMNVELAFL 185




>gi|19424272|ref|NP_598264.1|  (NM_133580) RAB26, member RAS oncogene family [Rattus
norvegicus]
 gi|1710022|sp|P51156|RB26_RAT  RAS-RELATED PROTEIN RAB-26
 gi|619734|gb|AAA69955.1|  (U18771) Rab26 [Rattus norvegicus]
          Length = 190

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 15/35 (42%), Positives = 26/35 (73%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           +RVV ++ GE+LA +  + ++ETSAK G  V++A+
Sbjct: 118 ERVVKREDGEKLAKEYGLPFMETSAKSGLNVDLAF 152




>gi|825440|gb|AAB07703.1|  (L42299) RAS [Aspergillus fumigatus]
          Length = 213

 Score = 34.7 bits (78), Expect = 0.27
 Identities = 16/26 (61%), Positives = 19/26 (72%)
 Frame = +2


Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +R VSQQ+GE LA Q    +IETSAK
Sbjct: 128 ERAVSQQEGEALARQFGCKFIETSAK 153




>gi|1845598|gb|AAB47925.1|  (L49400) Rab3 [Loligo pealei]
          Length = 220
```

```
 Score = 34.3 bits (77), Expect = 0.35
 Identities = 14/26 (53%), Positives = 23/26 (87%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVVS ++G++LA+QL +++ ETSAK
Sbjct: 141 ERVVSTERGKQLADQLGLEFFETSAK 166



>gi|392973|gb|AAA03315.1|  (U00986) Rab3 [Aplysia californica]
          Length = 219

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 14/26 (53%), Positives = 23/26 (87%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVVS ++G++LA+QL +++ ETSAK
Sbjct: 141 ERVVSTERGKQLADQLGLEFFETSAK 166



>gi|13375179|emb|CAC34553.1|  (AJ409198) putative GTPase [Plasmodium falciparum 3D7]
          Length = 200

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 15/34 (44%), Positives = 24/34 (70%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           R VS ++G+ELA+  N+ ++ETSAK+   V  A+
Sbjct: 129 RNVSYEEGKELADSCNIQFLETSAKIAHNVEQAF 162



>gi|3334326|sp|O14462|SEC4_CANAL  Ras-related protein SEC4
 gi|7493809|pir||T18242  ras protein homolog - yeast (Candida albicans)
 gi|2345152|gb|AAB67974.1|  (AF015306) small GTP-binding protein SEC4p [Candida albicans]
 gi|2642182|gb|AAC50022.1|  (AF017183) Sec4p [Candida albicans]
 gi|3859676|emb|CAA22013.1|  (AL033503) ras-related protein sec4p [Candida albicans]
          Length = 210

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 16/31 (51%), Positives = 23/31 (73%)
 Frame = +2


Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           R VS++QG+ELA +LNV ++E SAK    V+
Sbjct: 135 RQVSKEQGQELAAKLNVPFLEASAKSNENVD 165



>gi|17137220|ref|NP_477172.1|  (NM_057824) Rab6-P1; warthog [Drosophila melanogaster]
 gi|2313037|dbj|BAA21707.1|  (D84314) rab6 [Drosophila melanogaster]
```

```
 gi|7297923|gb|AAF53168.1| ☐ (AE003635) Rab6 gene product [Drosophila melanogaster]
 gi|16648070|gb|AAL25300.1| ☐ (AY060261) GH09086p [Drosophila melanogaster]
          Length = 208

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 132 KRQVSTEEGERKAKELNVMFIETSAKAGYNV 162



>gi|6912514|ref|NP_036351.1| ☐ (NM_012219) muscle RAS oncogene homolog; muscle and
microspikes RAS [Homo sapiens]
 gi|12729365|ref|XP_003111.2| ☐ (XM_003111) muscle RAS oncogene homolog [Homo sapiens]
 gi|6226045|sp|O14807|RASM_HUMAN  Ras-related protein M-Ras (Ras-related protein R-Ras3)
 gi|3228374|gb|AAC52085.1| ☐ (AF022080) R-ras3 [Homo sapiens]
 gi|4105178|gb|AAD02287.1| ☐ (AF043938) Ras-related protein M-Ras/R-Ras3 [Homo sapiens]
 gi|20147729|gb|AAM12632.1|AF493918_1  (AF493918) Ras family small GTP binding protein M-Ras
[Homo sapiens]
          Length = 208

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 16/32 (50%), Positives = 25/32 (78%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVNV 100
           R ++++QG+E+A + N+ YIETSAK  P +NV
Sbjct: 134 RKITREQGKEMATKHNIPYIETSAK-DPPLNV 164



>gi|131881|sp|P22126|RAS1_NEUCR  RAS-1 PROTEIN
 gi|101876|pir||S12892  transforming protein (ras) - Neurospora crassa
 gi|3074|emb|CAA37612.1|  (X53533) NC-ras protein [Neurospora crassa]
 gi|3452585|dbj|BAA32498.1|  (AB016991) NC-ras [Neurospora crassa]
 gi|227285|prf||1701291A  NC-ras protein [Neurospora crassa]
          Length = 213

 Score = 34.3 bits (77), Expect = 0.35
 Identities = 17/35 (48%), Positives = 23/35 (65%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           +RVVS+Q+G+ LA +    YIETSAK    V  A+
Sbjct: 128 ERVVSEQEGQALAAEFGTKYIETSAKTQHNVENAF 162




>gi|14423577|gb|AAK62471.1|AF389109_1  (AF389109) small GTP-binding protein Rab8 [Entamoeba
histolytica]
          Length = 199
```

```
 Score = 33.9 bits (76), Expect = 0.45
 Identities = 12/26 (46%), Positives = 23/26 (88%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RVV+ ++G+++AN+L + ++ETSAK
Sbjct: 127 KRVVTSEEGQQMANKLGIPFLETSAK 152



>gi|11558647|emb|CAC17832.1|  (AJ278658) secretion related GTPase, (SrgA) [Aspergillus
niger]
          Length = 206

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 13/26 (50%), Positives = 21/26 (80%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +R VS +QG++LAN+L + ++E SAK
Sbjct: 130 KRAVSTEQGQQLANELGIPFLEVSAK 155



>gi|12084563|pdb|1G16|A  Chain A, Crystal Structure Of Sec4-Gdp
 gi|12084564|pdb|1G16|B  Chain B, Crystal Structure Of Sec4-Gdp
 gi|12084565|pdb|1G16|C  Chain C, Crystal Structure Of Sec4-Gdp
 gi|12084566|pdb|1G16|D  Chain D, Crystal Structure Of Sec4-Gdp
          Length = 170

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 16/31 (51%), Positives = 21/31 (67%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           RVV+   QGE LA +L + +IE+SAK    VN
Sbjct: 123 RVVTADQGEALAKELGIPFIESSAKNDDNVN 153



>gi|12084567|pdb|1G17|A  Chain A, Crystal Structure Of Sec4-Guanosine-5'-(Beta,Gamma)-
Imidotriphosphate
 gi|12084568|pdb|1G17|B  Chain B, Crystal Structure Of Sec4-Guanosine-5'-(Beta,Gamma)-
Imidotriphosphate
          Length = 170

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 16/31 (51%), Positives = 21/31 (67%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           RVV+   QGE LA +L + +IE+SAK    VN
Sbjct: 123 RVVTADQGEALAKELGIPFIESSAKNDDNVN 153
```

```
>gi|7496249|pir||T15546  hypothetical protein C18A3.6 - Caenorhabditis elegans
          Length = 343

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 14/26 (53%), Positives = 21/26 (79%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
            +RVVS  +G +LA+QL +++ ETSAK
Sbjct: 207 ERVVSMDRGRQLADQLGLEFFETSAK 232



>gi|20071543|gb|AAH26915.1|  (BC026915) Similar to RAB6, member RAS oncogene family [Mus
musculus]
          Length = 184

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
            +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 109 KRQVSIEEGERKAKELNVMFIETSAKAGYNV 139



>gi|17535675|ref|NP_495129.1|   (NM_062728) synaptic vesicle-associated GTP-binding
protein [Caenorhabditis elegans]
 gi|1619869|gb|AAB16980.1|  (U68265) RAB-3 [Caenorhabditis elegans]
 gi|1619871|gb|AAB16981.1|  (U68266) RAB-3 [Caenorhabditis elegans]
 gi|14573838|gb|AAK68196.1|U28944_10  (U28944) C. elegans RAB-3 protein (corresponding
sequence C18A3.6b) [Caenorhabditis elegans]
          Length = 219

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 14/26 (53%), Positives = 21/26 (79%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
            +RVVS  +G +LA+QL +++ ETSAK
Sbjct: 142 ERVVSMDRGRQLADQLGLEFFETSAK 167



>gi|17535677|ref|NP_495128.1|   (NM_062727) synaptic vesicle-associated GTP-binding
protein [Caenorhabditis elegans]
 gi|14573837|gb|AAK68195.1|U28944_9  (U28944) C. elegans RAB-3 protein (corresponding
sequence C18A3.6a) [Caenorhabditis elegans]
          Length = 233
 Score = 33.5 bits (75), Expect = 0.59
 Identities = 14/26 (53%), Positives = 21/26 (79%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
            +RVVS  +G +LA+QL +++ ETSAK
Sbjct: 156 ERVVSMDRGRQLADQLGLEFFETSAK 181
```

```
>gi|12381851|emb|CAC24717.1|  (AJ304407) Sec4p [Pichia pastoris]
          Length = 204

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 14/36 (38%), Positives = 26/36 (71%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVNVAWVQ 112
           R V+++QGE+LA++L V ++E SAK    V+  +++
Sbjct: 135 RQVTKEQGEQLASELGVPFLEASAKSNKNVDAIFLE 170



>gi|13195674|ref|NP_077249.1|  (NM_024287) RAB6, member RAS oncogene family [Mus
musculus]
 gi|13638404|sp|P35279|RB6A_MOUSE  Ras-related protein Rab-6A
 gi|7670415|dbj|BAA95059.1|  (AB041575) unnamed protein product [Mus musculus]
          Length = 208

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 133 KRQVSIEEGERKAKELNVMFIETSAKAGYNV 163



>gi|17458313|ref|XP_035122.2|  (XM_035122) RAB6A, member RAS oncogene family [Homo
sapiens]
 gi|131796|sp|P20340|RB6A_HUMAN  Ras-related protein Rab-6A
 gi|106190|pir||G34323  GTP-binding protein Rab6 - human
 gi|550072|gb|AAA60246.1|  (M28212) GTP-binding protein [Homo sapiens]
 gi|4585565|gb|AAD25535.1|AF130986_1  (AF130986) RAS-related protein RAB6 [Homo sapiens]
 gi|20379054|gb|AAM21087.1|AF498939_1  (AF498939) small GTP binding protein RAB6A [Homo
sapiens]
          Length = 208

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 133 KRQVSIEEGERKAKELNVMFIETSAKAGYNV 163



>gi|17512290|gb|AAH19118.1|AAH19118  (BC019118) Similar to RAB6C, member RAS oncogene family
[Mus musculus]
          Length = 208

 Score = 33.5 bits (75), Expect = 0.59
```

```
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 133 KRQVSIEEGERKAKELNVMFIETSAKAGYNV 163



>gi|19923231|ref|NP_002860.2| ☐ (NM_002869) RAB6A, member RAS oncogene family; Oncogene
RAB6; RAB6, member RAS oncogene family; Rab GTPase [Homo sapiens]
 gi|4680635|gb|AAD27707.1|AF130122_1 ☐ (AF130122) small GTP binding protein RAB6 isoform
[Homo sapiens]
 gi|6684463|gb|AAF23593.1|AF119836_1 ☐ (AF119836) GTP-binding protein RAB6C [Homo sapiens]
 gi|8163754|gb|AAF73841.1|AF198616_1 ☐ (AF198616) Rab GTPase RAB6A' [Homo sapiens]
 gi|13177664|gb|AAH03617.1|AAH03617 ☐ (BC003617) RAB6, member RAS oncogene family [Homo
sapiens]
 gi|16552748|dbj|BAB71371.1|  (AK057157) unnamed protein product [Homo sapiens]
 gi|20379058|gb|AAM21089.1|AF498941_1  (AF498941) small GTP binding protein RAB6C [Homo
sapiens]
          Length = 208

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 133 KRQVSIEEGERKAKELNVMFIETSAKAGYNV 163



>gi|2959746|emb|CAA12071.1|  (AJ224685) putative SEC4 protein [Candida glabrata]
          Length = 215

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 16/31 (51%), Positives = 22/31 (70%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           RVV+ +QGE LA +L + + +IE+SAK    VN
Sbjct: 140 RVVTYEQGEALAKELGLPFIESSAKNDDNVN 170



>gi|6685834|sp|Q9WVB1|RB6A_RAT  Ras-related protein Rab-6A
 gi|5020094|gb|AAD38018.1|AF148210_1 ☐ (AF148210) GTPase Rab6 [Rattus norvegicus]
          Length = 140

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 17/31 (54%), Positives = 22/31 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS ++GE  A +LNV +IETSAK G  V
Sbjct: 65  KRQVSIEEGERKAKELNVMFIETSAKAGYNV 95
```

```
>gi|14318517|ref|NP_116650.1|  (NC_001138) Involved in transport and fusion of post-Golgi
secretory vesicles to the plasma membrane; Sec4p [Saccharomyces cerevisiae]
 gi|134394|sp|P07560|SEC4_YEAST  Ras-related protein SEC4
 gi|68957|pir||TVBYQ4  GTP-binding protein SEC4 - yeast (Saccharomyces cerevisiae)
 gi|172566|gb|AAA35032.1|  (M16507) ras-like protein [Saccharomyces cerevisiae]
 gi|836749|dbj|BAA09233.1|  (D50617) Ras-related protein [Saccharomyces cerevisiae]
          Length = 215

 Score = 33.5 bits (75), Expect = 0.59
 Identities = 16/31 (51%), Positives = 21/31 (67%)
 Frame = +2

Query: 5   RVVSQQQGEELANQLNVDYIETSAKLGPTVN 97
           RVV+  QGE LA +L + +IE+SAK    VN
Sbjct: 140 RVVTADQGEALAKELGIPFIESSAKNDDNVN 170



>gi|17553774|ref|NP_498993.1|  (NM_066592) Rab6 (ras protein) [Caenorhabditis elegans]
 gi|464527|sp|P34213|RAB6_CAEEL  Ras-related protein Rab-6 homolog F59B2.7
 gi|345349|pir||S31127  GTP-binding protein F59B2.7 - Caenorhabditis elegans
 gi|3877850|emb|CAA77590.1|  (Z11505) predicted using Genefinder~sequence similarity to ras-
related gtp-binding proteins, contains similarity to Pfam domain: PF00071 (Ras family),
Score=320.1, E-value=8.2e-93, N=1~cDNA EST yk681g4.3 comes from this gene~cDNA EST yk681g4.5
comes fr>
          Length = 205

 Score = 33.1 bits (74), Expect = 0.77
 Identities = 17/31 (54%), Positives = 21/31 (66%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTV 94
           +R VS + GE+ A  LNV +IETSAK G  V
Sbjct: 131 KRQVSTEDGEKKARDLNVMFIETSAKAGYNV 161



>gi|2342660|gb|AAB67800.1|  (AF014120) GTP-binding protein sprab3 [Strongylocentrotus
purpuratus]
          Length = 213

 Score = 33.1 bits (74), Expect = 0.77
 Identities = 13/26 (50%), Positives = 22/26 (84%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAK 79
           +RV+S  +G++LA+QL +++ ETSAK
Sbjct: 136 ERVISTDKGKQLADQLGLEFFETSAK 161



>gi|68942|pir||TVBYSR  transforming protein ras - fission yeast  (Schizosaccharomyces pombe)
(strain JY282)
          Length = 219
```

```
 Score = 32.7 bits (73), Expect = 1.0
 Identities = 16/35 (45%), Positives = 25/35 (70%)
 Frame = +2

Query: 2   QRVVSQQQGEELANQLNVDYIETSAKLGPTVNVAW 106
           +RVVS+++ E+LA  ++  Y+ETSAKL   V  A+
Sbjct: 128 ERVVSRREREQLAKSMHCLYVETSAKLRLNVEEAF 162




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 8, 2002  5:26 AM
  Number of letters in database: 292,803,047
  Number of sequences in database:  933,889

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 87,947,890
Number of Sequences: 933889
Number of extensions: 1117463
Number of successful extensions: 5943
Number of sequences better than 10.0: 294
Number of HSP's better than 10.0 without gapping: 5277
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 5926
length of database: 292,803,047
effective HSP length: 25
effective length of database: 269,455,822
effective search space used: 6466939728
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

RID: 1020950074-06174-28797

7.1.2 Query= hy-2_r43a30007

(297 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
933,889 sequences; 292,803,047 total letters

If you have any problems or questions with the results of this search
please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 150 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|17565854|ref|NP_506559.1|  (NM_074158) ATP synthase alpha...   89   1e-17
gi|7592738|dbj|BAA94379.1|  (AB033377) vacuolar H+-ATPase su...   85   1e-16
gi|3169287|gb|AAC17840.1|  (AF050673) vacuolar H+-ATPase cat...   84   4e-16
gi|12585490|sp|Q9SM09|VATA_CITUN  Vacuolar ATP synthase cata...   84   4e-16
gi|15982954|gb|AAL11505.1|AF367445_1  (AF367445) V-ATPase ca...   84   4e-16
gi|401322|sp|P31405|VATA_GOSHI  Vacuolar ATP synthase cataly...   84   4e-16
gi|1041768|gb|AAA79992.1|  (U36438) V-ATPase 66 kDa subunit ...   84   4e-16
gi|13235340|emb|CAC33578.1|  (AJ276326) putative vacuolar AT...   84   4e-16
gi|6721571|dbj|BAA89598.1|  (AB036926) vacuolar H+-ATPase A ...   84   4e-16
gi|14585885|gb|AAK67706.1|  (AY037143) V-type H+-ATPase subu...   84   4e-16
gi|14149011|emb|CAC39166.1|  (AJ278021) putative V-type H+-A...   84   4e-16
gi|1352830|sp|P49087|VATA_MAIZE  Vacuolar ATP synthase catal...   83   5e-16
gi|2493122|sp|Q39291|VATA_BRANA  Vacuolar ATP synthase catal...   83   7e-16
gi|137460|sp|P09469|VATA_DAUCA  Vacuolar ATP synthase cataly...   83   7e-16
gi|15219234|ref|NP_178011.1|  (NM_106539) ATPase 70 kDa subu...   83   7e-16
gi|11527563|dbj|BAB18682.1|  (AB032840) vacuolar proton-ATPa...   82   9e-16
gi|2493123|sp|Q40002|VATA_HORVU  Vacuolar ATP synthase catal...   82   9e-16
gi|1912395|gb|AAB50981.1|  (U49169) vacuolar H+-ATPase A sub...   82   1e-15
gi|7297989|gb|AAF53231.1|  (AE003638) CG3762 gene product [a...   82   1e-15
gi|19527547|gb|AAL89888.1|  (AY084150) RE30552p [Drosophila ...   82   1e-15
gi|1352828|sp|P48414|VATA_CYACA  Vacuolar ATP synthase catal...   82   1e-15
gi|14915706|dbj|BAB62103.1|  (AB066243) V-ATPase subunit A [...   82   1e-15
gi|3915247|sp|O16109|VATA_AEDAE  Vacuolar ATP synthase catal...   82   1e-15
```

```
gi|2506211|sp|P13548|VATA_PHAAU  Vacuolar ATP synthase catal...   82   2e-15
gi|12643368|sp|Q27331|VAA2_DROME  Vacuolar ATP synthase cata...   82   2e-15
gi|2118216|pir||I50716  A1 isoform of vacuolar H+-ATPase sub...   82   2e-15
gi|2118217|pir||I50715  A2 isoform of vacuolar H+-ATPase sub...   82   2e-15
gi|418177|sp|Q03498|VATA_PLAFA  Vacuolar ATP synthase cataly...   81   2e-15
gi|401323|sp|P31400|VATA_MANSE  Vacuolar ATP synthase cataly...   81   3e-15
gi|4502307|ref|NP_001682.1|  (NM_001691) ATPase, H+ transpor...   81   3e-15
gi|18859619|ref|NP_523560.1|  (NM_078836) vacuolar H+-ATPase...   80   4e-15
gi|12643429|sp|P48602|VAA1_DROME  Vacuolar ATP synthase cata...   80   4e-15
gi|108733|pir||S19659  H+-transporting ATPase (EC 3.6.1.35) ...   80   5e-15
gi|383059|prf||1902186A  vacuolar H ATPase:SUBUNIT=70kD [Bos...   80   5e-15
gi|17443303|ref|XP_056371.2|  (XM_056371) similar to vacuola...   80   5e-15
gi|6523821|gb|AAF14870.1|AF113129_1  (AF113129) vacuolar ATP...   80   5e-15
gi|401321|sp|P31404|VAA1_BOVIN  Vacuolar ATP synthase cataly...   80   5e-15
gi|1082432|pir||B46091  H+-transporting ATPase (EC 3.6.1.35)...   80   5e-15
gi|6680752|ref|NP_031534.1|  (NM_007508) ATPase, H+ transpor...   80   5e-15
gi|586210|sp|P38606|VAA1_HUMAN  Vacuolar ATP synthase cataly...   80   5e-15
gi|1082966|pir||A56807  H+-transporting ATP synthase (EC 3.6...   80   6e-15
gi|466358|gb|AAA21531.1|  (U04849) putative vacuolar proton-...   80   6e-15
gi|12231040|sp|Q29048|VAA1_PIG  Vacuolar ATP synthase cataly...   80   6e-15
gi|5869971|emb|CAB55557.1|  (AJ131823) H(+)-transporting ATP...   79   1e-14
gi|3334408|sp|Q38676|VAA1_ACEAT  Vacuolar ATP synthase catal...   79   1e-14
gi|2493120|sp|Q38677|VAA2_ACEAT  Vacuolar ATP synthase catal...   79   1e-14
gi|14600685|ref|NP_147205.1|  (NC_000854) membrane-associate...   79   1e-14
gi|108305|pir||S18887  H+-transporting ATPase (EC 3.6.1.35) ...   78   2e-14
gi|19114337|ref|NP_593425.1|  (NC_003424) V-type ATPase; vac...   78   2e-14
gi|4126679|dbj|BAA36691.1|  (AB016483) vacuolar-type H+-ATPa...   78   2e-14
gi|20094453|ref|NP_614300.1|  (NC_003551) Archaeal/vacuolar-...   78   2e-14
gi|2493121|sp|Q39442|VATA_BETVU  Vacuolar ATP synthase catal...   78   2e-14
gi|20129479|ref|NP_609595.1|  (NM_135751) CG5075 gene produc...   77   4e-14
gi|19527911|gb|AAL90070.1|  (AY089332) AT13860p [Drosophila ...   77   4e-14
gi|586209|sp|P38078|VATA_CANTR  Vacuolar ATP synthase cataly...   76   7e-14
gi|19074928|ref|NP_586434.1|  (NC_003237) VACUOLAR ATP SYNTH...   76   7e-14
gi|137461|sp|P11592|VATA_NEUCR  Vacuolar ATP synthase cataly...   76   9e-14
gi|172907|gb|AAB63978.1|  (M21609) transmembrane ATPase-like...   75   1e-13
gi|12585499|sp|Q9UVJ8|VATA_ASHGO  Vacuolar ATP synthase cata...   75   1e-13
gi|2493119|sp|Q26975|VATA_TRYCO  Vacuolar ATP synthase catal...   75   1e-13
gi|16444949|dbj|BAB70682.1|  (AB073302) vacuolar membrane AT...   75   2e-13
gi|12585409|sp|O32466|VATA_THESI  V-type ATP synthase alpha ...   74   3e-13
gi|6320016|ref|NP_010096.1|  (NC_001136) Encodes a protein w...   74   3e-13
gi|3417405|emb|CAA98761.1|  (Z74233) ORF YDL184w-b [Saccharo...   74   3e-13
gi|16417186|gb|AAL18608.1|  (AF389404) VMA1 [Saccharomyces s...   74   3e-13
gi|14591712|ref|NP_143800.1|  (NC_000961) H(+)-transporting ...   73   7e-13
gi|15674362|ref|NP_268536.1|  (NC_002737) putative V-type Na...   73   7e-13
gi|14521960|ref|NP_127437.1|  (NC_000868) H+-transporting AT...   73   7e-13
gi|7436320|pir||JC5532  vacuolar-type ATPase (EC 3.-.-.-) A ...   72   1e-12
gi|12585391|sp|O06504|VATA_DESSY  V-type ATP synthase alpha ...   72   1e-12
gi|15678973|ref|NP_276090.1|  (NC_000916) ATP synthase, subu...   72   2e-12
gi|19745323|ref|NP_606459.1|  (NC_003485) putative V-type Na...   72   2e-12
gi|18976554|ref|NP_577911.1|  (NC_003413) ATPase subunit A [...   72   2e-12
gi|80953|pir||S13589  H+-transporting ATP synthase (EC 3.6.1...   70   4e-12
gi|15901171|ref|NP_345775.1|  (NC_003028) v-type sodium ATP ...   70   6e-12
gi|1171780|sp|Q08636|NTPA_ENTHR  V-type sodium ATP synthase ...   68   2e-11
gi|18312083|ref|NP_558750.1|  (NC_003364) H+-transporting AT...   67   3e-11
gi|9229839|dbj|BAB00608.1|  (AB045977) A-ATPase A-subunit [T...   67   3e-11
gi|16081190|ref|NP_393482.1|  (NC_002578) probable ATP synth...   67   3e-11
gi|19705056|ref|NP_602551.1|  (NC_003454) V-type sodium ATP ...   67   4e-11
gi|13540883|ref|NP_110571.1|  (NC_002689) Vacuolar-type H+-A...   67   4e-11
gi|114528|sp|P09639|VATA_SULAC  V-type ATP synthase alpha ch...   66   7e-11
gi|15921725|ref|NP_377394.1|  (NC_003106) 595aa long membran...   66   7e-11
gi|15897484|ref|NP_342089.1|  (NC_002754) ATP synthase subun...   66   9e-11
```

```
gi|1777421|gb|AAB40515.1|  (U18938) vacuolar-ATPase catalyti...    65   1e-10
gi|15668390|ref|NP_247186.1|  (NC_000909) H+-transporting AT...    65   2e-10
gi|12585563|sp|Q57670|VATA_METJA  V-type ATP synthase alpha ...    65   2e-10
gi|11498766|ref|NP_069995.1|  (NC_000917) H+-transporting AT...    65   2e-10
gi|18310620|ref|NP_562554.1|  (NC_003366) V-type sodium ATP ...    64   3e-10
gi|12585457|sp|Q56403|VATA_THETH  V-type ATP synthase alpha ...    64   3e-10
gi|20092951|ref|NP_619026.1|  (NC_003552) H(+)-transporting ...    63   8e-10
gi|114520|sp|P22662|VATA_METBA  V-type ATP synthase alpha ch...    62   1e-09
gi|15805727|ref|NP_294423.1|  (NC_001263) v-type ATP synthas...    61   2e-09
gi|2493100|sp|Q60186|VATA_METMA  V-type ATP synthase alpha c...    60   5e-09
gi|2493099|sp|Q48332|VATA_HALVO  V-type ATP synthase alpha c...    59   1e-08
gi|15790973|ref|NP_280797.1|  (NC_002607) H+-transporting AT...    58   2e-08
gi|114516|sp|P25163|VATA_HALSA  V-type ATP synthase alpha ch...    58   2e-08
gi|12229704|sp|Q9HNE3|VATA_HALN1  V-type ATP synthase alpha ...    58   2e-08
gi|15835199|ref|NP_296958.1|  (NC_002620) ATP synthase, subu...    58   2e-08
gi|15639519|ref|NP_218969.1|  (NC_000919) V-type ATPase, sub...    57   3e-08

                                  Alignments


>gi|17565854|ref|NP_506559.1|   (NM_074158) ATP synthase alpha and beta subunits ; ATP
synthase ab C terminal [Caenorhabditis elegans]
 gi|7510027|pir||T27035  hypothetical protein Y49A3A.2 - Caenorhabditis elegans
 gi|4008438|emb|CAA22076.1|  (AL033512) predicted using Genefinder~contains similarity to
Pfam domain: PF00006 (ATP synthase alpha/beta family), Score=739.2, E-value=5.6e-219, N=1;
PF00306 (ATP synthase ab C terminal), Score=205.7, E-value=2.3e-58, N=1~cDNA EST yk4d5.3
comes from >
          Length = 606

 Score = 88.6 bits (218), Expect = 1e-17
 Identities = 40/80 (50%), Positives = 56/80 (70%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMGLNV+ + DST++W++A+ +I  +LGE+         L A++  FY RAGRV+CL
Sbjct: 325 YFRDMGLNVAMMADSTSRWAEALREISGRLGEMPADSGYPAYLAARLASFYERAGRVKCL 384

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSPER GS+T+VG  +   G
Sbjct: 385 GSPEREGSVTIVGAVSPPGG 404



>gi|7592738|dbj|BAA94379.1|  (AB033377) vacuolar H+-ATPase subunit A [Nepenthes alata]
          Length = 101

 Score = 85.1 bits (209), Expect = 1e-16
 Identities = 38/80 (47%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 17  YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 76

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSPER GS+T+VG  +   G
Sbjct: 77  GSPERTGSVTIVGAVSPPGG 96
```

```
>gi|3169287|gb|AAC17840.1|  (AF050673) vacuolar H+-ATPase catalytic subunit [Gossypium
hirsutum]
          Length = 623

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418




>gi|12585490|sp|Q9SM09|VATA_CITUN  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|6518112|dbj|BAA87891.1|  (AB004247) H+-ATPase catalytic subunit [Citrus unshiu]
          Length = 623

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418




>gi|15982954|gb|AAL11505.1|AF367445_1  (AF367445) V-ATPase catalytic subunit A [Prunus
persica]
          Length = 623

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2



Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418
```

```
>gi|401322|sp|P31405|VATA_GOSHI  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|167313|gb|AAA33050.1|  (L03186) vacuolar H+-ATPase catalytic subunit [Gossypium
hirsutum]
          Length = 623

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418



>gi|1041768|gb|AAA79992.1|  (U36438) V-ATPase 66 kDa subunit [Acer pseudoplatanus]
          Length = 336

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 235 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 294

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 295 GGPERTGSVTIVGAVSPPGG 314



>gi|13235340|emb|CAC33578.1|  (AJ276326) putative vacuolar ATP Synthase subunit A
[Mesembryanthemum crystallinum]
          Length = 624

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418
```

```
>gi|6721571|dbj|BAA89598.1|  (AB036926) vacuolar H+-ATPase A subunit [Citrus unshiu]
          Length = 317

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 33  YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 92

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 93  GGPERTGSVTIVGAVSPPGG 112



>gi|14585885|gb|AAK67706.1|  (AY037143) V-type H+-ATPase subunit A [Pisum sativum]
          Length = 151

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 36  YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 95

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 96  GGPERTGSVTIVGAVSPPGG 115



>gi|14149011|emb|CAC39166.1|  (AJ278021) putative V-type H+-ATPase catalytic subunit
[Lycopersicon esculentum]
          Length = 259

 Score = 83.6 bits (205), Expect = 4e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 51  YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 110

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 111 GGPERTGSVTIVGAVSPPGG 130



>gi|1352830|sp|P49087|VATA_MAIZE  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|1049253|gb|AAA80346.1|  (U36436) vacuolar ATPase 69 kDa subunit [Zea mays]
          Length = 561
```

```
 Score = 83.2 bits (204), Expect = 5e-16
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 277 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 336

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP+R GS+T+VG  +   G
Sbjct: 337 GSPDRNGSVTIVGAVSPPGG 356



>gi|2493122|sp|Q39291|VATA_BRANA  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit) (Tonoplast ATPase 70
kDa subunit) (BN59)
 gi|1362024|pir||S57790  H+-transporting ATPase (EC 3.6.1.35) 70K chain, vacuolar  (clone
BN59) - rape
 gi|558479|gb|AAA82881.1|  (U15604) tonoplast ATPase 70 kDa subunit [Brassica napus]
          Length = 623

 Score = 82.8 bits (203), Expect = 7e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERNGSVTIVGAVSPPGG 418



>gi|137460|sp|P09469|VATA_DAUCA  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|67952|pir||PXPZV9  H+-transporting ATPase (EC 3.6.1.35), vacuolar, 69K chain - carrot
 gi|167560|gb|AAA33139.1|  (J03769) vacular H+-ATPase [Daucus carota]
          Length = 623

 Score = 82.8 bits (203), Expect = 7e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398



Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERNGSVTIVGAVSPPGG 418
```

```
>gi|15219234|ref|NP_178011.1|  (NM_106539) ATPase 70 kDa subunit, putative [Arabidopsis
thaliana]
 gi|3334404|sp|O23654|VATA_ARATH  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|2266990|gb|AAB97128.1|  (U65638) vacuolar type ATPase subunit A [Arabidopsis thaliana]
 gi|3834305|gb|AAC83021.1|  (AC005679) Identical to gb|U65638 Arabidopsis thaliana vacuolar
type ATPase subunit A mRNA.  ESTs gb|N96435, gb|N96106, gb|N96189, gb|N96091, gb|AA042286,
gb|F14324, gb|W43643, gb|N96027, gb|N96299, gb|R29943, gb|T43460, gb|T43544, gb|T22472,
gb|T14078>
 gi|16649079|gb|AAL24391.1|  (AY059909) vacuolar type ATPase subunit A [Arabidopsis
thaliana]
 gi|19698899|gb|AAL91185.1|  (AY081296) ATPase subunit A [Arabidopsis thaliana]
          Length = 623

 Score = 82.8 bits (203), Expect = 7e-16
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAG+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERNGSVTIVGAVSPPGG 418




>gi|11527563|dbj|BAB18682.1|  (AB032840) vacuolar proton-ATPase [Hordeum vulgare subsp.
vulgare]
          Length = 621

 Score = 82.4 bits (202), Expect = 9e-16
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L +++  FY RAG+V+CL
Sbjct: 337 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLASRLASFYERAGKVQCL 396

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP+R GS+T+VG  +   G
Sbjct: 397 GSPDRTGSVTIVGAVSPPGG 416




>gi|2493123|sp|Q40002|VATA_HORVU  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|7436321|pir||T04409  probable H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar -
barley (fragment)
 gi|1051258|gb|AAB60306.1|  (U36939) vacuolar ATPase catalytic subunit [Hordeum vulgare
subsp. vulgare]
          Length = 580

 Score = 82.4 bits (202), Expect = 9e-16
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2
```

```
Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L +++  FY RAG+V+CL
Sbjct: 296 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLASRLASFYERAGKVQCL 355

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP+R GS+T+VG  +   G
Sbjct: 356 GSPDRTGSVTIVGAVSPPGG 375



>gi|1912395|gb|AAB50981.1|  (U49169) vacuolar H+-ATPase A subunit [Dictyostelium discoideum]
          Length = 618

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMGLNV+ + DST++W++A+ +I  +L E+         L A++  FY RAGRV C+
Sbjct: 337 YFRDMGLNVAMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVSCI 396

Query: 182 GSPERIGSITMVGIATQSEG 241
           G P RIGS+T+VG  +   G
Sbjct: 397 GHPTRIGSVTIVGAVSPPGG 416



>gi|7297989|gb|AAF53231.1|  (AE003638) CG3762 gene product [alt 1] [Drosophila
melanogaster]
 gi|7297990|gb|AAF53232.1|  (AE003638) CG3762 gene product [alt 2] [Drosophila
melanogaster]
 gi|7297991|gb|AAF53233.1|  (AE003638) CG3762 gene product [alt 3] [Drosophila
melanogaster]
          Length = 614

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 333 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 392

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 393 GNPEREGSVSIVGAVSPPGG 412



>gi|19527547|gb|AAL89888.1|  (AY084150) RE30552p [Drosophila melanogaster]
          Length = 614

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2
```

```
Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+          L A++  FY RAGRV+CL
Sbjct: 333 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 392

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 393 GNPEREGSVSIVGAVSPPGG 412



>gi|1352828|sp|P48414|VATA_CYACA  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|7436322|pir||T14360  H+-transporting ATPase (EC 3.6.1.35) chain A - red alga  (Cyanidium
caldarium)
 gi|576659|gb|AAA85820.1|  (U17100) V-ATPase A subunit [Cyanidium caldarium]
          Length = 587

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           Y+RDMGLN+S + DST++W++A+ +I  +L E+          L A++  FY RAG+V CL
Sbjct: 329 YYRDMGLNISMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGKVSCL 388

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP R GSIT+VG  +   G
Sbjct: 389 GSPNRQGSITIVGAVSPPGG 408



>gi|14915706|dbj|BAB62103.1|  (AB066243) V-ATPase subunit A [Fundulus heteroclitus]
          Length = 617

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+          L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415



>gi|3915247|sp|O16109|VATA_AEDAE  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|2454488|gb|AAB71659.1|  (AF008922) V-ATPase A-subunit [Aedes aegypti]
          Length = 615

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2
```

```
Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I   +L E+          L A++   FY RAGRV+CL
Sbjct: 334 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 393

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 394 GNPEREGSVSIVGAVSPPGG 413



>gi|2506211|sp|P13548|VATA_PHAAU  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit) (VAA3-1)
 gi|849136|gb|AAC49174.1|  (U26709) vacuolar H+-ATPase subunit A [Vigna radiata]
          Length = 623

 Score = 81.6 bits (200), Expect = 2e-15
 Identities = 36/80 (45%), Positives = 52/80 (65%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I   +L E+          L A++   FY R G+V+CL
Sbjct: 339 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERPGKVKCL 398

Query: 182 GSPERIGSITMVGIATQSEG 241
           G PER GS+T+VG  +   G
Sbjct: 399 GGPERTGSVTIVGAVSPPGG 418



>gi|12643368|sp|Q27331|VAA2_DROME  Vacuolar ATP synthase catalytic subunit A isoform 2 (V-
ATPase A subunit 2) (Vacuolar proton pump alpha subunit 2) (V-ATPase 69 kDa subunit 2)
 gi|1373433|gb|AAB02270.1| ☐ (U59146) vacuolar ATPase subunit A [Drosophila melanogaster]
 gi|1373435|gb|AAB02271.1| ☐ (U59147) vacuolar ATPase subunit A [Drosophila melanogaster]
          Length = 614

 Score = 81.6 bits (200), Expect = 2e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I   +L E+          L A++   FY RAGRV+CL
Sbjct: 333 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPRDSGYPAYLGARLASFYERAGRVKCL 392

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 393 GNPEREGSVSIVGAVSPPGG 412



>gi|2118216|pir||I50716  A1 isoform of vacuolar H+-ATPase subunit A - chicken
 gi|790617|gb|AAC59680.1|  (U22077) A1 isoform of vacuolar H+-ATPase subunit A [Gallus
gallus]
          Length = 617

 Score = 81.6 bits (200), Expect = 2e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2
```

```
Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+T+VG  +   G
Sbjct: 396 GNPEREGSVTIVGAVSPPGG 415



>gi|2118217|pir||I50715  A2 isoform of vacuolar H+-ATPase subunit A - chicken
 gi|790615|gb|AAC59679.1|  (U22076) A2 isoform of vacuolar H+-ATPase subunit A [Gallus
gallus]
 gi|1098356|prf||2115408A  vacuolar H ATPase:SUBUNIT=A:ISOTYPE=A2 [Gallus gallus]
          Length = 611

 Score = 81.6 bits (200), Expect = 2e-15
 Identities = 37/80 (46%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 330 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 389

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+T+VG  +   G
Sbjct: 390 GNPEREGSVTIVGAVSPPGG 409



>gi|418177|sp|Q03498|VATA_PLAFA  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|477307|pir||A48582  vacuolar ATPase A subunit homolog - malaria parasite  (Plasmodium
falciparum)
 gi|160736|gb|AAA29782.1|  (L08200) vacuolar ATPase [Plasmodium falciparum 3D7]
          Length = 611

 Score = 81.3 bits (199), Expect = 2e-15
 Identities = 36/80 (45%), Positives = 53/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG N + + DST++W++A+ +I  +L E+         L A++  FY RAG+V+C+
Sbjct: 330 YFRDMGYNATMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGKVKCI 389

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP RIGSIT+VG  +   G
Sbjct: 390 GSPSRIGSITIVGAVSPPGG 409



>gi|401323|sp|P31400|VATA_MANSE  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 69 kDa subunit)
 gi|102967|pir||S21107  H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar - tobacco
hornworm
 gi|11062|emb|CAA45537.1|  (X64233) H(+)-transporting  ATPase [Manduca sexta]
          Length = 617
```

```
 Score = 80.9 bits (198), Expect = 3e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
            YFRDMG NVS + DST++W++A+ +I  +L E+          L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
            G+P+R GS+++VG  +   G
Sbjct: 396 GNPDREGSVSIVGAVSPPGG 415




>gi|4502307|ref|NP_001682.1| ☐ (NM_001691) ATPase, H+ transporting, lysosomal, subunit A2;
vacuolar ATP synthase catalytic subunit A, osteoclast isoform; V-ATPase A subunit 2;
vacuolar proton pump alpha subunit 2; V-ATPase 69 kDa subunit 2; H(+)-transporting two-
sector ATPase, subun>
 gi|586208|sp|P38607|VAA2_HUMAN  Vacuolar ATP synthase catalytic subunit A, osteoclast
isoform (V-ATPase A subunit 2) (Vacuolar proton pump alpha subunit 2) (V-ATPase 69 kDa
subunit 2) (Isoform H068)
 gi|1082431|pir||A46091 ☐ H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar (HO68
type) - human
 gi|291866|gb|AAA35578.1| ☐ (L09234) ATPase [Homo sapiens]
          Length = 615

 Score = 80.9 bits (198), Expect = 3e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
            YFRDMG NVS + DST++W++A+ +I  +L E+          L A++  FY RAGRV+CL
Sbjct: 334 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 393

Query: 182 GSPERIGSITMVGIATQSEG 241
            G+P+R GS+++VG  +   G
Sbjct: 394 GNPDREGSVSIVGAVSPPGG 413



>gi|18859619|ref|NP_523560.1| ☐ (NM_078836) vacuolar H+-ATPase 68kD A subunit; Vacuolar
H[+]-ATPase A subunit; Vacuolar H[+]-ATPase 68kD A subunit
[Drosophila melanogaster]
 gi|641918|gb|AAA61760.1| ☐ (U19742) vacuolar ATPase subunit A [Drosophila melanogaster]
          Length = 614

 Score = 80.5 bits (197), Expect = 4e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
            YFRDMG NV+ + DST++W++A+ +I  +L E+          L A++  FY RAGRV+CL
Sbjct: 333 YFRDMGYNVAMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLATFYERAGRVKCL 392


Query: 182 GSPERIGSITMVGIATQSEG 241
            G+PER GS+++VG  +   G
Sbjct: 393 GNPEREGSVSIVGAVSPPGG 412
```

```
>gi|12643429|sp|P48602|VAA1_DROME  Vacuolar ATP synthase catalytic subunit A isoform 1 (V-
ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa subunit 1)
 gi|5851683|gb|AAA61761.2|   (U19745) V-ATPase A subunit [Drosophila melanogaster]
 gi|6003626|gb|AAF00515.1|AF185049_1   (AF185049) vacuolar ATPase subunit A [Drosophila
melanogaster]
 gi|7297994|gb|AAF53236.1|   (AE003638) Vha68-1 gene product [Drosophila melanogaster]
 gi|19528271|gb|AAL90250.1|   (AY089512) GH21132p [Drosophila melanogaster]
          Length = 614

 Score = 80.5 bits (197), Expect = 4e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NV+ + DST++W++A+ +I   +L E+          L A++   FY RAGRV+CL
Sbjct: 333 YFRDMGYNVAMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLATFYERAGRVKCL 392

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG   +   G
Sbjct: 393 GNPEREGSVSIVGAVSPPGG 412



>gi|108733|pir||S19659  H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar - bovine
 gi|163|emb|CAA41276.1|  (X58386) H(+)-transporting  ATPase; H(+)-ATPase subunit A [Bos
taurus]
 gi|228275|prf||1802274A  vacuolar H ATPase:SUBUNIT=A [Bos taurus]
          Length = 618

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I   +L E+          L A++   FY RAGRV+CL
Sbjct: 337 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 396

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG   +   G
Sbjct: 397 GNPEREGSVSIVGAVSPPGG 416



>gi|383059|prf||1902186A  vacuolar H ATPase:SUBUNIT=70kD [Bos taurus]
          Length = 618

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I   +L E+          L A++   FY RAGRV+CL
Sbjct: 337 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 396
```

```
Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 397 GNPEREGSVSIVGAVSPPGG 416




>gi|17443303|ref|XP_056371.2|  ☐ (XM_056371) similar to vacuolar ATPase isoform VA68 [Homo
sapiens]
 gi|19913424|ref|NP_001681.2|  ☐ (NM_001690) ATPase, H+ transporting, lysosomal, subunit
A1; H+-transporting ATPase chain A, vacuolar (VA68 type); vacuolar ATP synthase catalytic
subunit A, ubiquitous isoform; V-ATPase A subunit 1; vacuolar proton pump alpha subunit 1;
V-ATPase 69 kDa>
 gi|15341906|gb|AAH13138.1|AAH13138  ☐ (BC013138) Unknown (protein for MGC:21189) [Homo
sapiens]
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I   +L E+           L A++   FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415




>gi|6523821|gb|AAF14870.1|AF113129_1  ☐ (AF113129) vacuolar ATPase isoform VA68 [Homo
sapiens]
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I   +L E+           L A++   FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG  +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415




>gi|401321|sp|P31404|VAA1_BOVIN  Vacuolar ATP synthase catalytic subunit A, ubiquitous
isoform (V-ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa
subunit 1)
 gi|162705|gb|AAA30392.1|  (M80430) vacuolar H+-ATPase A subunit [Bos taurus]
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2
```

```
Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG   +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415



>gi|1082432|pir||B46091 ☐ H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar (VA68
type) - human
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG   +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415



>gi|6680752|ref|NP_031534.1| ☐ (NM_007508) ATPase, H+ transporting, lysosomal (vacuolar
proton pump), alpha 70 kDa, isoform 1; 70-kDa subunit; ATPase, H+ transporting, lysosomal
(vacuolar proton pump), alpha 70 kDa, isoform 2 [Mus musculus]
 gi|1718086|sp|P50516|VAA1_MOUSE  Vacuolar ATP synthase catalytic subunit A, ubiquitous
isoform (V-ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa
subunit 1)
 gi|1184659|gb|AAC52410.1| ☐ (U13837) vacuolar adenosine triphosphatase subunit A [Mus
musculus]
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG   +   G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415



>gi|586210|sp|P38606|VAA1_HUMAN  Vacuolar ATP synthase catalytic subunit A, ubiquitous
isoform (V-ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa
subunit 1) (Isoform VA68)
 gi|291868|gb|AAA83249.1| ☐ (L09235) ATPase [Homo sapiens]
```

```
          Length = 617

 Score = 80.1 bits (196), Expect = 5e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + DST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+++VG +    G
Sbjct: 396 GNPEREGSVSIVGAVSPPGG 415




>gi|1082966|pir||A56807  H+-transporting ATP synthase (EC 3.6.1.34), vacuolar type,
catalytic chain - pig
 gi|1968|emb|CAA44213.1|  (X62338) H+ ATPase [Sus scrofa]
          Length = 618

 Score = 79.7 bits (195), Expect = 6e-15
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + +ST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 337 YFRDMGYHVSMMANSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 396

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+T+VG +    G
Sbjct: 397 GNPEREGSVTIVGAVSPPGG 416




>gi|466358|gb|AAA21531.1|  (U04849) putative vacuolar proton-transporting ATPase catalytic
subunit [Entamoeba histolytica]
          Length = 607

 Score = 79.7 bits (195), Expect = 6e-15
 Identities = 35/80 (43%), Positives = 53/80 (65%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           Y+RDMG NV+ + DST++W++A+ +I  +L E+         L A++  FY RAG V CL
Sbjct: 328 YYRDMGYNVAMMADSTSRWAEALREISGRLAEMPADSGYPAYLAARLASFYERAGMVECL 387

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP+RIGS+++VG +    G
Sbjct: 388 GSPKRIGSVSIVGAVSPPGG 407



>gi|12231040|sp|Q29048|VAA1_PIG  Vacuolar ATP synthase catalytic subunit A, ubiquitous
isoform (V-ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa
subunit 1)
          Length = 617

 Score = 79.7 bits (195), Expect = 6e-15
```

```
 Identities = 36/80 (45%), Positives = 54/80 (67%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + +ST++W++A+ +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 336 YFRDMGYHVSMMANSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 395

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+T+VG  +   G
Sbjct: 396 GNPEREGSVTIVGAVSPPGG 415



>gi|5869971|emb|CAB55557.1|  (AJ131823) H(+)-transporting ATP synthase [Scherffelia dubia]
          Length = 615

 Score = 79.0 bits (193), Expect = 1e-14
 Identities = 36/80 (45%), Positives = 52/80 (65%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG N S + DST++W++A+ +I  +L E+         L A++  FY R+GRV+CL
Sbjct: 330 YFRDMGYNFSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERSGRVKCL 389

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP R GS+T+VG  +   G
Sbjct: 390 GSPGREGSVTIVGAVSPPGG 409



>gi|3334408|sp|Q38676|VAA1_ACEAT  Vacuolar ATP synthase catalytic subunit A isoform 1 (V-
ATPase A subunit 1) (Vacuolar proton pump alpha subunit 1) (V-ATPase 69 kDa subunit 1)
 gi|1303673|dbj|BAA09097.1|  (D50528) adenosine triphosphatase A subunit [Acetabularia
acetabulum]
          Length = 613

 Score = 78.6 bits (192), Expect = 1e-14
 Identities = 35/80 (43%), Positives = 52/80 (64%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG N + + DST++W++A+ +I  +L E+         L A++  FY R+GRV C+
Sbjct: 327 YFRDMGYNFAMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERSGRVACI 386


Query: 182 GSPERIGSITMVGIATQSEG 241
           GSPER GS+T+VG  +   G
Sbjct: 387 GSPEREGSVTIVGAVSPPGG 406



>gi|2493120|sp|Q38677|VAA2_ACEAT  Vacuolar ATP synthase catalytic subunit A isoform 2 (V-
ATPase A subunit 2) (Vacuolar proton pump alpha subunit 2) (V-ATPase 69 kDa subunit 2)
 gi|1303679|dbj|BAA09098.1|  (D50529) adenosine triphosphatase A subunit [Acetabularia
acetabulum]
          Length = 613

 Score = 78.6 bits (192), Expect = 1e-14
```

```
 Identities = 35/80 (43%), Positives = 52/80 (64%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG N + + DST++W++A+ +I  +L E+         L A++  FY R+GRV C+
Sbjct: 327 YFRDMGYNFAMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERSGRVACI 386

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSPER GS+T+VG  +   G
Sbjct: 387 GSPEREGSVTIVGAVSPPGG 406



>gi|14600685|ref|NP_147205.1|  (NC_000854) membrane-associated ATPase alpha chain [Aeropyrum
pernix]
 gi|12585522|sp|Q9YF35|VATA_AERPE  V-type ATP synthase alpha chain (V-type ATPase subunit A)
 gi|7436327|pir||E72733  probable membrane-associated ATPase alpha chain APE0405 - Aeropyrum
pernix (strain K1)
 gi|5104045|dbj|BAA79361.1|  (AP000059) 598aa long hypothetical membrane-associated ATPase
alpha chain [Aeropyrum pernix]
          Length = 598

 Score = 78.6 bits (192), Expect = 1e-14
 Identities = 35/80 (43%), Positives = 54/80 (66%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           Y+RDMG +V  + DST++W++A+ +I  +L E+   E     L +++ +FY RAGRV+ L
Sbjct: 326 YYRDMGYDVLLVADSTSRWAEALREIAGRLEEMPAEEGYPSYLASRLAEFYERAGRVKAL 385

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSPER GS+T+VG  +   G
Sbjct: 386 GSPERSGSVTVVGAVSPPGG 405



>gi|108305|pir||S18887  H+-transporting ATPase (EC 3.6.1.35) catalytic chain, renal - pig
(fragment)
          Length = 374

 Score = 78.2 bits (191), Expect = 2e-14
 Identities = 36/80 (45%), Positives = 53/80 (66%)
 Frame = +2


Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG +VS + +ST++W++A  +I  +L E+         L A++  FY RAGRV+CL
Sbjct: 179 YFRDMGYHVSMMANSTSRWAEAHREISGRLAEMPADSGYPAYLGARLASFYERAGRVKCL 238

Query: 182 GSPERIGSITMVGIATQSEG 241
           G+PER GS+T+VG  +   G
Sbjct: 239 GNPEREGSVTIVGAVSPPGG 258
```

```
>gi|19114337|ref|NP_593425.1|  (NC_003424) V-type ATPase; vacuolar ATP synthase catalytic
subunit A [Schizosaccharomyces pombe]
 gi|401324|sp|P31406|VATA_SCHPO  Vacuolar ATP synthase catalytic subunit A (V-ATPase A
subunit) (Vacuolar proton pump alpha subunit) (V-ATPase 67 kDa subunit)
 gi|283082|pir||S25334  H+-transporting ATPase (EC 3.6.1.35) chain A, vacuolar - fission
yeast (Schizosaccharomyces pombe)
 gi|1065902|emb|CAA48573.1|  (X68580) vacuolar H+-ATPase subunit A [Schizosaccharomyces
pombe]
 gi|5706506|emb|CAB52268.1|  (AL109739) V-type ATPase; vacuolar ATP synthase catalytic
subunit A [Schizosaccharomyces pombe]
          Length = 619

 Score = 78.2 bits (191), Expect = 2e-14
 Identities = 35/80 (43%), Positives = 52/80 (64%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           Y+RD G NVS + DST++W++A+ +I  +L E+         L AK+  FY RAGR RCL
Sbjct: 344 YYRDQGKNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGAKLASFYERAGRARCL 403

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP+R G++++VG  +   G
Sbjct: 404 GSPDREGTVSIVGAVSPPGG 423




>gi|4126679|dbj|BAA36691.1|  (AB016483) vacuolar-type H+-ATPase subunit A [Ascidia
sydneiensis samea]
          Length = 619

 Score = 78.2 bits (191), Expect = 2e-14
 Identities = 36/80 (45%), Positives = 51/80 (63%)
 Frame = +2

Query: 2   YFRDMGLNVSFIVDSTNQWSQAVNQIQEKLGEILPSETQSEVLTAKIGQFYNRAGRVRCL 181
           YFRDMG NVS + DST++W++A+ +I  +L E+         L A++  FY RAGRV C
Sbjct: 338 YFRDMGYNVSMMADSTSRWAEALREISGRLAEMPADSGYPAYLGARLASFYERAGRVTCQ 397

Query: 182 GSPERIGSITMVGIATQSEG 241
           GSP R GS+++VG  +   G
Sbjct: 398 GSPSREGSVSLVGAVSPPGG 417




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 8, 2002  5:26 AM
  Number of letters in database: 292,803,047
  Number of sequences in database:  933,889

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140
```

```
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 179,819,297
Number of Sequences: 933889
Number of extensions: 3547236
Number of successful extensions: 15209
Number of sequences better than 10.0: 322
Number of HSP's better than 10.0 without gapping: 14973
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 15205
length of database: 292,803,047
effective HSP length: 74
effective length of database: 223,695,261
effective search space used: 5368686264
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020951445-029814-4573
```

7.1.3 Query= **hy-3_11a245004**

```
         (390 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           933,889 sequences; 292,803,047 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 21 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                       (bits) Value

gi|16804941|ref|NP_472970.1|  (NC_000910) monosaccharide tra...    42   0.001
gi|20178387|ref|NP_619808.1|  (NC_003663) V018 [Cowpox virus...    34   0.25
gi|19424326|ref|NP_598295.1|  (NM_133611) solute carrier fam...    34   0.33
gi|17560496|ref|NP_506313.1|  (NM_073912) transport protein ...    33   0.56
gi|16418395|ref|NP_443117.1|  (NM_052885) proton (H+) myo-in...    33   0.73
gi|20428639|ref|NP_620469.1|  (NM_138984) ACXD gene product ...    31   2.1
gi|1808606|emb|CAA64088.1|  (X94355) D3L [Cowpox virus] >gi|...    31   2.8
gi|10956360|ref|NP_052809.1|  (NC_001496) pXO1-113 [Bacillus...    31   2.8
gi|6324469|ref|NP_014538.1|  (NC_001147) member of sugar tra...    31   2.8
gi|17568783|ref|NP_508569.1|  (NM_076168) glucose transporte...    30   3.6
gi|15792636|ref|NP_282459.1|  (NC_002163) possible flagellar...    30   3.6
gi|15219226|ref|NP_173089.1|  (NM_101505) putative transport...    30   4.7
gi|15829006|ref|NP_326366.1|  (NC_002771) unknown; predicted...    30   4.7
gi|20093146|ref|NP_619221.1|  (NC_003552) conserved hypothet...    30   4.7
gi|125895|sp|P06811|LAC1_NEUCR  Laccase precursor (Benzenedi...    30   6.2
gi|66288|pir||KSNCLO  laccase (EC 1.10.3.2) precursor - Neur...    30   6.2
gi|15230987|ref|NP_188627.1|  (NM_112883) monosaccharide tra...    30   6.2
gi|20089976|ref|NP_616051.1|  (NC_003552) predicted protein ...    30   6.2
gi|14331015|emb|CAC41012.1|  (AJ304839) putative chloroplast...    30   6.2
gi|17149213|gb|AAL35943.1|AF445192_1  (AF445192) occluded vi...    30   6.2
gi|4505343|ref|NP_002477.1|  (NM_002486) nuclear cap binding...    29   8.1

                               Alignments
```

```
>gi|16804941|ref|NP_472970.1|  (NC_000910) monosaccharide transporter [Plasmodium
falciparum]
 gi|7446735|pir||B71620  monosaccharide transporter PFB0210c - malaria parasite  (Plasmodium
falciparum)
 gi|3845125|gb|AAC71831.1|  (AE001381) monosaccharide transporter [Plasmodium falciparum
3D7]
 gi|4007665|emb|CAA10374.1|  (AJ131457) hexose transporter 1 [Plasmodium falciparum]
 gi|4574228|gb|AAD23958.1|AF101827_1  (AF101827) putative sugar transporter [Plasmodium
falciparum]
          Length = 504

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 28/99 (28%), Positives = 51/99 (51%)
 Frame = +2

Query: 2   RVXXLFPTVICILRSCIMQFVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIYKDQYV 181
           R+  LFP+VI ++    +   F  ++P  L ++G +E++K +LK+        IY+   V
Sbjct: 207 RLMFLFPSVISLIGILALVVFFKEETPYFLFEKGRIEESKNILKK--------IYETDNV 258

Query: 182 DEQFERYQKMATDQQKQSNKSFLSVFERKKLVTLQTGII 298
           DE      K A +Q + + K+ LS+    K+ + +  II
Sbjct: 259 DEPLNAI-KEAVEQNESAKKNSLSLLSALKIPSYRYVII 296



>gi|20178387|ref|NP_619808.1|  (NC_003663) V018 [Cowpox virus]
 gi|20153005|gb|AAM13466.1|AF482758_17  (AF482758) V018 [Cowpox virus]
          Length = 796

 Score = 34.3 bits (77), Expect = 0.25
 Identities = 13/50 (26%), Positives = 29/50 (58%)
 Frame = -1

Query: 363 KXCKHSTPYLYQHLIVKLLNTDIIPVYKVTSFFLSKTLKNDLLLYFYQSV 214
           K C + T   Y   ++K  N +I+P   +T+F ++  ++ND+++ +  S+
Sbjct: 512 KKCNYKTNDAYHFRMIKYYNNNILPFSGMTAFHVAVCIRNDVIMKYLLSI 561



>gi|19424326|ref|NP_598295.1|  (NM_133611) solute carrier family 2 (facilitated glucose
transporter), member 13 [Rattus norvegicus]
 gi|15211931|emb|CAC51117.1|  (AJ315643) proton myo-inositol transporter [Rattus
norvegicus]
          Length = 618

 Score = 33.9 bits (76), Expect = 0.33
 Identities = 16/60 (26%), Positives = 35/60 (57%)
 Frame = +2

Query: 59  FVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIYKDQYVDEQFERYQKMATDQQKQSN 238
           F+F  +SP  LIQ+G  +KA+ +L Q+          +Q +DE+++  +    +++K+++
Sbjct: 226 FLFLPESPRWLIQKGQTQKARRILSQM--------RGNQTIDEEYDSIRNSIEEEEKEAS 277
```

```
>gi|17560496|ref|NP_506313.1|  (NM_073912) transport protein [Caenorhabditis elegans]
 gi|7500373|pir||T21682  hypothetical protein F32H5.4 - Caenorhabditis elegans
 gi|3876630|emb|CAB04250.1|  (Z81524) predicted using Genefinder~contains similarity to Pfam
domain: PF00083 (Sugar (and other) transporter), Score=3.9, E-value=1.9e-07, N=1
[Caenorhabditis elegans]
          Length = 528

 Score = 33.1 bits (74), Expect = 0.56
 Identities = 21/77 (27%), Positives = 42/77 (54%), Gaps = 2/77 (2%)
 Frame = +2

Query: 59  FVFNYDSPEQLIQRGNMEKAKYVLKQLI-IQVICSIYKDQYVDEQFERYQKMATDQ-QKQ 232
           F+F  +SP  LIQ G +E+A+ VLK ++ +    S +    ++E  +  +K   ++ +K+
Sbjct: 248 FLFVKESPRWLIQAGRIEEARQVLKHIMEVDGNTSDHSWNEIEEMLQTEKKRQEERIKKR 307

Query: 233 SNKSFLSVFERKKLVTL 283
            N  F  +F  K + ++
Sbjct: 308 KNYDFRHLFWNKYMASV 324




>gi|16418395|ref|NP_443117.1|  (NM_052885) proton (H+) myo-inositol symporter [Homo
sapiens]
 gi|20177982|sp|Q96QE2|MYCT_HUMAN  Proton myo-inositol co-transporter (Hmit)
 gi|15211933|emb|CAC51116.1|  (AJ315644) proton myo-inositol transporter [Homo sapiens]
          Length = 629

 Score = 32.7 bits (73), Expect = 0.73
 Identities = 16/58 (27%), Positives = 33/58 (56%)
 Frame = +2

Query: 59  FVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIYKDQYVDEQFERYQKMATDQQKQ 232
           F+F  +SP  LIQ+G  +KA+ +L Q+          +Q +DE+++  +    +++K+
Sbjct: 237 FLFLPESPRWLIQKGQTQKARRILSQM--------RGNQTIDEEYDSIKNNIEEEEKE 286




>gi|20428639|ref|NP_620469.1|  (NM_138984) ACXD gene product [Drosophila melanogaster]
 gi|6959484|gb|AAF33110.1|AF177931_1  (AF177931) adenylyl cyclase ACXD [Drosophila
melanogaster]
          Length = 1162

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 15/39 (38%), Positives = 26/39 (66%)
 Frame = -1

Query: 318 VKLLNTDIIPVYKVTSFFLSKTLKNDLLLYFYQSVAIFW 202
           +K+L T+I+P + V  F+LS  L+N+L    Y +VA+ +
Sbjct: 853 IKVLLTNILPTH-VADFYLSNQLQNELYYEEYDNVAVMF 890




>gi|1808606|emb|CAA64088.1|  (X94355) D3L [Cowpox virus]
 gi|3096979|emb|CAA72553.1|  (Y11842) C3L [Cowpox virus]
          Length = 833

 Score = 30.8 bits (68), Expect = 2.8
```

```
 Identities = 12/43 (27%), Positives = 24/43 (54%)
 Frame = -1

Query: 357 CKHSTPYLYQHLIVKLLNTDIIPVYKVTSFFLSKTLKNDLLLY 229
           C + T  +Y   ++K  N +I+P   +T+F ++  L+ D + Y
Sbjct: 558 CNYETNDVYHRQMIKYYNKNIVPFSGMTAFHVAVCLRKDAMKY 600



>gi|10956360|ref|NP_052809.1|  (NC_001496) pXO1-113 [Bacillus anthracis]
 gi|17367296|sp|Q9ZFB4|GRXA_BACAN  Spore germination protein XA
 gi|7474221|pir||A59105  hypothetical protein pXO1-113 - Bacillus anthracis virulence
plasmid pXO1
 gi|4092084|gb|AAC99415.1|  (AF108144) spore germination protein GerXA [Bacillus anthracis]
 gi|4894329|gb|AAD32417.1|AAD32417  (AF065404) pXO1-113 [Bacillus anthracis]
 gi|16031482|emb|CAC93927.1|  (AJ413932) GerXA protein [Bacillus anthracis]
 gi|16031486|emb|CAC93930.1|  (AJ413933) GerXA protein [Bacillus anthracis]
          Length = 492

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 16/44 (36%), Positives = 24/44 (54%)
 Frame = -1

Query: 132 FNTYFAFSMFPLCINYSGLSQLNTNCIIHERRMQMTVGNKXXTL 1
           F T   FS+FPL I Y+G      NC++H R + +  G+   T+
Sbjct: 216 FITGDQFSLFPL-IEYTGRPDYAVNCLLHGRFILLVDGSPTATI 258



>gi|6324469|ref|NP_014538.1|  (NC_001147) member of sugar transporter superfamily; Itr2p
[Saccharomyces cerevisiae]
 gi|266396|sp|P30606|ITR2_YEAST  Myo-inositol transporter 2
 gi|101449|pir||B40538  myo-inositol transport protein ITR2 - yeast (Saccharomyces
cerevisiae)
 gi|218439|dbj|BAA14367.1|  (D90353) myo-inositol transporter [Saccharomyces cerevisiae]
 gi|663251|emb|CAA88159.1|  (Z48149) ORF [Saccharomyces cerevisiae]
 gi|1419961|emb|CAA99119.1|  (Z74845) ORF YOL103w [Saccharomyces cerevisiae]
          Length = 612

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 30/119 (25%), Positives = 58/119 (48%), Gaps = 2/119 (1%)
 Frame = +2

Query: 14  LFPTVICILRSCIMQFVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIYKDQYVDEQF 193
           L PTV+     C +      D+P   + +G++++AK VLK+  +       +D+ +D++
Sbjct: 279 LIPTVLQFSFFCFLP-----DTPRYYVMKGDLKRAKMVLKRSYVNT-----EDEIIDQKV 328

Query: 194 ERYQKMATDQQKQSNKSFLSVFER--KKLVTLQTGIISVFNNLTIRCQYKYGVECLQXF 364
           E   ++++  Q    K+ ++ F    K+L T    + S F  L I C    G++ +Q F
Sbjct: 329 E---ELSSLNQSIPGKNPITKFWNMVKELHT----VPSNFRALIIGC----GLQAIQQF 376



>gi|17568783|ref|NP_508569.1|  (NM_076168) glucose transporter [Caenorhabditis elegans]
 gi|7505626|pir||T16593  hypothetical protein K09C4.1 - Caenorhabditis elegans
 gi|1125851|gb|AAA83627.1|  (U43375) Hypothetical protein K09C4.1 [Caenorhabditis elegans]
          Length = 520
```

```
 Score = 30.4 bits (67), Expect = 3.6
 Identities = 24/88 (27%), Positives = 44/88 (49%), Gaps = 7/88 (7%)
 Frame = +2

Query: 14  LFPTVICILRSCIMQFVFNYDSPEQLIQRGNMEKAK-----YVLKQLIIQVICSIYKDQY 178
           LF  V  ++  C+M  +   DSP+ LIQ+  ++KA+     Y  K  + +V+ S+ K++
Sbjct: 189 LFEMVCSVMILCLM--IHLPDSPKWLIQQNKVKKAEDSIKFYYGKHFLDEVVTSLIKEKN 246

Query: 179 V--DEQFERYQKMATDQQKQSNKSFLSV 256
           +  D +    Q    D  ++S K   +V
Sbjct: 247 LTKDNRISLRQIWENDTLRESLKILFAV 274




>gi|15792636|ref|NP_282459.1|  (NC_002163) possible flagellar protein [Campylobacter jejuni]
 gi|11282857|pir||F81274  probable flagellar protein Cj1313 [imported] - Campylobacter
jejuni (strain NCTC 11168)
 gi|6968747|emb|CAB73740.1|  (AL139078) possible flagellar protein [Campylobacter jejuni]
          Length = 157

 Score = 30.4 bits (67), Expect = 3.6
 Identities = 18/77 (23%), Positives = 37/77 (47%)
 Frame = +2

Query: 104 NMEKAKYVLKQLIIQVICSIYKDQYVDEQFERYQKMATDQQKQSNKSFLSVFERKKLVTL 283
           N ++ K + K      I    K +++D  FE + +   +  + SNK +  VF+ ++++
Sbjct: 11  NSQEIKLIFKWRNHPDISQFMKTKHID--FEEHLRFIRNLHQDSNKKYFLVFQDEQII-- 66

Query: 284 QTGIISVFNNLTIRCQY 334
             G+I   N  T  C++
Sbjct: 67  --GVIDFVNITTKSCEF 81




>gi|15219226|ref|NP_173089.1|  (NM_101505) putative transport protein [Arabidopsis thaliana]
 gi|4966360|gb|AAD34691.1|AC006341_19  (AC006341) Is a member of the PF|00083 sugar
transporter family. [Arabidopsis thaliana]
 gi|18958009|gb|AAL79578.1|  (AY078972) At1g16390/F3O9_19 [Arabidopsis thaliana]
          Length = 518

 Score = 30.0 bits (66), Expect = 4.7
 Identities = 23/89 (25%), Positives = 40/89 (44%), Gaps = 14/89 (15%)
 Frame = -1

Query: 291 PVYKVTSFFLSKTLKNDLLLYFYQSVAIFWYLSNCSSTY*SLQXEQM------TQMISCF 130
           P + +T FF+ K  + D L+ F         LS  SS   ++  +Q+       +++S F
Sbjct: 367 PAFLITFFFIDKINRRDALIGFTA-------LSGISSALIAVLGQQLGSLQIVLELVSFF 419

Query: 129 NTYFAFSM--------FPLCINYSGLSQL 67
           +   AF+M        FP C+  S +S +
Sbjct: 420 SACTAFNMTLIYTIEMFPTCVRNSAISMV 448




>gi|15829006|ref|NP_326366.1|  (NC_002771) unknown; predicted coding region [Mycoplasma
pulmonis]
 gi|14089950|emb|CAC13708.1|  (AL445565) unknown; predicted coding region [Mycoplasma
pulmonis]
```

```
          Length = 299

 Score = 30.0 bits (66), Expect = 4.7
 Identities = 25/94 (26%), Positives = 39/94 (40%), Gaps = 11/94 (11%)
 Frame = +2

Query: 14  LFPTVICILRSCIMQFVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIY--------- 166
           LF  +  IL S +M F+  Y     + +  N+E  K  +    I V+ S
Sbjct: 204 LFFNISLILISSLMGFLSAYS----IAKNSNLESVKIFIVISFITVVVSFLTNLLVTFSL 259

Query: 167 --KDQYVDEQFERYQKMATDQQKQSNKSFLSVFE 262
              K Q +DEQ E+   M  D + +S+     +FE
Sbjct: 260 SSKAQKIDEQIEKIDLMLKDIKNESDIKINELFE 293



>gi|20093146|ref|NP_619221.1|  (NC_003552) conserved hypothetical protein [Methanosarcina
acetivorans str. C2A]
 gi|19918486|gb|AAM07701.1|  (AE011154) conserved hypothetical protein [Methanosarcina
acetivorans str. C2A]
          Length = 330

 Score = 30.0 bits (66), Expect = 4.7
 Identities = 14/49 (28%), Positives = 23/49 (46%)
 Frame = +2

Query: 137 LIIQVICSIYKDQYVDEQFERYQKMATDQQKQSNKSFLSVFERKKLVTL 283
           L++ ++ SI      D+Q +RYQK+  D Q        + F   KL  +
Sbjct: 31  LVLMLVSSITGINNYDKQLDRYQKLQADTQGLEENDVFTYFPEPKLTAV 79



>gi|125895|sp|P06811|LAC1_NEUCR  Laccase precursor (Benzenediol:oxygen oxidoreductase)
(Urishiol oxidase) (Laccase allele OR)
 gi|168826|gb|AAA33591.1|  (M18333) laccase [Neurospora crassa]
          Length = 619

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 13/40 (32%), Positives = 25/40 (62%), Gaps = 2/40 (5%)
 Frame = -1

Query: 312 LLNTDIIPV--YKVTSFFLSKTLKNDLLLYFYQSVAIFWY 199
           +++ D++PV  YKV S FL    + D+++   Q+V  +W+
Sbjct: 300 IISADLVPVQPYKVDSLFLGVGQRYDVIIDANQAVGNYWF 339



>gi|66288|pir||KSNCLO  laccase (EC 1.10.3.2) precursor - Neurospora crassa (strain OR)
          Length = 619

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 13/40 (32%), Positives = 25/40 (62%), Gaps = 2/40 (5%)
 Frame = -1

Query: 312 LLNTDIIPV--YKVTSFFLSKTLKNDLLLYFYQSVAIFWY 199
           +++ D++PV  YKV S FL    + D+++   Q+V  +W+
Sbjct: 300 IISADLVPVQPYKVDSLFLGVGQRYDVIIDANQAVGNYWF 339
```

```
>gi|15230987|ref|NP_188627.1|  (NM_112883) monosaccharide transport protein, STP4
[Arabidopsis thaliana]
 gi|99758|pir||S25009  monosaccharide transport protein STP4 - Arabidopsis thaliana
 gi|16524|emb|CAA47325.1|  (X66857) sugar transport protein [Arabidopsis thaliana]
 gi|11994205|dbj|BAB01308.1|  (AB025631) monosaccharide transporter STP4 [Arabidopsis
thaliana]
 gi|13605906|gb|AAK32938.1|AF367352_1  (AF367352) AT3g19930/MPN9_17 [Arabidopsis thaliana]
 gi|16226824|gb|AAL16272.1|AF428342_1  (AF428342) AT3g19930/MPN9_17 [Arabidopsis thaliana]
          Length = 514

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 28/101 (27%), Positives = 48/101 (46%), Gaps = 1/101 (0%)
 Frame = +2

Query: 20  PTVICILRSCIMQFVFNYDSPEQLIQRGNMEKAKYVLKQLIIQVICSIYKDQYVDEQFER 199
           P V+ ++ + I+      D+P  LI+RG  E+AK +L+        SI     VDE+F+
Sbjct: 210 PAVMIMIGALILP-----DTPNSLIERGYTEEAKEMLQ--------SIRGTNEVDEEFQD 256

Query: 200 YQKMATDQQKQSNKSFLSV-FERKKLVTLQTGIISVFNNLT 319
               A+++ KQ    + ++   R +   + T  I  F  LT
Sbjct: 257 LID-ASEESKQVKHPWKNIMLPRYRPQLIMTCFIPFFQQLT 296



>gi|20089976|ref|NP_616051.1|  (NC_003552) predicted protein [Methanosarcina acetivorans
str. C2A]
 gi|19914939|gb|AAM04531.1|  (AE010777) predicted protein [Methanosarcina acetivorans str.
C2A]
          Length = 542

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 20/48 (41%), Positives = 27/48 (55%), Gaps = 3/48 (6%)
 Frame = -2

Query: 317 LNY*IPILFQSIKQQVSFFQKHSRMIYYFI---SINQWPFSGTFQIVH 183
           LNY   +LF SI     FF  H+ ++YY +   +IN    SGTF IV+
Sbjct: 447 LNY--VLLFVSIIAMSVFFSVHNMVLYYLLQPYNINLEMKSGTFAIVN 492


>gi|14331015|emb|CAC41012.1|  (AJ304839) putative chloroplast terpene synthase [Quercus
ilex]
          Length = 597

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 11/24 (45%), Positives = 16/24 (65%)
 Frame = +3

Query: 78  VLNN*YKGETWKRQNMYQNSQSFR 149
           V NN + G+TWK +N+Y  +  FR
Sbjct: 135 VYNNDHGGDTWKAENLYATALKFR 158
```

```
>gi|17149213|gb|AAL35943.1|AF445192_1  (AF445192) occluded virion protein [Spodoptera litura
nucleopolyhedrovirus]
          Length = 337

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 21/69 (30%), Positives = 34/69 (48%), Gaps = 3/69 (4%)
 Frame = -1

Query: 201 YLSNCSSTY*SLQXEQMT-QMISCFNTY--FAFSMFPLCINYSGLSQLNTNCIIHERRMQ 31
           Y++     Y +    Q+T QM+S  NT        +P+ IN +     +TN I + +R+Q
Sbjct: 47  YVNKIVRYYRTNDMSQLTPQMLSLINTIRNVCIETYPVDINVTKRFDNDTNLINNYKRLQ 106

Query: 30  MTVGNKXXT 4
             +GNK  T
Sbjct: 107 KELGNKPIT 115




>gi|4505343|ref|NP_002477.1| ☐ (NM_002486) nuclear cap binding protein subunit 1, 80kD;
nuclear cap binding protein, 80kD; nuclear cap binding protein 1, 80kD [Homo sapiens]
 gi|1705654|sp|Q09161|CB80_HUMAN  80 kDa nuclear cap binding protein (NCBP 80 kDa subunit)
(CBP80)
 gi|1082651|pir||S50082  nuclear cap binding protein - human
 gi|563368|emb|CAA56334.1| ☐ (X80030) cap binding protein [Homo sapiens]
 gi|577628|dbj|BAA06769.1| ☐ (D32002) nuclear cap binding protein [Homo sapiens]
 gi|12655187|gb|AAH01450.1|AAH01450 ☐ (BC001450) nuclear cap binding protein subunit 1,
80kD [Homo sapiens]
          Length = 790

 Score = 29.3 bits (64), Expect = 8.1
 Identities = 20/78 (25%), Positives = 31/78 (39%), Gaps = 10/78 (12%)
 Frame = +2

Query: 167 KDQYVDEQFERYQKMATDQQKQSNKSFLSVFERKKLVTLQTGIISVFNNLTIRCQ----- 331
           KD  ++EQ ER Q+     Q +    FL +F+R          I +     +RC+
Sbjct: 684 KDGVLEEQIERLQEKVESAQSEQKNLFLVIFQR---------FIMILTEHLVRCETDGTS 734

Query: 332 -----YKYGVECLQXFII 370
                YK  +E LQ   +
Sbjct: 735 VLTPWYKNCIERLQQIFL 752




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 8, 2002  5:26 AM
  Number of letters in database: 292,803,047
  Number of sequences in database:  933,889

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140
```

```
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 201,135,928
Number of Sequences: 933889
Number of extensions: 3809791
Number of successful extensions: 11974
Number of sequences better than 10.0: 42
Number of HSP's better than 10.0 without gapping: 11685
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 11969
length of database: 292,803,047
effective HSP length: 105
effective length of database: 194,744,702
effective search space used: 4673872848
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

RID: 1020951731-05181-2777

7.1.4 Query= **hy-4_11b250001**

(361 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
933,889 sequences; 292,803,047 total letters

If you have any problems or questions with the results of this search
please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 150 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|17539958|ref|NP_501507.1|  (NM_069106) signal recognition...   127   2e-29
gi|6755660|ref|NP_036029.1|  (NM_011899) signal recognition ...   123   3e-28
gi|91300|pir||S05198  signal recognition particle 54K protei...   123   3e-28
gi|12848354|dbj|BAB27921.1|  (AK011928) data source:MGD, sou...   123   3e-28
gi|226430|prf||1512311A  SRP protein 54kD [Mus musculus]          123   3e-28
gi|4507215|ref|NP_003127.1|  (NM_003136) signal recognition ...   123   3e-28
gi|2119055|pir||S51597  signal recognition particle 54K prot...   122   6e-28
gi|1711508|sp|P49968|SR51_HORVU  SIGNAL RECOGNITION PARTICLE...   122   7e-28
gi|1711511|sp|P49969|SR52_HORVU  SIGNAL RECOGNITION PARTICLE...   122   7e-28
gi|20151307|gb|AAM11013.1|  (AY094660) AT23778p [Drosophila ...   120   2e-27
gi|17647949|ref|NP_523931.1|  (NM_079207) Signal recognition...   120   2e-27
gi|1711509|sp|P49971|SR51_LYCES  SIGNAL RECOGNITION PARTICLE...   120   2e-27
gi|20160917|dbj|BAB89854.1|  (AP003295) putative signal reco...   120   2e-27
gi|1711512|sp|P49972|SR52_LYCES  SIGNAL RECOGNITION PARTICLE...   120   4e-27
gi|13542673|gb|AAH05543.1|AAH05543  (BC005543) Similar to si...   119   6e-27
gi|11094805|gb|AAG29734.1|AC084414_2  (AC084414) signal reco...   115   7e-26
gi|1711513|sp|P49967|SR53_ARATH  SIGNAL RECOGNITION PARTICLE...   115   7e-26
gi|18402407|ref|NP_564535.1|  (NM_103783) signal recognition...   115   7e-26
gi|1711514|sp|P49970|SR53_HORVU  SIGNAL RECOGNITION PARTICLE...   114   2e-25
gi|15239909|ref|NP_199761.1|  (NM_124327) SRP54 (signal reco...   114   2e-25
gi|1711510|sp|P49966|SR52_ARATH  Signal recognition particle...   113   3e-25
gi|3334344|sp|O42816|SR54_CANAL  SIGNAL RECOGNITION PARTICLE...   107   2e-23
gi|19075709|ref|NP_588209.1|  (NC_003421) signal recognition...   103   5e-22
gi|226739|prf||1604366A  signal recognition particle 54kD pr...   103   5e-22
gi|4538|emb|CAA34781.1|  (X16908) Srh1p [Saccharomyces cerev...   102   8e-22
```

```
gi|6325345|ref|NP_015413.1|  (NC_001148) Signal recognition ...   102    8e-22
gi|18394225|ref|NP_563970.1|  (NM_101400) expressed protein ...   100    5e-21
gi|2500887|sp|Q99150|SR54_YARLI  SIGNAL RECOGNITION PARTICLE...    96    6e-20
gi|19569901|gb|AAL92215.1|  (AC115580) Signal recognition pa...    91    2e-18
gi|15679321|ref|NP_276438.1|  (NC_000916) signal recognition...    91    2e-18
gi|14520695|ref|NP_126170.1|  (NC_000868) signal recognition...    89    9e-18
gi|15594157|dbj|BAB64926.1|  (AB057373) signal recognition p...    89    1e-17
gi|18978103|ref|NP_579460.1|  (NC_003413) signal recognition...    89    1e-17
gi|12230630|sp|O59307|SR54_PYRHO  Signal recognition 54 kDa ...    87    3e-17
gi|14591457|ref|NP_143537.1|  (NC_000961) signal recognition...    87    3e-17
gi|3334340|sp|O15821|SR54_ENTHI  Signal recognition particle...    86    6e-17
gi|15668272|ref|NP_247065.1|  (NC_000909) signal recognition...    86    8e-17
gi|20095048|ref|NP_614895.1|  (NC_003551) Signal recognition...    85    2e-16
gi|19074175|ref|NP_584781.1|  (NC_003231) SIGNAL RECOGNITION...    83    6e-16
gi|3334345|sp|Q00179|SR54_ASPNG  SIGNAL RECOGNITION PARTICLE...    82    1e-15
gi|15921563|ref|NP_377232.1|  (NC_003106) 356aa long hypothe...    82    1e-15
gi|20093371|ref|NP_619446.1|  (NC_003552) signal recognition...    81    2e-15
gi|15612144|ref|NP_223796.1|  (NC_000921) putative SIGNAL RE...    80    3e-15
gi|15645766|ref|NP_207943.1|  (NC_000915) signal recognition...    80    3e-15
gi|12230633|sp|P70722|SR54_ACIAM  Signal recognition 54 kDa ...    78    2e-14
gi|15897850|ref|NP_342455.1|  (NC_002754) Signal recognition...    78    2e-14
gi|13541823|ref|NP_111511.1|  (NC_002689) Signal recognition...    77    4e-14
gi|16081627|ref|NP_393991.1|  (NC_002578) probable signal re...    76    6e-14
gi|12230625|sp|O07853|SR54_SULAC  Signal recognition 54 kDa ...    76    8e-14
gi|15595039|ref|NP_212828.1|  (NC_001318) signal recognition...    76    8e-14
gi|15791232|ref|NP_281056.1|  (NC_002607) signal recognition...    75    1e-13
gi|15277690|gb|AAK93963.1|AF395887_1  (AF395887) signal reco...    74    2e-13
gi|15836678|ref|NP_297366.1|  (NC_002488) signal recognition...    74    3e-13
gi|15828813|ref|NP_326173.1|  (NC_002771) SIGNAL RECOGNITION...    73    5e-13
gi|15792058|ref|NP_281881.1|  (NC_002163) signal recognition...    73    7e-13
gi|401118|sp|Q01442|SR54_MYCMY  SIGNAL RECOGNITION PARTICLE ...    73    7e-13
gi|15892137|ref|NP_359851.1|  (NC_003103) signal recognition...    72    9e-13
gi|19704725|ref|NP_604287.1|  (NC_003454) Signal recognition...    72    9e-13
gi|18314019|ref|NP_560686.1|  (NC_003364) signal recognition...    72    1e-12
gi|13473694|ref|NP_105262.1|  (NC_002678) signal recognition...    71    2e-12
gi|15606801|ref|NP_214181.1|  (NC_000918) signal recognition...    71    3e-12
gi|15604048|ref|NP_220563.1|  (NC_000963) SIGNAL RECOGNITION...    70    6e-12
gi|15924227|ref|NP_371761.1|  (NC_002758) signal recognition...    69    7e-12
gi|17547530|ref|NP_520932.1|  (NC_003295) PUTATIVE SIGNAL RE...    69    1e-11
gi|3746964|gb|AAC64139.1|  (AF092168) signal recognition par...    69    1e-11
gi|15237580|ref|NP_196014.1|  (NM_120476) signal recognition...    69    1e-11
gi|17936572|ref|NP_533362.1|  (NC_003304) signal recognition...    68    2e-11
gi|15889951|ref|NP_355632.1|  (NC_003062) AGR_C_4890p [Agrob...    68    2e-11
gi|15616997|ref|NP_240210.1|  (NC_002528) signal recognition...    67    3e-11
gi|17986509|ref|NP_539143.1|  (NC_003317) SIGNAL RECOGNITION...    67    3e-11
gi|3746903|gb|AAC64109.1|  (AF089724) signal recognition par...    67    3e-11
gi|15966993|ref|NP_387346.1|  (NC_003047) PROBABLE SIGNAL RE...    67    4e-11
gi|15834914|ref|NP_296673.1|  (NC_002620) signal recognition...    67    4e-11
gi|15604743|ref|NP_219527.1|  (NC_000117) Signal Recognition...    67    4e-11
gi|14601593|ref|NP_148133.1|  (NC_000854) signal recognition...    67    5e-11
gi|7799381|emb|CAB90997.1|  (AJ239043) signal recognition pa...    66    6e-11
gi|11498230|ref|NP_069456.1|  (NC_000917) signal recognition...    66    6e-11
gi|7769071|gb|AAF69242.1|AF173880_5  (AF173880) signal recog...    66    6e-11
gi|15615047|ref|NP_243350.1|  (NC_002570) signal recognition...    66    8e-11
gi|17229444|ref|NP_485992.1|  (NC_003272) signal recognition...    66    8e-11
gi|19568856|gb|AAL91941.1|AF482013_1  (AF482013) signal reco...    65    1e-10
gi|2500884|sp|Q55311|SR54_SYNP7  SIGNAL SEQUENCE BINDING PRO...    65    1e-10
gi|18310695|ref|NP_562629.1|  (NC_003366) signal recognition...    65    1e-10
gi|6900424|emb|CAB72034.1|  (AJ391261) signal recognition pa...    65    1e-10
gi|15795060|ref|NP_284882.1|  (NC_003116) signal recognition...    65    1e-10
gi|2126515|pir||S68160  probable RNA binding protein ffh - S...    65    1e-10
```

```
gi|16761530|ref|NP_457147.1|  (NC_003198) signal recognition...   65   2e-10
gi|13357648|ref|NP_077922.1|  (NC_002162) signal recognition...   65   2e-10
gi|15895031|ref|NP_348380.1|  (NC_003030) Signal recognition...   64   2e-10
gi|15644313|ref|NP_229365.1|  (NC_000853) signal recognition...   64   2e-10
gi|15676217|ref|NP_273349.1|  (NC_003112) signal recognition...   64   2e-10
gi|78628|pir||S07178  hypothetical protein, 48K (rpsP 5' reg...   64   4e-10
gi|1800015|dbj|BAA16495.1|  (D90888) SIGNAL RECOGNITION PART...   64   4e-10
gi|15803132|ref|NP_289163.1|  (NC_002655) GTP-binding export...   64   4e-10
gi|16123449|ref|NP_406762.1|  (NC_003143) signal recognition...   64   4e-10
gi|19746132|ref|NP_607268.1|  (NC_003485) putative signal re...   63   5e-10
gi|2500883|sp|Q54431|SR54_STRMU  Signal recognition particle...   63   5e-10
gi|16127883|ref|NP_422447.1|  (NC_002696) signal recognition...   63   7e-10
gi|15903209|ref|NP_358759.1|  (NC_003098) Signal recognition...   62   1e-09
gi|15901147|ref|NP_345751.1|  (NC_003028) signal recognition...   62   1e-09
```

Alignments

```
>gi|17539958|ref|NP_501507.1| ☐ (NM_069106) signal recognition particle protein (SRP54)
[Caenorhabditis elegans]
 gi|7499565|pir||T21140  hypothetical protein F21D5.7 - Caenorhabditis elegans  (fragment)
 gi|3876111|emb|CAA91040.1|  (Z54271) contains similarity to Pfam domain: PF00448 (SRP54-
type protein), Score=445.2, E-value=1.8e-130, N=1~cDNA EST EMBL:M88956 comes from this
gene~cDNA EST EMBL:Z14505 comes from this gene~cDNA EST EMBL:Z14541 comes from this
gene~cDNA EST EMBL:Z1>
 gi|3876124|emb|CAA92301.1|  (Z68161) contains similarity to Pfam domain: PF00448 (SRP54-
type protein), Score=445.2, E-value=1.8e-130, N=1~cDNA EST EMBL:M88956 comes from this
gene~cDNA EST EMBL:Z14505 comes from this gene~cDNA EST EMBL:Z14541 comes from this
gene~cDNA EST EMBL:Z1>
          Length = 496

 Score =  127 bits (319), Expect = 2e-29
 Identities = 56/120 (46%), Positives = 88/120 (72%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           A T+SP++FIG GEH DD E F+  SFV++LLG+GDI  L   V D+  ++D  +L+ +L
Sbjct: 264 AVTKSPVIFIGTGEHIDDFEIFKPKSFVQKLLGMGDIAGLVDMVNDI-GIQDNKELVGRL 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K+G+F++RD+  QF +++K+G  +Q M  IPG GS  ++KGNE+ES+ R++R + +M+SM
Sbjct: 323 KQGQFTLRDMYEQFQNIMKMGPFSQIMGMIPGFGSEFMTKGNEQESVNRLKRMMTVMDSM 382




>gi|6755660|ref|NP_036029.1| ☐ (NM_011899) signal recognition particle 54 kDa; 54K subunit
of signal recognition particle [Mus musculus]
 gi|134886|sp|P14576|SR54_MOUSE  Signal recognition particle 54 kDa protein (SRP54)
 gi|54194|emb|CAA34386.1| ☐ (X16319) SRP 54K subunit (AA 1-504) [Mus musculus]
          Length = 504
 Score =  123 bits (309), Expect = 3e-28
 Identities = 56/120 (46%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DD E F+   F+ +LLG+GDI  L   V ++  + D   LI+KL
Sbjct: 264 AATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNEALIEKL 322
```

```
Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382



>gi|91300|pir||S05198  signal recognition particle 54K protein - mouse
          Length = 504

 Score =  123 bits (309), Expect = 3e-28
 Identities = 56/120 (46%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DD E F+   F+ +LLG+GDI  L   V ++   + D   LI+KL
Sbjct: 264 AATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNEALIEKL 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382



>gi|12848354|dbj|BAB27921.1|  (AK011928) data source:MGD, source key:MGI:1346087,
evidence:ISS~putative~signal recognition particle 54 kDa [Mus musculus]
 gi|18043829|gb|AAH19683.1|AAH19683  (BC019683) signal recognition particle 54 kDa [Mus
musculus]
          Length = 504

 Score =  123 bits (309), Expect = 3e-28
 Identities = 56/120 (46%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DD E F+   F+ +LLG+GDI  L   V ++   + D   LI+KL
Sbjct: 264 AATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNEALIEKL 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382


>gi|226430|prf||1512311A  SRP protein 54kD [Mus musculus]
          Length = 504

 Score =  123 bits (309), Expect = 3e-28
 Identities = 56/120 (46%), Positives = 85/120 (70%)
 Frame = -2


Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DD E F+   F+ +LLG+GDI  L   V ++   + D   LI+KL
Sbjct: 264 AATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNEALIEKL 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382
```

```
>gi|4507215|ref|NP_003127.1| ☐ (NM_003136) signal recognition particle 54kD [Homo sapiens]
 gi|11433513|ref|XP_007239.1| ☐ (XM_007239) signal recognition particle 54kD [Homo
sapiens]
 gi|134884|sp|P13624|SR54_HUMAN  Signal recognition particle 54 kDa protein (SRP54)
 gi|89065|pir||S05197  signal recognition particle 54K protein - dog
 gi|1082802|pir||S54143  SRP 54 protein - human (fragment)
 gi|930|emb|CAA34385.1|  (X16318) 54k protein (AA 1-504) [Canis familiaris]
 gi|784999|emb|CAA60132.1| ☐ (X86373) SRP 54 [Homo sapiens]
 gi|1256820|gb|AAC50994.1| ☐ (U51920) signal recognition particle [Homo sapiens]
 gi|12653733|gb|AAH00652.1|AAH00652 ☐ (BC000652) signal recognition particle 54kD [Homo
sapiens]
 gi|13097267|gb|AAH03389.1|AAH03389 ☐ (BC003389) signal recognition particle 54kD [Homo
sapiens]
 gi|226429|prf||1512310A  SRP protein 54kD [Canis familiaris]
          Length = 504

 Score =  123 bits (309), Expect = 3e-28
 Identities = 56/120 (46%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
            AAT+SPI+FIG GEH DD E F+   F+ +LLG+GDI  L   V ++  + D   LI+KL
Sbjct: 264 AATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNEALIEKL 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382



>gi|2119055|pir||S51597  signal recognition particle 54K protein - tomato (cv. Rentita)
          Length = 496

 Score =  122 bits (307), Expect = 6e-28
 Identities = 56/120 (46%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
            AAT+SP++ IG GEH D+ E F+   FV RLLG+GD++ L   ++DVV M  QP+L+QKL
Sbjct: 264 AATKSPVIVIGTGEHMDEFEVFDVKPFVSRLLGMGDLSGLVNKIQDVVPMDQQPELLQKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
             EG F++R +  QF S+LK+G L Q  S +PG  + ++ +G EKES  + +R++ +M+SM
Sbjct: 324 SEGHFTLRIMYEQFQSMLKMGPLGQVFSMLPGFSAEMMPQGREKESQAKFKRYMTMMDSM 383



>gi|1711508|sp|P49968|SR51_HORVU  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 1 (SRP54)
 gi|7440067|pir||T06185  signal recognition particle 54 K protein - barley
 gi|1020000|gb|AAA79354.1|  (L48284) signal recognition particle 54 kDa subunit [Hordeum
vulgare]
          Length = 497

 Score =  122 bits (306), Expect = 7e-28
```

```
 Identities = 56/120 (46%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD++ L   ++DV+    QP+L+ KL
Sbjct: 264 AATKSPVIFIGTGEHIDEFEIFDVKPFVSRLLGMGDLSGLMDKIQDVMPADQQPELLAKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R L  QF ++LK+G + Q  S +PG  S ++ KG+EKE   +I+R++ IM+SM
Sbjct: 324 AEGTFTLRLLYEQFQNLLKMGPIGQVFSMLPGFSSELMPKGHEKEGQAKIKRYMTIMDSM 383



>gi|1711511|sp|P49969|SR52_HORVU  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 2 (SRP54)
 gi|7440068|pir||T06186  signal recognition particle 54 K protein 2 - barley
 gi|1020002|gb|AAA79355.1|  (L48285) signal recognition particle 54 kDa subunit [Hordeum
vulgare]
          Length = 497

 Score =  122 bits (306), Expect = 7e-28
 Identities = 56/120 (46%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD++ L   ++DV+    QP+L+ KL
Sbjct: 264 AATKSPVIFIGTGEHIDEFEIFDVKPFVSRLLGMGDLSGLMDKIQDVMPADQQPELLAKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R L  QF ++LK+G + Q  S +PG  S ++ KG+EKE   +I+R++ IM+SM
Sbjct: 324 AEGTFTLRLLYEQFQNLLKMGPIGQVFSMLPGFSSELMPKGHEKEGQAKIKRYMTIMDSM 383



>gi|20151307|gb|AAM11013.1|  (AY094660) AT23778p [Drosophila melanogaster]
          Length = 569

 Score =  120 bits (302), Expect = 2e-27
 Identities = 55/120 (45%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DDLE F+   FV +LLG+GDI  L   V ++  +    +L++K+
Sbjct: 264 AATQSPIIFIGTGEHIDDLEPFKTKPFVSKLLGMGDIEGLIDKVNEL-KLDGNDELLEKI 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G F+IRD+  QF +++K+G  +QFM+ IPG     ++KG E+ES+ R++R + +M+SM
Sbjct: 323 KHGHFTIRDMYEQFQNIMKMGPFSQFMNMIPGFSQDFMTKGGEQESMARVKRMMTMMDSM 382



>gi|17647949|ref|NP_523931.1|  (NM_079207) Signal recognition particle protein 54k
[Drosophila melanogaster]
 gi|5679038|gb|AAD46831.1|AF160891_1  (AF160891) BcDNA.GM09489 [Drosophila melanogaster]
 gi|7295492|gb|AAF50806.1|  (AE003567) Srp54k gene product [Drosophila melanogaster]
          Length = 508

 Score =  120 bits (302), Expect = 2e-27
```

```
 Identities = 55/120 (45%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH DDLE F+   FV +LLG+GDI  L   V ++  +    +L++K+
Sbjct: 264 AATQSPIIFIGTGEHIDDLEPFKTKPFVSKLLGMGDIEGLIDKVNEL-KLDGNDELLEKI 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G F+IRD+  QF +++K+G  +QFM+ IPG     ++KG E+ES+ R++R + +M+SM
Sbjct: 323 KHGHFTIRDMYEQFQNIMKMGPFSQFMNMIPGFSQDFMTKGGEQESMARVKRMMTMMDSM 382



>gi|1711509|sp|P49971|SR51_LYCES  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 1 (SRP54)
 gi|556900|emb|CAA84275.1|  (Z34518) 54-kD signal recognition particle (SRP) specific
protein [Lycopersicon esculentum]
          Length = 496

 Score =  120 bits (302), Expect = 2e-27
 Identities = 57/120 (47%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD++ L   ++DVV M  QP+L+QKL
Sbjct: 265 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGMGDLSGLVNKIQDVVPMDQQPELLQKL 324

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R +  QF S+LK+G L  F S +PG  + ++ +G EKES  + +R++ +M+SM
Sbjct: 325 SEGHFTLRIMYEQFQSMLKMGPLGVF-SMLPGFSAEMMPQGREKESQAKFKRYMTMMDSM 383



>gi|20160917|dbj|BAB89854.1|  (AP003295) putative signal recognition particle 54K protein
[Oryza sativa (japonica cultivar-group)]
          Length = 496

 Score =  120 bits (302), Expect = 2e-27
 Identities = 55/120 (45%), Positives = 83/120 (68%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD +     + +VV    QP+L+QKL
Sbjct: 264 AATKSPVIFIGTGEHIDEFEVFDVKPFVSRLLGMGDWSGFMDKIHEVVPTDQQPELLQKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R +  QF ++LK+G + Q  S +PG  S ++ KG+EKES  +I+R++ +M+SM
Sbjct: 324 SEGTFTLRLMYEQFQNILKMGPIGQVFSMLPGFSSELMPKGHEKESQAKIKRYMTMMDSM 383



>gi|1711512|sp|P49972|SR52_LYCES  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 2 (SRP54)
 gi|1076577|pir||S51598  signal recognition particle 54K protein - tomato (cv. UC82-B)
 gi|556902|emb|CAA84288.1|  (Z34527) 54-kD signal recognition particle (SRP) specific
protein [Lycopersicon esculentum]
          Length = 499

 Score =  120 bits (300), Expect = 4e-27
 Identities = 54/120 (45%), Positives = 82/120 (68%)
```

```
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD +     + +VV M  QP+L+QKL
Sbjct: 264 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGMGDWSGFMDKIHEVVPMDQQPELLQKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R +  QF ++LK+G + Q  S +PG  + ++ KG E ES  +I+R++ +M+SM
Sbjct: 324 SEGHFTLRIMYEQFQNILKMGPIGQVFSMLPGFSAELMPKGRENESQAKIKRYMTMMDSM 383



>gi|13542673|gb|AAH05543.1|AAH05543 ☐ (BC005543) Similar to signal recognition particle 54
kDa [Mus musculus]
          Length = 500

 Score =  119 bits (298), Expect = 6e-27
 Identities = 54/120 (45%), Positives = 84/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GEH D+ E F+   F+ +LLG+GDI  L   V ++  + D   LI+K
Sbjct: 264 AATKSPIIFIGTGEHIDNFEPFKTQPFISKLLGMGDIEGLIDKVNEL-KLDDNKALIEKW 322

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           K G+F++RD+  QF +++K+G  +Q +  IPG G+  +SKGNE+ES+ R+++ + IM+SM
Sbjct: 323 KHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDFMSKGNEQESMARLKKLMTIMDSM 382



>gi|11094805|gb|AAG29734.1|AC084414_2  (AC084414) signal recognition particle 54 kDa protein
2 (SRP54), putative [Arabidopsis thaliana]
          Length = 522

 Score =  115 bits (289), Expect = 7e-26
 Identities = 51/120 (42%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD +     +++VV    QP+L++KL
Sbjct: 291 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGMGDWSGFVDKLQEVVPKDQQPELLEKL 350

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            +G F++R +  QF ++L +G L +  S +PG+ + ++ KG+EKES  +I+R++ +M+SM
Sbjct: 351 SQGNFTLRIMYDQFQNILNMGPLKEVFSMLPGISAEMMPKGHEKESQAKIKRYMTMMDSM 410



>gi|1711513|sp|P49967|SR53_ARATH  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 3 (SRP54)
 gi|515681|gb|AAA66200.1|  (U12127) signal recognition particle 54 kDa subunit [Arabidopsis
thaliana]
          Length = 495

 Score =  115 bits (289), Expect = 7e-26
 Identities = 51/120 (42%), Positives = 84/120 (69%)
 Frame = -2
```

```
Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD +     +++VV    QP+L++KL
Sbjct: 264 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGMGDWSGFVDKLQEVVPKDQQPELLEKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            +G F++R +  QF ++L +G L +  S +PG+ + ++ KG+EKES  +I+R++ +M+SM
Sbjct: 324 SQGNFTLRIMYDQFQNILNMGPLKEVFSMLPGISAEMMPKGHEKESQAKIKRYMTMMDSM 383
```

```
>gi|18402407|ref|NP_564535.1|  (NM_103783) signal recognition particle 54 kDa protein 2
(SRP54), putative [Arabidopsis thaliana]
 gi|15450461|gb|AAK96524.1|  (AY052331) At1g48900/F27K7_8 [Arabidopsis thaliana]
 gi|15810010|gb|AAL06932.1|  (AY054273) At1g48900/F27K7_8 [Arabidopsis thaliana]
 gi|17386102|gb|AAL38597.1|AF446864_1  (AF446864) At1g48900/F27K7_8 [Arabidopsis thaliana]
          Length = 495

 Score =  115 bits (289), Expect = 7e-26
 Identities = 51/120 (42%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG+GD +     +++VV    QP+L++KL
Sbjct: 264 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGMGDWSGFVDKLQEVVPKDQQPELLEKL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            +G F++R +  QF ++L +G L +  S +PG+ + ++ KG+EKES  +I+R++ +M+SM
Sbjct: 324 SQGNFTLRIMYDQFQNILNMGPLKEVFSMLPGISAEMMPKGHEKESQAKIKRYMTMMDSM 383
```

```
>gi|1711514|sp|P49970|SR53_HORVU  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN 3 (SRP54)
 gi|7440069|pir||T06187  signal recognition particle 54 K protein 3 - barley
 gi|1020004|gb|AAA79356.1|  (L48286) signal recognition particle 54 kDa subunit [Hordeum
vulgare]
          Length = 493

 Score =  114 bits (286), Expect = 2e-25
 Identities = 53/120 (44%), Positives = 80/120 (66%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH DD + F    FV RLLG GD+  L   ++ +V    Q +L+ KL
Sbjct: 263 AATKSPVIFIGTGEHIDDFDVFNVEPFVARLLGRGDLPGLIDKMESIVPADQQSELVAKL 322


Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG F++R L  QF ++LK+G ++Q  S +PG  S ++ KG EK+S ++ +R++ IM+SM
Sbjct: 323 SEGAFTLRLLYEQFQNLLKMGPMSQIFSMLPGFSSELMPKGQEKQSKEKFKRYMTIMDSM 382
```

```
>gi|15239909|ref|NP_199761.1|  (NM_124327) SRP54 (signal recognition particle 54 KDa)
protein [Arabidopsis thaliana]
 gi|10177616|dbj|BAB10763.1|  (AB023033) SRP54 (signal recognition particle 54 KDa) protein
[Arabidopsis thaliana]
          Length = 497
```

```
 Score =  114 bits (285), Expect = 2e-25
 Identities = 52/120 (43%), Positives = 84/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+A  FV RLLG GD++     +++VV    QP+L++ L
Sbjct: 266 AATKSPVIFIGTGEHMDEFEVFDAKPFVSRLLGNGDMSGFVNKLQEVVPKDQQPELLEML 325

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
             GKF++R +  QF ++L +G L +  S +PGM + ++ +G+EKES  +I+R++ +M+SM
Sbjct: 326 SHGKFTLRIMYDQFQNMLNMGPLKEVFSMLPGMRAEMMPEGHEKESQAKIKRYMTMMDSM 385



>gi|1711510|sp|P49966|SR52_ARATH  Signal recognition particle 54 kDa protein 2 (SRP54)
 gi|515679|gb|AAA66199.1|  (U12126) signal recognition particle 54 kDa subunit [Arabidopsis
thaliana]
          Length = 495

 Score =  113 bits (283), Expect = 3e-25
 Identities = 52/120 (43%), Positives = 83/120 (68%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+A  FV RLLG GD++     +++VV    QP+L++ L
Sbjct: 264 AATKSPVIFIGTGEHMDEFEVFDAKPFVSRLLGNGDMSGFVNKLQEVVPKDQQPELLEML 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
             GKF++R +  QF ++L +G L +  S +PGM + ++ +G+EKES   I+R++ +M+SM
Sbjct: 324 SHGKFTLRIMYDQFQNMLNMGPLKEVFSMLPGMRAEMMPEGHEKESQAEIKRYMTMMDSM 383



>gi|3334344|sp|O42816|SR54_CANAL  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN HOMOLOG
 gi|2739331|emb|CAA10999.1|  (AJ222805) Srp54 protein [Candida albicans]
          Length = 556

 Score =  107 bits (268), Expect = 2e-23
 Identities = 55/122 (45%), Positives = 86/122 (70%), Gaps = 2/122 (1%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQ--PQLIQ 187
           AAT++PIVFIG GEH  DLE F+ ++F+ +LLG+GDI  L + V+ +   +D+   Q I+
Sbjct: 281 AATKTPIVFIGTGEHVGDLEIFKPTTFISKLLGIGDIQGLIEHVQSLNLHQDEGHKQTIE 340


Query: 186 KLKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMN 7
            +KEGKF++RD Q Q N+ LK+G L    S IPG+ S+++S+  ++E+ K+I+  + IM+
Sbjct: 341 HIKEGKFTLRDFQNQMNNFLKMGPLTNIASMIPGL-SNIMSQVGDEETSKKIKNMIYIMD 399

Query: 6   SM 1
           SM
Sbjct: 400 SM 401



>gi|19075709|ref|NP_588209.1|  (NC_003421) signal recognition particle 54 kd protein homolog
[Schizosaccharomyces pombe]
```

```
 gi|134887|sp|P21565|SR54_SCHPO  Signal recognition particle 54 kDa protein homolog (SRP54)
 gi|101076|pir||A33644  signal recognition particle 54K protein [imported] - fission yeast
(Schizosaccharomyces pombe)
 gi|5096|emb|CAA35951.1|  (X51613) signal recognition particle [Schizosaccharomyces pombe]
 gi|173491|gb|AAA35344.1|  (M55518) signal recognition particle 54 kDa subunit
[Schizosaccharomyces pombe]
 gi|4678686|emb|CAB41226.1|  (AL049662) signal recognition particle 54 kd protein homolog
[Schizosaccharomyces pombe]
          Length = 522

 Score =  103 bits (256), Expect = 5e-22
 Identities = 52/120 (43%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT++PIVFIG GEH +DLE F   SF+ +LLGLGD+  L + V+ +    D+  +++ L
Sbjct: 264 AATKTPIVFIGTGEHINDLERFSPRSFISKLLGLGDLEGLMEHVQSL--DFDKKNMVKNL 321

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           ++GKF++RD + Q  +++KLG L++  S IPGM S++++  N++E   R++R L I++SM
Sbjct: 322 EQGKFTVRDFRDQLGNIMKLGPLSKMASMIPGM-SNMMNGMNDEEGSLRMKRMLYIVDSM 380



>gi|226739|prf||1604366A  signal recognition particle 54kD protein [Schizosaccharomyces
pombe]
          Length = 522

 Score =  103 bits (256), Expect = 5e-22
 Identities = 52/120 (43%), Positives = 85/120 (70%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT++PIVFIG GEH +DLE F   SF+ +LLGLGD+  L + V+ +    D+  +++ L
Sbjct: 264 AATKTPIVFIGTGEHINDLERFSPRSFISKLLGLGDLEGLMEHVQSL--DFDKKNMVKNL 321

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           ++GKF++RD + Q  +++KLG L++  S IPGM S++++  N++E   R++R L I++SM
Sbjct: 322 EQGKFTVRDFRDQLGNIMKLGPLSKMASMIPGM-SNMMNGMNDEEGSLRMKRMLYIVDSM 380



>gi|4538|emb|CAA34781.1|  (X16908) Srh1p [Saccharomyces cerevisiae]
          Length = 541

 Score =  102 bits (254), Expect = 8e-22
 Identities = 45/120 (37%), Positives = 82/120 (67%)
 Frame = -2


Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT +PI+FIG GEH  DLE F   SF+ +LLG+GDI  LF+ ++ V N  D    ++ +
Sbjct: 272 AATNTPIIFIGTGEHIHDLEKFSPKSFISKLLGIGDIESLFEQLQTVSNKEDAKATMENI 331

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           ++GKF++ D + Q  +++K+G L+     IPGM S+++++  E+E+ +++++ + +++SM
Sbjct: 332 QKGKFTLLDFKKQMQTIMKMGPLSNIAQMIPGM-SNMMNQVGEEETSQKMKKMVYVLDSM 390
```

```
>gi|6325345|ref|NP_015413.1|  (NC_001148) Signal recognition particle subunit (homolog of
mammalian SRP54); Srp54p [Saccharomyces cerevisiae]
 gi|1711516|sp|P20424|SR54_YEAST  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN HOMOLOG (SRP54)
 gi|2119053|pir||JX0112  signal recognition particle 54K protein - yeast  (Saccharomyces
cerevisiae)
 gi|4542|emb|CAA35952.1|  (X51614) signal recognition particle [Saccharomyces cerevisiae]
 gi|172705|gb|AAA35092.1|  (M55517) signal recognition particle 54 kDa subunit
[Saccharomyces cerevisiae]
 gi|1230693|gb|AAB68136.1|  (U51033) Signal recognition particle 54 kd subunit (Swiss Prot.
accession number P20424) [Saccharomyces cerevisiae]
 gi|226740|prf||1604366B  signal recognition particle 54kD protein [Saccharomyces
cerevisiae]
          Length = 541

 Score =  102 bits (254), Expect = 8e-22
 Identities = 45/120 (37%), Positives = 82/120 (67%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT +PI+FIG GEH  DLE F   SF+ +LLG+GDI  LF+ ++ V N  D    ++ +
Sbjct: 272 AATNTPIIFIGTGEHIHDLEKFSPKSFISKLLGIGDIESLFEQLQTVSNKEDAKATMENI 331

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           ++GKF++ D + Q  +++K+G L+     IPGM S+++++  E+E+ +++++ + +++SM
Sbjct: 332 QKGKFTLLDFKKQMQTIMKMGPLSNIAQMIPGM-SNMMNQVGEEETSQKMKKMVYVLDSM 390



>gi|18394225|ref|NP_563970.1|  (NM_101400) expressed protein [Arabidopsis thaliana]
 gi|586036|sp|P37106|SR51_ARATH  Signal recognition particle 54 kDa protein 1 (SRP54)
 gi|629560|pir||S42550  signal recognition particle 54K protein - Arabidopsis thaliana
 gi|304111|gb|AAA19728.1|  (L19997) signal recognition particle 54 kDa subunit [Arabidopsis
thaliana]
 gi|5103829|gb|AAD39659.1|AC007591_24  (AC007591) Identical to gb|L19997 signal recognition
particle 54 kDa subunit (Srp54-1) from Arabidopsis thaliana.  ESTs gb|T88590 and gb|T20603
come from this gene
 gi|12745001|gb|AAK06880.1|AF344329_1  (AF344329) unknown protein [Arabidopsis thaliana]
          Length = 479

 Score = 99.8 bits (247), Expect = 5e-21
 Identities = 48/120 (40%), Positives = 80/120 (66%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SP++FIG GEH D+ E F+   FV RLLG GD + L   +++VV    Q +L++ L
Sbjct: 264 AATKSPVIFIGTGEHMDEFEVFDVKPFVSRLLGKGDWSGLVDKLQEVVPKDLQNELVENL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            +G F++R +   QF   L++  LNQ  S +PG+ + ++ KG+ +ES  +++R++ +M+SM
Sbjct: 324 SQGNFTLRSMYDQFQCSLRI-PLNQLFSMLPGISAEMMPKGHGEESRVKMKRYMTMMDSM 382



>gi|2500887|sp|Q99150|SR54_YARLI  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN HOMOLOG (SRP54)
 gi|1147795|gb|AAC49735.1|  (U42418) Srp54p [Yarrowia lipolytica]
          Length = 536

 Score = 96.3 bits (238), Expect = 6e-20
```

```
 Identities = 48/120 (40%), Positives = 83/120 (69%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT +PI+FIG GEH  DLE+F    F+ +LLG+GD+  L ++++  +N+ DQ + ++ +
Sbjct: 266 AATNTPIIFIGTGEHIHDLEAFSPKQFISKLLGIGDLQGLMETMQS-LNL-DQKKTMEHI 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           +EG F++ DL+ Q  ++LK+GSL+     IPG+ S + S  +++E  +RI+R + I++SM
Sbjct: 324 QEGIFTLADLRDQMGNMLKMGSLSSIAGMIPGL-SGMASSISDEEGTRRIKRMIYILDSM 382



>gi|19569901|gb|AAL92215.1|  (AC115580) Signal recognition particle 54 kDa protein
[Dictyostelium discoideum]
          Length = 542

 Score = 91.3 bits (225), Expect = 2e-18
 Identities = 41/103 (39%), Positives = 69/103 (66%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT +PI+FIG GEH  DLE F+ S+FV +LLG GD+  + + +K+V+   +    ++++
Sbjct: 264 AATNTPIIFIGTGEHLTDLELFDPSTFVSKLLGYGDMKGMLEKIKEVI--PEDSTSLKEI 321

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNE 52
            +GKF++R +Q QF  +++LG +++ +  IPGM      +GNE
Sbjct: 322 AQGKFTLRSMQQQFQQIMQLGPIDKLVQMIPGMNQLPQLQGNE 364



>gi|15679321|ref|NP_276438.1|  (NC_000916) signal recognition particle protein SRP54
[Methanothermobacter thermautotrophicus]
 gi|3334341|sp|O27376|SR54_METTH  Signal recognition 54 kDa protein (SRP54)
 gi|7440075|pir||F69042  signal recognition particle protein SRP54 - Methanobacterium
thermoautotrophicum (strain Delta H)
 gi|2622427|gb|AAB85799.1|  (AE000896) signal recognition particle protein SRP54
[Methanothermobacter thermautotrophicus]
          Length = 442

 Score = 90.9 bits (224), Expect = 2e-18
 Identities = 46/116 (39%), Positives = 71/116 (60%)
 Frame = -2

Query: 348 SPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLKEGK 169
           +PI FIG GE  DDLE F+   F+ RLLG+GD+  L + V++V       + +  +  GK
Sbjct: 264 APIKFIGTGERIDDLEVFDPERFISRLLGMGDLRSLLEKVEEVSEEEIAEESLDAILSGK 323

Query: 168 FSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           F++RD++ QF  + K+G L Q MS +PG G   L K   + + + I+++L IM+SM
Sbjct: 324 FTLRDMRVQFEMMGKMGPLQQVMSMLPGAGK--LPKDASRMTEETIRKYLIIMDSM 377



>gi|14520695|ref|NP_126170.1|  (NC_000868) signal recognition particle, subunit SRP54
[Pyrococcus abyssi]
 gi|12230641|sp|Q9V1E8|SR54_PYRAB  Signal recognition 54 kDa protein (SRP54)
 gi|7440081|pir||B75165  signal recognition particle, chain srp54 (srp54) PAB0320 -
Pyrococcus abyssi (strain Orsay)
```

```
 gi|5457911|emb|CAB49401.1|  (AJ248284) signal recognition particle, subunit SRP54 (srp54)
[Pyrococcus abyssi]
          Length = 443

 Score = 89.0 bits (219), Expect = 9e-18
 Identities = 51/125 (40%), Positives = 77/125 (60%), Gaps = 5/125 (4%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV-VNMRDQPQLIQK 184
           AAT +PI FIG GE  DDLE F+ + FV RLLGLGDI  L +  K++   +    + +++
Sbjct: 263 AATGAPIKFIGTGEKIDDLEPFDPARFVSRLLGLGDIQGLLEKFKELEKEVEFTEEDLER 322

Query: 183 LKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSS----VLSKGNEKESIKRIQRFLC 16
             +GKF+++D+  Q  ++ K+G L Q +  IPG+G S    V+S G E     R+++F
Sbjct: 323 FLKGKFTLKDMYAQLEAMRKMGPLKQILRMIPGLGYSLPDEVISVGEE-----RLRKFKV 377

Query: 15  IMNSM 1
           IM+SM
Sbjct: 378 IMDSM 382




>gi|15594157|dbj|BAB64926.1|  (AB057373) signal recognition particle protein 54 [Pyrococcus
furiosus]
          Length = 443

 Score = 88.6 bits (218), Expect = 1e-17
 Identities = 52/127 (40%), Positives = 78/127 (60%), Gaps = 7/127 (5%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV---VNMRDQPQLI 190
           AAT +PI FIG GE  DD+E F+   FV RLLGLGDI  L +  K++   V ++++   I
Sbjct: 263 AATGAPIKFIGTGEKIDDIEPFDPPRFVSRLLGLGDIQGLLEKFKELEKEVEIKEED--I 320

Query: 189 QKLKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSS----VLSKGNEKESIKRIQRF 22
           ++   GKF+++D+  Q  ++ K+G L Q +  IPG+G S    V+S G E     R+++F
Sbjct: 321 ERFLRGKFTLKDMYAQLEAMRKMGPLKQILRMIPGLGYSLPDDVISIGEE-----RLKKF 375

Query: 21  LCIMNSM 1
             IM+SM
Sbjct: 376 KVIMDSM 382




>gi|18978103|ref|NP_579460.1|  (NC_003413) signal recognition particle protein srp54
[Pyrococcus furiosus DSM 3638]
 gi|18893898|gb|AAL81855.1|  (AE010270) signal recognition particle protein srp54
[Pyrococcus furiosus DSM 3638]
          Length = 443

 Score = 88.6 bits (218), Expect = 1e-17
 Identities = 52/127 (40%), Positives = 78/127 (60%), Gaps = 7/127 (5%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV---VNMRDQPQLI 190
           AAT +PI FIG GE  DD+E F+   FV RLLGLGDI  L +  K++   V ++++   I
Sbjct: 263 AATGAPIKFIGTGEKIDDIEPFDPPRFVSRLLGLGDIQGLLEKFKELEKEVEIKEED--I 320
```

```
Query: 189 QKLKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSS----VLSKGNEKESIKRIQRF 22
           ++   GKF+++D+  Q  ++ K+G L Q +  IPG+G S    V+S G E     R+++F
Sbjct: 321 ERFLRGKFTLKDMYAQLEAMRKMGPLKQILRMIPGLGYSLPDDVISIGEE-----RLKKF 375

Query: 21  LCIMNSM 1
             IM+SM
Sbjct: 376 KVIMDSM 382



>gi|12230630|sp|O59307|SR54_PYRHO  Signal recognition 54 kDa protein (SRP54)
          Length = 443

 Score = 87.0 bits (214), Expect = 3e-17
 Identities = 51/125 (40%), Positives = 75/125 (59%), Gaps = 5/125 (4%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV-VNMRDQPQLIQK 184
           AAT +PI FIG GE  DDLE F+ + FV RLLGLGDI  L +  K++   +    + I +
Sbjct: 263 AATGAPIKFIGVGEKIDDLEPFDPARFVSRLLGLGDIQGLLEKFKELEKEVEFTEEDIDR 322

Query: 183 LKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSS----VLSKGNEKESIKRIQRFLC 16
              GKF+++D+  Q  ++ K+G L Q +  IPG+G S    ++S G E     R+++F
Sbjct: 323 FLRGKFTLKDMYAQLEAMRKMGPLKQILRMIPGLGYSLPDELISVGEE-----RLRKFKV 377

Query: 15  IMNSM 1
           IM+SM
Sbjct: 378 IMDSM 382



>gi|14591457|ref|NP_143537.1|  (NC_000961) signal recognition particle protein [Pyrococcus
horikoshii]
 gi|7440073|pir||H71176  probable signal recognition particle protein - Pyrococcus
horikoshii
          Length = 445

 Score = 87.0 bits (214), Expect = 3e-17
 Identities = 51/125 (40%), Positives = 75/125 (59%), Gaps = 5/125 (4%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV-VNMRDQPQLIQK 184
           AAT +PI FIG GE  DDLE F+ + FV RLLGLGDI  L +  K++   +    + I +
Sbjct: 265 AATGAPIKFIGVGEKIDDLEPFDPARFVSRLLGLGDIQGLLEKFKELEKEVEFTEEDIDR 324

Query: 183 LKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSS----VLSKGNEKESIKRIQRFLC 16
              GKF+++D+  Q  ++ K+G L Q +  IPG+G S    ++S G E     R+++F
Sbjct: 325 FLRGKFTLKDMYAQLEAMRKMGPLKQILRMIPGLGYSLPDELISVGEE-----RLRKFKV 379

Query: 15  IMNSM 1
           IM+SM
Sbjct: 380 IMDSM 384



>gi|3334340|sp|O15821|SR54_ENTHI  Signal recognition particle 54 kDa protein (SRP54)
 gi|17980209|gb|AAL50553.1|  (AF386797) signal recognition particle 54 kDa subunit SRP54
[Entamoeba histolytica]
```

```
          Length = 487

 Score = 86.3 bits (212), Expect = 6e-17
 Identities = 44/120 (36%), Positives = 75/120 (61%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+SPI+FIG GE  +++E F+A SFVR+LLG+GD+  + +  KD     +   +++ L
Sbjct: 264 AATKSPIIFIGTGEKVNEIEEFDAESFVRQLLGMGDLKGIAKLAKDFAENAEYKTMVKHL 323

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           +EG  ++RD + Q +++ K+G L   M  I G+   +   GN +   K+ + F+ I++SM
Sbjct: 324 QEGTLTVRDWKEQLSNLQKMGQLGNIMQMI-GLNHPMFQGGNIE---KKFKVFMVILDSM 379



>gi|15668272|ref|NP_247065.1|  (NC_000909) signal recognition particle, subunit SRP54
           [Methanococcus jannaschii]
 gi|2500886|sp|Q57565|SR54_METJA  Signal recognition 54 kDa protein (SRP54)
 gi|2129283|pir||E64312  signal recognition particle protein - Methanococcus jannaschii
 gi|1590878|gb|AAB98081.1|  (U67467) signal recognition particle, subunit SRP54
[Methanococcus jannaschii]
          Length = 451

 Score = 85.9 bits (211), Expect = 8e-17
 Identities = 43/120 (35%), Positives = 75/120 (61%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           A T++PI FIG GE  DDLE F+   F+ RLLG+GD+  L +  +D+V+ + + + I  +
Sbjct: 263 AETKAPIKFIGIGEGIDDLEPFDPKKFISRLLGMGDLESLLEKAEDMVDEKTE-ESIDAI 321

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
             GKF++ +L TQ  ++  +GS+ + +S IPG G ++  + +     K I+++  I++SM
Sbjct: 322 MRGKFTLNELMTQLEAIENMGSMKKILSMIPGFGGAMPKELSHLTEAK-IKKYKVIISSM 380



>gi|20095048|ref|NP_614895.1|  (NC_003551) Signal recognition particle GTPase [Methanopyrus
kandleri AV19]
 gi|19888324|gb|AAM02825.1|  (AE010451) Signal recognition particle GTPase [Methanopyrus
kandleri AV19]
          Length = 447

 Score = 84.7 bits (208), Expect = 2e-16
 Identities = 42/120 (35%), Positives = 76/120 (63%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           A T +PI F+G GE  DDLE F   SFV RLLG+GDI++L +  ++   M ++ +  + +
Sbjct: 262 ARTGAPIKFVGTGERVDDLEEFNPRSFVARLLGIGDIDELLRRTEE---MLEEEEKAEDV 318

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
            EG+F+++DL  Q  ++ K+G +++ +  +PGMG     +   + + +R++++  IM+SM
Sbjct: 319 LEGEFTLKDLYEQLEALSKMGPVDKLLQYVPGMGGGRNVRKISQITEERLKKYKVIMDSM 378
```

```
>gi|19074175|ref|NP_584781.1|  (NC_003231) SIGNAL RECOGNITION PARTICLE 54kDa SUBUNIT (SRP54)
[Encephalitozoon cuniculi]
 gi|19068817|emb|CAD25285.1|  (AL590444) SIGNAL RECOGNITION PARTICLE 54kDa SUBUNIT (SRP54)
[Encephalitozoon cuniculi]
          Length = 466

 Score = 82.8 bits (203), Expect = 6e-16
 Identities = 38/92 (41%), Positives = 62/92 (67%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT+ PI F+G GE  DDLE+F+A  FV R+LG+GD+  L + V  +    D+ ++++KL
Sbjct: 261 AATKCPIEFVGTGEGMDDLEAFDARRFVSRMLGMGDVEGLMEKVGSL--GIDEKEVVKKL 318

Query: 180 KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPG 85
           ++G+F++ D   QF  +L LG +++ +  IPG
Sbjct: 319 RQGRFTLGDFYDQFQKILSLGPISKLLEMIPG 350




>gi|3334345|sp|Q00179|SR54_ASPNG  SIGNAL RECOGNITION PARTICLE 54 KD PROTEIN HOMOLOG
 gi|2119054|pir||JC4572  signal recognition particle 54K protein homolog - Aspergillus niger
 gi|598061|gb|AAB04946.1|  (L38317) srpA gene product [Aspergillus niger]
 gi|1586559|prf||2204256A  srpA gene [Aspergillus niger]
          Length = 534

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 41/121 (33%), Positives = 74/121 (60%), Gaps = 1/121 (0%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVV-NMRDQPQLIQK 184
           AAT +PI+++G GEH  DLE FE  +F+++LLG+GD+  L + V+ V  +     +  +
Sbjct: 265 AATHTPIIYLGTGEHLMDLERFEPKAFIQKLLGMGDMAGLVEHVQAVTKDSASAKETYKH 324

Query: 183 LKEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESIKRIQRFLCIMNS 4
           + EG +++RD +    S++K+G L++    IPG+ +      +E  S+K ++R + I +S
Sbjct: 325 ISEGIYTLRDFRENITSIMKMGPLSKLSGMIPGLSNLTAGLDDEDGSMK-LRRMIYIFDS 383

Query: 3   M 1
           M
Sbjct: 384 M 384




>gi|15921563|ref|NP_377232.1|  (NC_003106) 356aa long hypothetical signal recognition
particle protein [Sulfolobus tokodaii]
 gi|15622349|dbj|BAB66341.1|  (AP000985) 356aa long hypothetical signal recognition particle
protein [Sulfolobus tokodaii]
          Length = 356

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 47/126 (37%), Positives = 74/126 (58%), Gaps = 6/126 (4%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT + I FIG GE  D+LE F    FV R+LG+GDI  + + +K    M +  Q+ +K+
Sbjct: 171 AATGAAIKFIGTGEKLDELEVFNPRRFVSRILGMGDIESIIEKIK---GMEEYEQIQKKM 227
```

```
Query: 180 KE-----GKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMG-SSVLSKGNEKESIKRIQRFL 19
           +E      K ++RD+  Q  ++ K+G LN+ +   +PG G  S + +    K   ++I++FL
Sbjct: 228 EEVMSGKAKLTLRDIYKQLTALRKMGPLNKILQMLPGFGVFSQIPEEQLKLGEEKIKKFL 287

Query: 18  CIMNSM 1
            IMNSM
Sbjct: 288 VIMNSM 293



>gi|20093371|ref|NP_619446.1|  (NC_003552) signal recognition particle, 54 kDa protein
[Methanosarcina acetivorans str. C2A]
 gi|19918737|gb|AAM07926.1|  (AE011180) signal recognition particle, 54 kDa protein
[Methanosarcina acetivorans str. C2A]
          Length = 440

 Score = 80.9 bits (198), Expect = 2e-15
 Identities = 41/119 (34%), Positives = 73/119 (60%), Gaps = 1/119 (0%)
 Frame = -2

Query: 354 TESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLKE 175
           T++PI FIG GE  +D E FEA  F+ RLLG+GD+  L +  ++  ++ ++   ++ L +
Sbjct: 260 TKAPIAFIGVGETPEDFEKFEADRFISRLLGMGDLKSLMEKAEE--SLSEEDVNVEALMQ 317

Query: 174 GKFSIRDLQTQFNSVLKLGSLNQFMSAIP-GMGSSVLSKGNEKESIKRIQRFLCIMNSM 1
           G+F+++D+  Q  ++ K+G L Q MS +P GMG    S    + +  +++ +  IM+SM
Sbjct: 318 GRFTLKDMYKQLEAMNKMGPLKQIMSMLPMGMGGMKFSDEMFQATSDKMKNYKVIMDSM 376



>gi|15612144|ref|NP_223796.1|  (NC_000921) putative SIGNAL RECOGNITION PARTICLE PROTEIN
[Helicobacter pylori J99]
 gi|11387213|sp|Q9ZK62|SR54_HELPJ  Signal recognition particle protein (Fifty-four homolog)
 gi|7440076|pir||D71852  probable signal recognition particle protein - Helicobacter pylori
(strain J99)
 gi|4155673|gb|AAD06659.1|  (AE001535) putative SIGNAL RECOGNITION PARTICLE PROTEIN
[Helicobacter pylori J99]
          Length = 448

 Score = 80.5 bits (197), Expect = 3e-15
 Identities = 43/116 (37%), Positives = 65/116 (55%), Gaps = 1/116 (0%)
 Frame = -2

Query: 345 PIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLKEGKF 166
           P+ FIG GE   DL+ F     V RL+G GDI  L +    V+N  +   L +KLK+G+F
Sbjct: 261 PLRFIGSGEKIPDLDVFMPERIVGRLMGAGDIISLAEKTASVLNPNEAKDLSKKLKKGQF 320

Query: 165 SIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLS-KGNEKESIKRIQRFLCIMNSM 1
           +  D   Q   V KLGS++  +S IPG+G+   + K  + ES   +++   ++NSM
Sbjct: 321 TFNDFLNQIEKVKKLGSMSSLISMIPGLGNMASALKDTDLESSLEVKKIKAMVNSM 376



>gi|15645766|ref|NP_207943.1|  (NC_000915) signal recognition particle protein (ffh)
[Helicobacter pylori 26695]
 gi|2500881|sp|P56005|SR54_HELPY  Signal recognition particle protein (Fifty-four homolog)
 gi|7440077|pir||H64663  signal recognition particle protein - Helicobacter pylori  (strain
26695)
```

```
 gi|2314305|gb|AAD08194.1|  (AE000621) signal recognition particle protein (ffh)
[Helicobacter pylori 26695]
          Length = 448

 Score = 80.5 bits (197), Expect = 3e-15
 Identities = 43/116 (37%), Positives = 65/116 (55%), Gaps = 1/116 (0%)
 Frame = -2

Query: 345 PIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLKEGKF 166
           P+ FIG GE   DL+ F     V RL+G GDI  L +    V+N  +   L +KLK+G+F
Sbjct: 261 PLRFIGSGEKIPDLDVFVPERIVGRLMGAGDIVSLAEKTASVLNPNEAKDLSKKLKKGQF 320

Query: 165 SIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLS-KGNEKESIKRIQRFLCIMNSM 1
           +  D   Q   V KLGS++  +S IPG+G+   + K  + ES   +++   ++NSM
Sbjct: 321 TFNDFLNQIEKVKKLGSMSSLISMIPGLGNMASALKDTDLESSLEVKKIKAMVNSM 376



>gi|12230633|sp|P70722|SR54_ACIAM  Signal recognition 54 kDa protein (SRP54)
 gi|4165432|emb|CAA69991.1|  (Y08735) fifty-four homologue of SRP54 [Acidianus ambivalens]
          Length = 451

 Score = 78.2 bits (191), Expect = 2e-14
 Identities = 42/123 (34%), Positives = 74/123 (60%), Gaps = 3/123 (2%)
 Frame = -2

Query: 360 AATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKL 181
           AAT + I FIG GE  D+LE F    FV R+LG+GDI  + + +K+V N     + ++++
Sbjct: 263 AATGATIKFIGTGEKIDELEVFNPRRFVARILGMGDIETILEKIKEVENYDKMQKKMEEV 322

Query: 180 --KEGKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGS-SVLSKGNEKESIKRIQRFLCIM 10
              +GK ++RD+  Q  ++ K+G L++    +PG+G    + +   K   ++++++L IM
Sbjct: 323 ISGKGKLTLRDVYNQLIALRKMGPLSKLFQLLPGIGMLGQIPEDQLKVGEEKMRKWLAIM 382

Query: 9   NSM 1
           NSM
Sbjct: 383 NSM 385



>gi|15897850|ref|NP_342455.1|  (NC_002754) Signal recognition particle protein subunit SRP54
(srp54) [Sulfolobus solfataricus]
 gi|13814155|gb|AAK41245.1|  (AE006717) Signal recognition particle protein subunit SRP54
(srp54) [Sulfolobus solfataricus]
          Length = 447

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 46/127 (36%), Positives = 75/127 (58%), Gaps = 8/127 (6%)
 Frame = -2

Query: 357 ATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLK 178
           AT + I FIG GE  D+LE+F A  FV R+LG+GDI  + + VK +    ++   IQK
Sbjct: 262 ATGATIKFIGTGEKIDELETFNAKRFVSRILGMGDIESILEKVKGL----EEYDKIQKKM 317

Query: 177 E------GKFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKESI--KRIQRF 22
           E      GK ++RD+  Q  ++ K+G L++ +   IPG+G  + +   ++  I  ++I+R+
Sbjct: 318 EDVMEGKGKLTLRDVYAQIIALRKMGPLSKVLQHIPGLGIMLPTPSEDQLKIGEEKIRRW 377
```

```
Query: 21  LCIMNSM 1
           L  +NSM
Sbjct: 378 LAALNSM 384




>gi|13541823|ref|NP_111511.1|  (NC_002689) Signal recognition particle GTPase [Thermoplasma
volcanium]
 gi|14325260|dbj|BAB60164.1|  (AP000995) signal recognition particle protein srp54
[Thermoplasma volcanium]
          Length = 455

 Score = 77.0 bits (188), Expect = 4e-14
 Identities = 40/119 (33%), Positives = 70/119 (58%), Gaps = 4/119 (3%)
 Frame = -2

Query: 345 PIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV-VNMRDQPQLIQKLKEGK 169
           PI FIG GEH DD+E F+   F+ RLLGLGD+  LF++VK+  +   +  +  +KL  GK
Sbjct: 263 PIYFIGTGEHMDDMEIFDPKKFLSRLLGLGDLETLFETVKEANITEEEAQKSFEKLMTGK 322

Query: 168 FSIRDLQTQFNSVLKLGSLNQFMSAIPGM---GSSVLSKGNEKESIKRIQRFLCIMNSM 1
           F+++D+   +    + G + + + ++P     GS  ++      + +++QR+  IM+SM
Sbjct: 323 FNLKDMYDVWEKFSRPGLMKKLVDSLPLAKLPGSQKINDDQILTAEQKLQRYRIIMDSM 381




>gi|16081627|ref|NP_393991.1|  (NC_002578) probable signal recognition particle protein
[Thermoplasma acidophilum]
 gi|12230635|sp|Q9HKT0|SR54_THEAC  Signal recognition 54 kDa protein (SRP54)
 gi|10639683|emb|CAC11655.1|  (AL445064) probable signal recognition particle protein
[Thermoplasma acidophilum]
          Length = 456

 Score = 76.3 bits (186), Expect = 6e-14
 Identities = 40/119 (33%), Positives = 68/119 (56%), Gaps = 4/119 (3%)
 Frame = -2

Query: 345 PIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDV-VNMRDQPQLIQKLKEGK 169
           PI FIG GEH DDLE F+   F+ RLLGLGD+  LF++V++  +   +  +  +KL  GK
Sbjct: 263 PIYFIGTGEHMDDLEVFDPKKFLSRLLGLGDLESLFETVQEADITEEEAQESFEKLMTGK 322

Query: 168 FSIRDLQTQFNSVLKLGSLNQFMSAIPGM---GSSVLSKGNEKESIKRIQRFLCIMNSM 1
           F+++D+   +    K G + + + A+P     GS  +     + +  +++ +  IM+SM
Sbjct: 323 FNLKDMYDVWEKFSKPGLMKKLVDALPLARIPGSQKIDDSKIQSAEDKLRMYRIIMDSM 381




>gi|12230625|sp|O07853|SR54_SULAC  Signal recognition 54 kDa protein (SRP54)
 gi|2231060|emb|CAA73234.1|  (Y12702) fifty-four homologue of SRP54 [Sulfolobus
acidocaldarius]
          Length = 446

 Score = 75.9 bits (185), Expect = 8e-14
 Identities = 45/124 (36%), Positives = 72/124 (57%), Gaps = 5/124 (4%)
 Frame = -2

Query: 357 ATESPIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLK 178
           AT + I FIG GE  D+LE F    FV RLLGLGDI  + + +K V +  +  + ++ +
Sbjct: 263 ATGAQIKFIGTGEKLDELEVFNPRRFVSRLLGLGDIESIIEKIKSVEDYENLEKRMEDVI 322
```

```
Query: 177 EG--KFSIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKE---SIKRIQRFLCI 13
            G  K ++RD+  Q  ++ K+G L +    +PG G  +LS+  E++     ++I+ F+ I
Sbjct: 323 SGKTKLTLRDIYKQLIALRKMGPLGKIFQMLPGAG--ILSQVPEEQLKLGEEKIRTFMAI 380

Query: 12  MNSM 1
           MNSM
Sbjct: 381 MNSM 384



>gi|15595039|ref|NP_212828.1|  (NC_001318) signal recognition particle protein (ffh)
[Borrelia burgdorferi]
 gi|7440078|pir||E70186  signal recognition particle protein (ffh) homolog - Lyme disease
           spirochete
 gi|2688634|gb|AAC67049.1|  (AE001170) signal recognition particle protein (ffh) [Borrelia
burgdorferi]
          Length = 447

 Score = 75.9 bits (185), Expect = 8e-14
 Identities = 35/100 (35%), Positives = 59/100 (59%)
 Frame = -2

Query: 345 PIVFIGEGEHFDDLESFEASSFVRRLLGLGDINKLFQSVKDVVNMRDQPQLIQKLKEGKF 166
           PI FIG GE  +DL+SF       R+LG+GD+  L + V+ VV+   +  +L +K+ +  F
Sbjct: 272 PIKFIGVGEKIEDLDSFYPERIASRILGMGDVVSLVEKVQSVVDKEEAIKLEEKINKASF 331

Query: 165 SIRDLQTQFNSVLKLGSLNQFMSAIPGMGSSVLSKGNEKE 46
           +  D  +QF  + ++G  + F+S +PG+  S+L+  N  E
Sbjct: 332 NFEDYLSQFRRIRQVGGFSNFVSFLPGVSKSMLNSNNLNE 371




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 8, 2002  5:26 AM
  Number of letters in database: 292,803,047
  Number of sequences in database:  933,889

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 179,593,055
Number of Sequences: 933889
Number of extensions: 3407747
Number of successful extensions: 12808
Number of sequences better than 10.0: 367
Number of HSP's better than 10.0 without gapping: 12564
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 12759
```

```
length of database: 292,803,047
effective HSP length: 95
effective length of database: 204,083,592
effective search space used: 4898006208
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020952001-010439-23814
```

7.1.5 Query= **hy-5_11b250002**

```
         (393 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 209 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                        (bits) Value

gi|2271461|gb|AAC13355.1|  (AF009561) calcium-dependent prot...     93   6e-19
gi|1279425|emb|CAA96439.1|  (Z71757) calmodulin-domain prote...     89   8e-18
gi|1279423|emb|CAA96438.1|  (Z71756) calmodulin-domain prote...     89   1e-17
gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. prot...     79   9e-15
gi|2854042|gb|AAC02532.1|  (AF043629) protein kinase 4 [Toxo...     78   2e-14
gi|12484153|gb|AAG53993.1|AF333958_1  (AF333958) calmodulin-...     78   2e-14
gi|13561063|emb|CAA65500.1|  (X96723) protein kinase [Medica...     75   1e-13
gi|15223629|ref|NP_175485.1|  (NM_103952) calcium-dependent ...     75   2e-13
gi|15231140|ref|NP_188676.1|  (NM_112932) calmodulin-domain ...     74   3e-13
gi|15220658|ref|NP_176386.1|  (NM_104875) calcium-dependent ...     74   4e-13
gi|7428012|pir||T02139  calcium-dependent protein kinase (EC...     74   4e-13
gi|15222931|ref|NP_177731.1|  (NM_106253) calcium-dependent ...     73   5e-13
gi|6721111|gb|AAF26765.1|AC007396_14  (AC007396) T4O12.25 [A...     73   5e-13
gi|20260246|gb|AAM13021.1|  (AY093022) calcium-dependent pro...     73   5e-13
gi|13774101|gb|AAK38161.1|  (AY030280) calcium-dependent pro...     73   6e-13
gi|116054|sp|P28583|CDPK_SOYBN  Calcium-dependent protein ki...     72   1e-12
gi|2271459|gb|AAC13354.1|  (AF009560) calcium-dependent prot...     70   3e-12
gi|7434371|pir||T08874  calcium-dependent protein kinase (EC...     70   3e-12
gi|19171502|emb|CAC87494.1|  (AJ308296) calcium-dependent pr...     70   5e-12
gi|7428008|pir||T03024  calcium-dependent protein kinase (EC...     70   5e-12
gi|15234656|ref|NP_193925.1|  (NM_118315) calcium-dependent ...     69   7e-12
gi|15236560|ref|NP_194096.1|  (NM_118496) calcium-dependent ...     69   7e-12
gi|15234435|ref|NP_192381.1|  (NM_116710) putative calcium d...     69   7e-12
gi|15292915|gb|AAK92828.1|  (AY050891) putative calcium depe...     69   7e-12
```

```
gi|4115945|gb|AAD03455.1|  (AF118223) contains similarity to...   69   9e-12
gi|7428009|pir||T01989  calcium-dependent protein kinase (EC...   69   1e-11
gi|15450437|gb|AAK96512.1|  (AY052319) AT4g23650/F9D16_120 [...   69   1e-11
gi|6469601|gb|AAF13351.1|AF121337_1  (AF121337) calcium/calm...   69   1e-11
gi|2827773|sp|P28582|CDPK_DAUCA  Calcium-dependent protein k...   68   2e-11
gi|82056|pir||S17759  protein kinase, calcium-dependent (EC ...   68   2e-11
gi|12313676|dbj|BAB21081.1|  (AP002819) putative calcium-dep...   67   3e-11
gi|4336426|gb|AAD17800.1|  (AF090835) Ca2+-dependent protein...   67   3e-11
gi|15234441|ref|NP_192383.1|  (NM_116712) putative calcium d...   67   3e-11
gi|13872969|dbj|BAB44074.1|  (AP003073) putative calcium-dep...   67   3e-11
gi|7428003|pir||S71770  calcium-dependent protein kinase (EC...   67   4e-11
gi|1705734|sp|P53683|CDP2_ORYSA  Calcium-dependent protein k...   67   4e-11
gi|15289760|dbj|BAB63464.1|  (AB051809) calcium dependent pr...   66   6e-11
gi|15236966|ref|NP_195257.1|  (NM_119697) calmodulin-domain ...   66   6e-11
gi|5162877|dbj|BAA81748.1|  (AB017515) calcium-dependent pro...   66   6e-11
gi|5162878|dbj|BAA81749.1|  (AB017515) calcium-dependent pro...   66   6e-11
gi|15234432|ref|NP_192380.1|  (NM_116709) putative calcium d...   66   8e-11
gi|7428004|pir||T03271  calcium-dependent protein kinase (EC...   65   1e-10
gi|14970920|emb|CAC44471.1|  (AJ320242) calcium dependent ca...   65   1e-10
gi|5732059|gb|AAD48958.1|AF149414_7  (AF149414) similar to P...   65   1e-10
gi|16754824|emb|CAC83060.1|  (AJ298072) calcium dependent ca...   65   1e-10
gi|1345719|sp|P49101|CDP2_MAIZE  Calcium-dependent protein k...   65   1e-10
gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falci...   65   1e-10
gi|10568116|gb|AAD28192.2|  (AF115406) calcium-dependent pro...   65   2e-10
gi|14029712|gb|AAK52801.1|AF363784_1  (AF363784) calcium-dep...   65   2e-10
gi|18076255|emb|CAC79947.1|  (AJ278957) protein kinase [Nyct...   65   2e-10
gi|7434369|pir||T10938  calcium-dependent protein kinase (EC...   64   2e-10
gi|7428011|pir||T02993  calcium-dependent protein kinase (EC...   64   2e-10
gi|18416872|ref|NP_568281.1|  (NM_121286) calcium-dependent ...   64   3e-10
gi|18398377|ref|NP_565411.1|  (NM_127284) putative calmoduli...   64   3e-10
gi|603473|dbj|BAA05918.1|  (D28582) calcium-dependent protei...   64   3e-10
gi|1399275|gb|AAB03246.1|  (U31835) calmodulin-domain protei...   64   3e-10
gi|2944385|gb|AAC05270.1|  (AF048691) calcium dependent prot...   64   4e-10
gi|5706728|gb|AAD03451.2|  (AF118223) contains similarity to...   64   4e-10
gi|2982257|gb|AAC32116.1|  (AF051211) probable calcium depen...   64   4e-10
gi|6063536|dbj|BAA85396.1|  (AP000615) ESTs C22369(C12239),C...   63   5e-10
gi|1705735|sp|P53684|CDP3_ORYSA  Calcium-dependent protein k...   63   5e-10
gi|15239716|ref|NP_197437.1|  (NM_121941) calcium-dependent ...   63   5e-10
gi|1084335|pir||S46284  calcium-dependent protein kinase (EC...   63   6e-10
gi|15219693|ref|NP_174807.1|  (NM_103271) calcium-dependent ...   63   6e-10
gi|8778378|gb|AAF79386.1|AC007887_45  (AC007887) F15O4.8 [Ar...   63   6e-10
gi|15227525|ref|NP_181133.1|  (NM_129148) putative calcium-d...   63   6e-10
gi|7428005|pir||T03263  calcium-dependent protein kinase (EC...   62   8e-10
gi|20453015|gb|AAL68972.1|  (AY072802) calmodulin-like-domai...   62   8e-10
gi|18175903|gb|AAL59948.1|  (AY072126) putative calcium-depe...   62   1e-09
gi|15239888|ref|NP_196779.1|  (NM_121256) calcium-dependent ...   62   1e-09
gi|7434364|pir||T09940  calcium-dependent protein kinase (EC...   62   1e-09
gi|20453013|gb|AAL68971.1|  (AY072801) phloem calmodulin-lik...   62   1e-09
gi|2129552|pir||S71777  calcium-dependent protein kinase 19 ...   62   1e-09
gi|15239742|ref|NP_197446.1|  (NM_121950) calcium-dependent ...   62   1e-09
gi|15231060|ref|NP_190753.1|  (NM_115044) calcium-dependent ...   61   2e-09
gi|15233947|ref|NP_192695.1|  (NM_117025) calmodulin-domain ...   61   2e-09
gi|10944296|dbj|BAB16888.1|  (AB042550) OsCDPK7 [Oryza sativ...   61   2e-09
gi|15224978|ref|NP_181425.1|  (NM_129449) putative calcium-d...   61   2e-09
gi|17064926|gb|AAL32617.1|  (AY062539) calcium-dependent pro...   61   2e-09
gi|6644464|gb|AAF21062.1|  (AF216527) calcium-dependent prot...   61   2e-09
gi|15237791|ref|NP_197748.1|  (NM_122264) calcium-dependent ...   61   2e-09
gi|15289758|dbj|BAB63463.1|  (AB051808) calcium dependent pr...   60   3e-09
gi|2129555|pir||S71775  calcium-dependent protein kinase 9 -...   60   3e-09
gi|16215467|emb|CAC82998.1|  (AJ344154) calcium-dependent pr...   60   4e-09
gi|7428007|pir||T02784  calcium-dependent protein kinase (EC...   60   4e-09
```

```
gi|15233737|ref|NP_195536.1|  (NM_119985) calcium-dependent ...    60   4e-09
gi|16215471|emb|CAC82999.1|  (AJ344155) calcium-dependent pr...    60   5e-09
gi|7434370|pir||T08873  calcium-dependent protein kinase (EC...    59   7e-09
gi|1362190|pir||S56717  calcium-dependent protein kinase (EC...    59   7e-09
gi|2315983|gb|AAB70706.1|  (U82087) calmodulin-like domain p...    59   7e-09
gi|15227452|ref|NP_181717.1|  (NM_129750) putative calcium-d...    59   9e-09
gi|5326544|emb|CAB46228.1|  (Y18055) calcium dependent prote...    59   1e-08
gi|15225092|ref|NP_180708.1|  (NM_128707) putative calcium-d...    59   1e-08
gi|15228350|ref|NP_187677.1|  (NM_111902) calmodulin-domain ...    58   2e-08
gi|16215475|emb|CAC83000.1|  (AJ344156) calcium-dependent pr...    58   2e-08
gi|6730697|gb|AAF27092.1|AC011809_1  (AC011809) calcium-depe...    58   2e-08
gi|14484895|gb|AAK62812.1|  (AF276999) calcium-dependent pro...    58   2e-08
gi|1084334|pir||S46283  calcium-dependent protein kinase (EC...    58   2e-08
gi|18394667|ref|NP_564066.1|  (NM_101746) calcium-dependent ...    58   2e-08
gi|15230332|ref|NP_191312.1|  (NM_115613) calcium-dependent ...    57   3e-08

                                   Alignments


>gi|2271461|gb|AAC13355.1|  (AF009561) calcium-dependent protein kinase-b [Paramecium
tetraurelia]
          Length = 493

 Score = 92.8 bits (229), Expect = 6e-19
 Identities = 47/102 (46%), Positives = 65/102 (63%)
 Frame = +2

Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQINSLTLSKLQDFDSKNKLK 181
           K+S   K+LI+RML  +   RISA +A N  WI+ N   G I+   +  L  F  KNK++
Sbjct: 279 KISEDAKNLIKRMLTKDYQLRISAQEAYNDPWIQKNAPNGPIDMKAIKNLSSFFGKNKVR 338

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQ 307
            A+ QFIT  ++TN EK  LL  F++IDK+G+G +SKDEL Q
Sbjct: 339 AALMQFITTNLMTNTEKEGLLNEFKKIDKDGNGQISKDELLQ 380



>gi|1279425|emb|CAA96439.1|  (Z71757) calmodulin-domain protein kinase [Eimeria tenella]
          Length = 487

 Score = 89.0 bits (219), Expect = 8e-18
 Identities = 49/108 (45%), Positives = 65/108 (59%), Gaps = 5/108 (4%)
 Frame = +2

Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKG---QINSL--TLSKLQDFDS 166
           KVS   KDLIR+ML Y PT RISA DAL HEW+K   +      + SL  T+  ++ F
Sbjct: 253 KVSEPAKDLIRKMLAYVPTMRISARDALEHEWLKTTDAATDSIDVPSLESTILNIRQFQG 312

Query: 167 KNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQG 310
             KL  A   ++  ++ TN+E  +L KIFQ +DKNGDG + K EL +G
Sbjct: 313 TQKLAAAALLYMGSKLTTNEETVELNKIFQRMDKNGDGQLDKQELMEG 360




>gi|1279423|emb|CAA96438.1|  (Z71756) calmodulin-domain protein kinase [Eimeria maxima]
          Length = 414

 Score = 88.6 bits (218), Expect = 1e-17
```

```
 Identities = 49/109 (44%), Positives = 67/109 (60%), Gaps = 6/109 (5%)
 Frame = +2

Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNK--SKGQIN----SLTLSKLQDFD 163
           KVS   KDLIR+ML Y P+ RISA DAL+H WIK+    +K  IN      T+  ++ F
Sbjct: 178 KVSEPAKDLIRKMLAYVPSMRISAKDALDHPWIKSTDVTAKDSINLPSLESTILNIRQFQ 237

Query: 164 SKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQG 310
              KL  A   ++  ++ TN+E  +L KIFQ++DKNGDG + K EL +G
Sbjct: 238 GTQKLAAAALLYMGSKLTTNEETDELNKIFQKMDKNGDGQLDKQELMEG 286




 Score = 29.3 bits (64), Expect = 7.9
 Identities = 14/49 (28%), Positives = 24/49 (48%)
 Frame = +2

Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDEL 301
           DFD    ++Y+ F  + +   T   +  L + F+  D +G G +S  EL
Sbjct: 318 DFDKNGFIEYSEFVTVAMDRRTLLSRQRLERAFEMFDSDGSGKISSSEL 366



>gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. protein kinase (C-term. EF hand)
[Plasmodium falciparum]
 gi|422320|pir||A45472  protein kinase (EC 2.7.1.37) - malaria parasite (Plasmodium
falciparum)
 gi|9878|emb|CAA47704.1|  (X67288) protein kinase [Plasmodium falciparum]
 gi|3845284|gb|AAC71952.1|  (AE001419) calcium-dept. protein kinase (C-term. EF hand)
[Plasmodium falciparum 3D7]
          Length = 524

 Score = 79.0 bits (193), Expect = 9e-15
 Identities = 46/136 (33%), Positives = 74/136 (53%), Gaps = 17/136 (12%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK---NNKSKGQINSL--TLSKLQDFDSK 169
           +S   K+LI+ ML Y+   RI+A +ALN +WIK   NN +K    +L   LS ++ F+
Sbjct: 294 ISEEAKELIKLMLTYDYNKRITAKEALNSKWIKKYANNINKSDQKTLCGALSNMRKFEGS 353

Query: 170 NKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ----------- 316
            KL  A   FI  ++ T +E+ +L  IF+++DKNGDG + K EL +G
Sbjct: 354 QKLAQAAILFIGSKLTTLEERKELTDIFKKLDKNGDGQLDKKELIEGYNILRSFKNELGE 413

Query: 317 -NNKQQPIQKFIKEIN 361
             N ++ +   +KE++
Sbjct: 414 LKNVEEEVDNILKEVD 429



>gi|2854042|gb|AAC02532.1|  (AF043629) protein kinase 4 [Toxoplasma gondii]
          Length = 501

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 44/109 (40%), Positives = 64/109 (58%), Gaps = 6/109 (5%)
 Frame = +2
```

```
Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQIN------SLTLSKLQDFD 163
           KVS S KDLIR+ML Y P+ RISA DAL+HEWI+   +K QI+         +  ++ F
Sbjct: 276 KVSESAKDLIRKMLTYVPSMRISARDALDHEWIQ-TYTKEQISVDVPSLDNAILNIRQFQ 334

Query: 164 SKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQG 310
              KL  A   ++  ++ +  E  +L  IF ++DKNGDG + + EL +G
Sbjct: 335 GTQKLAQAALLYMGSKLTSQDETKELTAIFHKMDKNGDGQLDRAELIEG 383




 Score = 30.0 bits (66), Expect = 4.6
 Identities = 14/46 (30%), Positives = 23/46 (49%)
 Frame = +2

Query: 158 FDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKD 295
           FDS N  K +  +  T+  V++ +      +  E+DKN DG +  D
Sbjct: 450 FDSDNSGKISSTELATIFGVSDVDSETWKSVLSEVDKNNDGEIDFD 495



>gi|12484153|gb|AAG53993.1|AF333958_1  (AF333958) calmodulin-domain protein kinase 1
[Toxoplasma gondii]
          Length = 507

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 44/109 (40%), Positives = 64/109 (58%), Gaps = 6/109 (5%)
 Frame = +2

Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQIN------SLTLSKLQDFD 163
           KVS S KDLIR+ML Y P+ RISA DAL+HEWI+   +K QI+         +  ++ F
Sbjct: 276 KVSESAKDLIRKMLTYVPSMRISARDALDHEWIQ-TYTKEQISVDVPSLDNAILNIRQFQ 334

Query: 164 SKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQG 310
              KL  A   ++  ++ +  E  +L  IF ++DKNGDG + + EL +G
Sbjct: 335 GTQKLAQAALLYMGSKLTSQDETKELTAIFHKMDKNGDGQLDRAELIEG 383




 Score = 33.9 bits (76), Expect = 0.32
 Identities = 17/50 (34%), Positives = 25/50 (50%)
 Frame = +2

Query: 158 FDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQ 307
           FDS N  K +  +  T+  V++ +      +  E+DKN DG V  DE  Q
Sbjct: 450 FDSDNSGKISSTELATIFGVSDVDSETWKSVLSEVDKNNDGEVDFDEFQQ 499



>gi|13561063|emb|CAA65500.1|  (X96723) protein kinase [Medicago sativa]
          Length = 542

 Score = 75.5 bits (184), Expect = 1e-13
 Identities = 42/120 (35%), Positives = 70/120 (58%), Gaps = 1/120 (0%)
 Frame = +2
```

```
Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIK-NNKSKGQINSLTLSKLQDFDSKNKL 178
           K+S+S KDL+R+ML+  P  RI+AA  L H WIK  N S   I+S  LS+++ F + NKL
Sbjct: 314 KISDSAKDLVRKMLIQEPKKRITAAQVLEHPWIKGGNASDKPIDSAVLSRMKQFRAMNKL 373

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L  +F  +D +  GT++ +EL  G+Q    +  +  +K++
Sbjct: 374 KKLALKVI-AENMSEEEIRGLKAMFTNMDTDNSGTITYEELKAGLQRLGSKLSEAEVKQL 432



>gi|15223629|ref|NP_175485.1|  (NM_103952) calcium-dependent protein kinase [Arabidopsis
thaliana]
 gi|12322336|gb|AAG51192.1|AC079279_13  (AC079279) calcium-dependent protein kinase
[Arabidopsis thaliana]
          Length = 521

 Score = 74.7 bits (182), Expect = 2e-13
 Identities = 42/120 (35%), Positives = 68/120 (56%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +SNS KDL+RRML  +P  RISAA+ L H W++     S   I+S  LS+++ F + NKL
Sbjct: 300 ISNSAKDLVRRMLTQDPKRRISAAEVLKHPWLREGGEASDKPIDSAVLSRMKQFRAMNKL 359

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I   + T +E   L  +F  ID +  GT++ +EL +G+     +  +  +K++
Sbjct: 360 KKLALKVIAENIDT-EEIQGLKAMFANIDTDNSGTITYEELKEGLAKLGSRLTEAEVKQL 418




 Score = 30.8 bits (68), Expect = 2.7
 Identities = 15/49 (30%), Positives = 23/49 (46%)
 Frame = +2

Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDEL 301
           D D    + Y  F   T+     +   ++ K FQ  DK+G G ++ DEL
Sbjct: 423 DVDGNGSIDYIEFITATMHRHRLESNENVYKAFQHFDKDGSGYITTDEL 471



>gi|15231140|ref|NP_188676.1|  (NM_112932) calmodulin-domain protein kinase CDPK isoform 9
[Arabidopsis thaliana]
 gi|1399265|gb|AAB03242.1|  (U31751) calmodulin-domain protein kinase CDPK isoform 9
[Arabidopsis thaliana]
 gi|9294561|dbj|BAB02824.1|  (AB024036) calmodulin-domain protein kinase CDPK isoform 9
[Arabidopsis thaliana]
          Length = 541

 Score = 73.9 bits (180), Expect = 3e-13
 Identities = 42/120 (35%), Positives = 68/120 (56%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDL+RRML  +P  RISAAD L H W++     S   I+S  LS+++ F + NKL
Sbjct: 318 ISSSAKDLVRRMLTADPKRRISAADVLQHPWLREGGEASDKPIDSAVLSRMKQFRAMNKL 377
```

```
Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I   + T +E   L  +F  ID +  GT++ +EL +G+     +  +  +K++
Sbjct: 378 KKLALKVIAENIDT-EEIQGLKAMFANIDTDNSGTITYEELKEGLAKLGSKLTEAEVKQL 436




 Score = 29.3 bits (64), Expect = 7.9
 Identities = 15/54 (27%), Positives = 24/54 (43%)
 Frame = +2

Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ 316
           D D    + Y  F   T+     +   +L K FQ  DK+  G ++ DEL   ++
Sbjct: 441 DVDGNGSIDYIEFITATMHRHRLESNENLYKAFQHFDKDSSGYITIDELESALK 494



>gi|15220658|ref|NP_176386.1|  (NM_104875) calcium-dependent protein kinase, putative
[Arabidopsis thaliana]
          Length = 547

 Score = 73.6 bits (179), Expect = 4e-13
 Identities = 41/120 (34%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S S KDL+R ML Y+P  R +AA  L H WI+     S   I+S  LS+++   + NKL
Sbjct: 326 ISESAKDLVRNMLKYDPKKRFTAAQVLEHPWIREGGEASDKPIDSAVLSRMKQLRAMNKL 385

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K   F+FI  Q +  +E   L  +F  +D +  GT++ DEL  G++    +  +  +K++
Sbjct: 386 KKLAFKFI-AQNLKEEELKGLKTMFANMDTDKSGTITYDELKSGLEKLGSRLTETEVKQL 444



>gi|7428012|pir||T02139  calcium-dependent protein kinase (EC 2.7.1.-) F8K4.14 - Arabidopsis
thaliana
 gi|3367525|gb|AAC28510.1|  (AC004392) Similar to gb|AF072908 calcium-dependent protein
kinase from Nicotiana tabacum. [Arabidopsis thaliana]
          Length = 553

 Score = 73.6 bits (179), Expect = 4e-13
 Identities = 41/120 (34%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S S KDL+R ML Y+P  R +AA  L H WI+     S   I+S  LS+++   + NKL
Sbjct: 328 ISESAKDLVRNMLKYDPKKRFTAAQVLEHPWIREGGEASDKPIDSAVLSRMKQLRAMNKL 387

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K   F+FI  Q +  +E   L  +F  +D +  GT++ DEL  G++    +  +  +K++
Sbjct: 388 KKLAFKFI-AQNLKEEELKGLKTMFANMDTDKSGTITYDELKSGLEKLGSRLTETEVKQL 446
```

```
>gi|15222931|ref|NP_177731.1|  (NM_106253) calcium-dependent protein kinase, putative
[Arabidopsis thaliana]
          Length = 534

 Score = 73.2 bits (178), Expect = 5e-13
 Identities = 42/119 (35%), Positives = 70/119 (58%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNK-SKGQINSLTLSKLQDFDSKNKLK 181
           +S S KDLIR+ML+ +P  RI+AA+AL H W+ + K S   INS  L +++ F + NKLK
Sbjct: 312 ISESAKDLIRKMLIRDPKKRITAAEALEHPWMTDTKISDKPINSAVLVRMKQFRAMNKLK 371

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L + F+ +D +  GT++ DEL  G+     +  +  IK++
Sbjct: 372 KLALKVI-AENLSEEEIKGLKQTFKNMDTDESGTITFDELRNGLHRLGSKLTESEIKQL 429
```

```
>gi|6721111|gb|AAF26765.1|AC007396_14  (AC007396) T4O12.25 [Arabidopsis thaliana]
          Length = 980

 Score = 73.2 bits (178), Expect = 5e-13
 Identities = 42/119 (35%), Positives = 70/119 (58%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNK-SKGQINSLTLSKLQDFDSKNKLK 181
           +S S KDLIR+ML+ +P  RI+AA+AL H W+ + K S   INS  L +++ F + NKLK
Sbjct: 339 ISESAKDLIRKMLIRDPKKRITAAEALEHPWMTDTKISDKPINSAVLVRMKQFRAMNKLK 398

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L + F+ +D +  GT++ DEL  G+     +  +  IK++
Sbjct: 399 KLALKVI-AENLSEEEIKGLKQTFKNMDTDESGTITFDELRNGLHRLGSKLTESEIKQL 456
```

```
>gi|20260246|gb|AAM13021.1|  (AY093022) calcium-dependent protein kinase, putative
[Arabidopsis thaliana]
          Length = 323

 Score = 73.2 bits (178), Expect = 5e-13
 Identities = 42/119 (35%), Positives = 70/119 (58%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNK-SKGQINSLTLSKLQDFDSKNKLK 181
           +S S KDLIR+ML+ +P  RI+AA+AL H W+ + K S   INS  L +++ F + NKLK
Sbjct: 101 ISESAKDLIRKMLIRDPKKRITAAEALEHPWMTDTKISDKPINSAVLVRMKQFRAMNKLK 160

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L + F+ +D +  GT++ DEL  G+     +  +  IK++
Sbjct: 161 KLALKVI-AENLSEEEIKGLKQTFKNMDTDESGTITFDELRNGLHRLGSKLTESEIKQL 218
```

```
>gi|13774101|gb|AAK38161.1|  (AY030280) calcium-dependent protein kinase [Psophocarpus
tetragonolobus]
          Length = 347

 Score = 72.8 bits (177), Expect = 6e-13
```

```
 Identities = 42/120 (35%), Positives = 70/120 (58%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWI--KNNKSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDLIRRML  NP  R++A + L H WI   N      ++S  LS+L+ F + NKL
Sbjct: 145 ISDSAKDLIRRMLDQNPKTRLTAHEVLRHPWIVDDNLAPDKPLDSAVLSRLKQFSAMNKL 204

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L ++F+ ID +  GT++ DEL  G++    + ++  IK++
Sbjct: 205 KKMALRVI-AERLSEEEIGGLKELFKMIDTDNSGTITFDELKDGLKRVGSELMESEIKDL 263




 Score = 34.7 bits (78), Expect = 0.19
 Identities = 21/72 (29%), Positives = 34/72 (47%), Gaps = 2/72 (2%)
 Frame = +2

Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN--NKQ 328
           D D    + Y  F   TV +   + + +L+  F   DK+G G ++ DE+ Q  ++
Sbjct: 268 DIDKSGTIDYGEFIAATVHLNKLEREENLVSAFSYFDKDGSGYITLDEIQQACKDFGLDD 327

Query: 329 QPIQKFIKEINQ 364
             I   IKEI+Q
Sbjct: 328 VHIDDMIKEIDQ 339



>gi|116054|sp|P28583|CDPK_SOYBN  Calcium-dependent protein kinase SK5 (CDPK)
 gi|280393|pir||A43713  calcium-dependent protein kinase (EC 2.7.1.-) - soybean
 gi|169931|gb|AAB00806.1|  (M64987) Glycine max calcium dependent protein kinase mRNA
          Length = 508

 Score = 71.6 bits (174), Expect = 1e-12
 Identities = 41/120 (34%), Positives = 70/120 (58%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWI--KNNKSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDLIR+ML  NP  R++A + L H WI   N      ++S  LS+L+ F + NKL
Sbjct: 261 ISDSAKDLIRKMLDQNPKTRLTAHEVLRHPWIVDDNIAPDKPLDSAVLSRLKQFSAMNKL 320

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L ++F+ ID +  GT++ DEL  G++    + ++  IK++
Sbjct: 321 KKMALRVI-AERLSEEEIGGLKELFKMIDTDNSGTITFDELKDGLKRVGSELMESEIKDL 379




 Score = 34.7 bits (78), Expect = 0.19
 Identities = 21/72 (29%), Positives = 34/72 (47%), Gaps = 2/72 (2%)
 Frame = +2


Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQP 334
           D D    + Y  F   TV +   + + +L+  F   DK+G G ++ DE+ Q  ++
Sbjct: 384 DIDKSGTIDYGEFIAATVHLNKLEREENLVSAFSYFDKDGSGYITLDEIQQACKDFGLDD 443
```

```
Query: 335 --IQKFIKEINQ 364
             I   IKEI+Q
Sbjct: 444 IHIDDMIKEIDQ 455




>gi|2271459|gb|AAC13354.1|  (AF009560) calcium-dependent protein kinase-a [Paramecium
tetraurelia]
 gi|2271463|gb|AAC13356.1|  (AF009562) calcium-dependent protein kinase-a [Paramecium
tetraurelia]
          Length = 481

 Score = 70.5 bits (171), Expect = 3e-12
 Identities = 36/104 (34%), Positives = 61/104 (58%), Gaps = 3/104 (2%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK---NNKSKGQINSLTLSKLQDFDSKNK 175
           VS   KDL+ +ML Y+   R+SA   L H +++     K   QI    L  L +F ++ K
Sbjct: 237 VSAHAKDLVSQMLQYDVQKRLSAKQVLEHPFLQLQHQEKVDKQIVQCRLKNLVNFRAEQK 296

Query: 176 LKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQ 307
           L+ A   FI   +++ +EK  L++ F+E+D+NGDG ++K+E+ +
Sbjct: 297 LQQATLMFIGTTMISKEEKNQLMQAFKEMDQNGDGILTKEEILE 340




 Score = 32.0 bits (71), Expect = 1.2
 Identities = 23/76 (30%), Positives = 34/76 (44%), Gaps = 4/76 (5%)
 Frame = +2

Query: 149 LQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVS----KDELYQGIQ 316
           L D D    + Y  F   T+      +K  L + FQ  DK+G+G +S    KD L   I
Sbjct: 361 LVDMDGSGTIDYTEFIIATMDRKKAVQKEKLKEAFQIFDKDGNGFISEQEIKDVLGPSIT 420

Query: 317 NNKQQPIQKFIKEINQ 364
              ++     IKEI++
Sbjct: 421 GIDEKYWMNMIKEIDK 436




>gi|7434371|pir||T08874  calcium-dependent protein kinase (EC 2.7.1.-) gamma - soybean
 gi|2501766|gb|AAB80693.1|  (U69174) calmodulin-like domain protein kinase isoenzyme gamma
[Glycine max]
          Length = 538

 Score = 70.5 bits (171), Expect = 3e-12
 Identities = 39/119 (32%), Positives = 70/119 (58%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKN-NKSKGQINSLTLSKLQDFDSKNKLK 181
           +S+S KDL+R+ML+ +P  RI++A  L H WIK+ N S   I+S  LS+++ F + NKLK
Sbjct: 311 ISDSAKDLVRKMLIQDPKKRITSAQVLEHPWIKDGNASDKPIDSAVLSRMKQFRAMNKLK 370

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L  +F  +D +  GT++ +EL  G+     +  +  +K++
Sbjct: 371 KLALKVI-AENMSAEEIQGLKAMFTNMDTDKSGTITYEELKSGLHRLGSKLTEAEVKQL 428
```

```
 Score = 31.2 bits (69), Expect = 2.1
 Identities = 19/77 (24%), Positives = 33/77 (42%), Gaps = 3/77 (3%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+     +    L K FQ  DK+  G +++DEL   ++
Sbjct: 428 LMEAADVDGNGSIDYIEFITATMHRHKLERDDQLFKAFQYFDKDNSGFITRDELESAMKE 487

Query: 320 ---NKQQPIQKFIKEIN 361
                   I++ I E++
Sbjct: 488 YGMGDDATIKEIISEVD 504



>gi|19171502|emb|CAC87494.1|  (AJ308296) calcium-dependent protein kinase [Lycopersicon
esculentum]
          Length = 553

 Score = 69.7 bits (169), Expect = 5e-12
 Identities = 39/119 (32%), Positives = 66/119 (54%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK-NNKSKGQINSLTLSKLQDFDSKNKLK 181
           +SNS KDLIR+ML   P  RI++A  L H W++    S   I+S  LS+++ F + NKLK
Sbjct: 332 ISNSAKDLIRKMLTQEPRKRITSAQVLEHPWLRLGEASDKPIDSAVLSRMKQFRAMNKLK 391

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L  +F  ID +  GT++ +EL  G+     +  +  +K++
Sbjct: 392 KLALKVI-AENLSEEEIKGLKAMFHNIDTDNSGTITYEELKSGLARLGSKLTETEVKQL 449




 Score = 31.6 bits (70), Expect = 1.6
 Identities = 21/79 (26%), Positives = 39/79 (48%), Gaps = 5/79 (6%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTD--LLKIFQEIDKNGDGTVSKDELYQGI 313
           L +  D D    + Y   +FIT  +  ++ + D  L K FQ  DK+  G +++DEL   +
Sbjct: 449 LMEAADVDGNGSIDY--IEFITATMHRHRLERDEHLFKAFQHFDKDHSGFITRDELENAM 506

Query: 314 QN---NKQQPIQKFIKEIN 361
           +      +  I++ I E++
Sbjct: 507 KEYGMGDEATIKEIIAEVD 525



>gi|7428008|pir||T03024  calcium-dependent protein kinase (EC 2.7.1.-), calmodulin-
independent - maize (fragment)
 gi|639722|gb|AAA61682.1|  (L27484) calcium-dependent protein kinase [Zea mays]
          Length = 465

 Score = 69.7 bits (169), Expect = 5e-12
 Identities = 36/105 (34%), Positives = 62/105 (58%), Gaps = 2/105 (1%)
```

```
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ--INSLTLSKLQDFDSKNKL 178
           +S   KDL+++ML  NP  R++A   LNH WIK +       ++++ L +L+ F + N+
Sbjct: 235 ISPGAKDLVKKMLNINPKERLTAFQVLNHPWIKEDGDAPDTPLDNVVLDRLKQFRAMNQF 294

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K A  + I    ++ +E T L ++F+ IDK+  GT++ DEL  G+
Sbjct: 295 KKAALRII-AGCLSEEEITGLKEMFKNIDKDNSGTITLDELKHGL 338




 Score = 32.3 bits (72), Expect = 0.94
 Identities = 24/94 (25%), Positives = 36/94 (37%)
 Frame = +2

Query: 77  DALNHEWIKNNKSKGQINSLTLSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQ 256
           D L H   K+           L +  D D    + Y  F   TV +     +  L   FQ
Sbjct: 332 DELKHGLAKHGPKLSDSEMEKLMEAADADGNGLIDYDEFVTATVHMNKLDREEHLYTAFQ 391

Query: 257 EIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
             DK+  G ++K+EL   ++        K IK+I
Sbjct: 392 YFDKDNSGYITKEELEHALKEQGLYDADK-IKDI 424



>gi|15234656|ref|NP_193925.1|  (NM_118315) calcium-dependent protein kinase - like protein
[Arabidopsis thaliana]
 gi|7434360|pir||T05476  calcium-dependent protein kinase (EC 2.7.1.-) T8O5.150 -
Arabidopsis thaliana
 gi|2894572|emb|CAA17161.1|  (AL021890) calcium-dependent protein kinase - like protein
[Arabidopsis thaliana]
 gi|7269039|emb|CAB79149.1|  (AL161556) calcium-dependent protein kinase-like protein
[Arabidopsis thaliana]
 gi|14334644|gb|AAK59500.1|  (AY034995) putative calcium-dependent protein kinase
[Arabidopsis thaliana]
 gi|17104579|gb|AAL34178.1|  (AY063004) putative calcium-dependent protein kinase
[Arabidopsis thaliana]
          Length = 554

 Score = 69.3 bits (168), Expect = 7e-12
 Identities = 38/119 (31%), Positives = 69/119 (57%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +S S KDL+R++L  +P  RISAA AL H WI+  ++  + I+S  LS+++ F + NKLK
Sbjct: 329 ISESAKDLVRKLLTKDPKQRISAAQALEHPWIRGGEAPDKPIDSAVLSRMKQFRAMNKLK 388

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L  +F  +D +  GT++ +EL  G+     +  +  +K++
Sbjct: 389 KLALKVI-AESLSEEEIKGLKTMFANMDTDKSGTITYEELKNGLAKLGSKLTEAEVKQL 446
```

```
>gi|15236560|ref|NP_194096.1|  (NM_118496) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|2129550|pir||S71197  calcium-dependent protein kinase (EC 2.7.1.-) CDPK6 - Arabidopsis
thaliana
 gi|2129554|pir||S71901  calcium-dependent protein kinase 6 - Arabidopsis thaliana
 gi|7428013|pir||S71774  calcium-dependent protein kinase (EC 2.7.1.-) 6 - Arabidopsis
thaliana
 gi|836940|gb|AAA67654.1|  (U20623) calcium-dependent protein kinase [Arabidopsis thaliana]
 gi|836944|gb|AAA67656.1|  (U20625) calcium-dependent protein kinase [Arabidopsis thaliana]
 gi|4454034|emb|CAA23031.1|  (AL035394) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|7269213|emb|CAB79320.1|  (AL161559) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|14326514|gb|AAK60302.1|AF385710_1  (AF385710) AT4g23650/F9D16_120 [Arabidopsis thaliana]
 gi|19548043|gb|AAL87385.1|  (AY081732) AT4g23650/F9D16_120 [Arabidopsis thaliana]
          Length = 529

 Score = 69.3 bits (168), Expect = 7e-12
 Identities = 35/105 (33%), Positives = 67/105 (63%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+  KDL+R+ML Y+P +R++AA+ LNH WI+ +   S   +++  LS+++ F + NKL
Sbjct: 305 LSDGAKDLVRKMLKYDPKDRLTAAEVLNHPWIREDGEASDKPLDNAVLSRMKQFRAMNKL 364

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L ++F+ +D + +G V+ +EL  G+
Sbjct: 365 KKMALKVI-AENLSEEEIIGLKEMFKSLDTDNNGIVTLEELRTGL 408



>gi|15234435|ref|NP_192381.1|  (NM_116710) putative calcium dependent protein kinase
[Arabidopsis thaliana]
 gi|4115943|gb|AAD03453.1|  (AF118223) contains similarity to eukaryotic protein kinase
domains (Pfam: PF00069, score=312.6, E=4.7e-90, N=1) and EF hand domains (Pfam: PF00036,
score=131, E=2.1e-35, N=4) [Arabidopsis thaliana]
 gi|7267230|emb|CAB80837.1|  (AL161501) putative calcium dependent protein kinase
[Arabidopsis thaliana]
          Length = 531

 Score = 69.3 bits (168), Expect = 7e-12
 Identities = 38/104 (36%), Positives = 62/104 (59%), Gaps = 1/104 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +S S KDL+R+ML  +P  RI+AA  L H WIK  ++  + I+S  LS+++ F + NKLK
Sbjct: 307 ISESAKDLVRKMLTKDPKRRITAAQVLEHPWIKGGEAPDKPIDSAVLSRMKQFRAMNKLK 366

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
                + I   + ++ +E   L  +F  ID +  GT++ +EL  G+
Sbjct: 367 KLALKVI-AESLSEEEIKGLKTMFANIDTDKSGTITYEELKTGL 409




 Score = 29.6 bits (65), Expect = 6.1
 Identities = 18/77 (23%), Positives = 33/77 (42%), Gaps = 3/77 (3%)
 Frame = +2
```

```
Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+          + K FQ  DK+  G +++DEL   ++
Sbjct: 424 LMEAADVDGNGTIDYYEFISATMHRYKLDRDEHVYKAFQHFDKDNSGHITRDELESAMKE 483

Query: 320 ---NKQQPIQKFIKEIN 361
                +  I++ I E++
Sbjct: 484 YGMGDEASIKEVISEVD 500



>gi|15292915|gb|AAK92828.1|  (AY050891) putative calcium dependent protein kinase
[Arabidopsis thaliana]
          Length = 531

 Score = 69.3 bits (168), Expect = 7e-12
 Identities = 38/104 (36%), Positives = 62/104 (59%), Gaps = 1/104 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +S S KDL+R+ML  +P  RI+AA  L H WIK  ++  + I+S  LS+++ F + NKLK
Sbjct: 307 ISESAKDLVRKMLTKDPKRRITAAQVLEHPWIKGGEAPDKPIDSAVLSRMKQFRAMNKLK 366

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
               + I  + ++ +E   L  +F  ID +  GT++ +EL  G+
Sbjct: 367 KLALKVI-AESLSEEEIKGLKTMFANIDTDKSGTITYEELKTGL 409




 Score = 29.6 bits (65), Expect = 6.1
 Identities = 18/77 (23%), Positives = 33/77 (42%), Gaps = 3/77 (3%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+          + K FQ  DK+  G +++DEL   ++
Sbjct: 424 LMEAADVDGNGTIDYYEFISATMHRYKLDRDEHVYKAFQHFDKDSSGHITRDELESAMKE 483

Query: 320 ---NKQQPIQKFIKEIN 361
                +  I++ I E++
Sbjct: 484 YGMGDEASIKEVISEVD 500



>gi|4115945|gb|AAD03455.1|  (AF118223) contains similarity to eukaryotic protein kinase
domains (Pfam: PF00069, score=253.1, E=3.8e-72, N=1) and EF hand domains (Pfam: PF00036,
score=94.6, E=2e-24 , N=4) [Arabidopsis thaliana]
          Length = 584

 Score = 68.9 bits (167), Expect = 9e-12
 Identities = 38/124 (30%), Positives = 70/124 (55%), Gaps = 5/124 (4%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKS-KGQINSLTLSKLQDFDSKNKLK 181
           +S+S KDL+ +ML  +P  RI+AA  L H WIK  ++ +  I+S  LS+++ F + NKLK
Sbjct: 308 ISDSAKDLVEKMLTEDPKRRITAAQVLEHPWIKGGEAPEKPIDSTVLSRMKQFRAMNKLK 367

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNK----QQPIQKFI 349
               +   V  ++ +E   L  +F  +D N  GT++ ++L  G+   +    +  +Q+ +
Sbjct: 368 KLALKVSAVS-LSEEEIKGLKTLFANMDTNRSGTITYEQLQTGLSRLRSRLSETEVQQLV 426
```

```
Query: 350 KEIN 361
           + +N
Sbjct: 427 EAVN 430



 Score = 31.6 bits (70), Expect = 1.6
 Identities = 19/88 (21%), Positives = 39/88 (43%), Gaps = 3/88 (3%)
 Frame = +2

Query: 107 NKSKGQINSLTLSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTV 286
           N  K  +  + ++   D D    + Y  F   T+          + K FQ +DK+ +G +
Sbjct: 430 NLEKDSLKHMMITFQSDVDGNGTIDYYEFISATMHRYKLHHDEHVHKAFQHLDKDKNGHI 489

Query: 287 SKDELYQGIQN---NKQQPIQKFIKEIN 361
           ++DEL   ++      +  I++ I E++
Sbjct: 490 TRDELESAMKEYGMGDEASIKEVISEVD 517



>gi|7428009|pir||T01989  calcium-dependent protein kinase (EC 2.7.1.-) 1 - common tobacco
 gi|3283996|gb|AAC25423.1|  (AF072908) calcium-dependent protein kinase [Nicotiana tabacum]
          Length = 540

 Score = 68.6 bits (166), Expect = 1e-11
 Identities = 37/119 (31%), Positives = 68/119 (57%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +SNS KDLIR+ML   P  R+++A  L H W++  ++  + I+S  LS+++ F + NKLK
Sbjct: 320 ISNSAKDLIRKMLTQEPRKRMTSAQVLEHPWLRMGEASDKPIDSAVLSRMKQFRAMNKLK 379

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               + I  + ++ +E   L  +F  ID +  GT++ +EL  G+     +  +  +K++
Sbjct: 380 KLALKVI-AENLSEEEIKGLKAMFANIDTDNSGTITYEELKSGLARLGSKLTETEVKQL 437




 Score = 30.4 bits (67), Expect = 3.6
 Identities = 21/79 (26%), Positives = 39/79 (48%), Gaps = 5/79 (6%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTD--LLKIFQEIDKNGDGTVSKDELYQGI 313
           L +  D D    + Y   +FIT  +  ++ + D  L K FQ  DK+  G +++DEL   +
Sbjct: 437 LMEAADVDGNGTIDY--IEFITATMHRHRLERDEHLFKAFQYFDKDHSGFITRDELESAM 494

Query: 314 QN---NKQQPIQKFIKEIN 361
           +      +  I++ I E++
Sbjct: 495 KEYGMGDEATIKEIIAEVD 513
```

```
>gi|15450437|gb|AAK96512.1|  (AY052319) AT4g23650/F9D16_120 [Arabidopsis thaliana]
 gi|17386100|gb|AAL38596.1|AF446863_1  (AF446863) AT4g23650/F9D16_120 [Arabidopsis thaliana]
          Length = 529

 Score = 68.6 bits (166), Expect = 1e-11
 Identities = 35/105 (33%), Positives = 66/105 (62%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+  KDL+R ML Y+P +R++AA+ LNH WI+ +   S   +++  LS+++ F + NKL
Sbjct: 305 LSDGAKDLVREMLKYDPKDRLTAAEVLNHPWIREDGEASDKPLDNAVLSRMKQFRAMNKL 364

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L ++F+ +D + +G V+ +EL  G+
Sbjct: 365 KKMALKVI-AENLSEEEIIGLKEMFKSLDTDNNGIVTLEELRTGL 408



>gi|6469601|gb|AAF13351.1|AF121337_1  (AF121337) calcium/calmodulin-dependent protein kinase
[Eufolliculina uhligi]
          Length = 368

 Score = 68.6 bits (166), Expect = 1e-11
 Identities = 36/107 (33%), Positives = 62/107 (57%), Gaps = 3/107 (2%)
 Frame = +2

Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQINSL---TLSKLQDFDSKN 172
           KVS   ++L+++ML  +P   +SA  A  H WI  N S+   + L   +L  L+ F
Sbjct: 151 KVSKDAQNLLKKMLERDPVKILSAVQAAKHPWIVRNMSEESTSPLLIESLMNLKHFKVTQ 210

Query: 173 KLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K+  AIF +IT Q+++ +E+  L ++F  +D + +G +S+ EL QG+
Sbjct: 211 KVDKAIFYYITYQILSPKEQNSLTELFNSLDMDRNGKLSRAELEQGL 257



>gi|2827773|sp|P28582|CDPK_DAUCA  Calcium-dependent protein kinase (CDPK)
 gi|7434365|pir||T14335  protein kinase, calcium-dependent (EC 2.7.1.-) - carrot
 gi|1765912|emb|CAA39936.1|  (X56599) calcium- dependent protein kinase [Daucus carota]
          Length = 532

 Score = 68.2 bits (165), Expect = 2e-11
 Identities = 37/105 (35%), Positives = 62/105 (58%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           VSNS KDL+R+ML  +P  RI++A  L+H W++     S   I+S  LS+++ F + NKL
Sbjct: 308 VSNSAKDLVRKMLTQDPRRRITSAQVLDHPWMREGGEASDKPIDSAVLSRMKQFRAMNKL 367

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L  +F  +D +  GT++ +EL  G+
Sbjct: 368 KQLALKVI-AESLSEEEIKGLKSMFANMDTDKSGTITYEELKSGL 411
```

```
>gi|82056|pir||S17759  protein kinase, calcium-dependent (EC 2.7.1.-) - carrot  (fragment)
          Length = 425

 Score = 68.2 bits (165), Expect = 2e-11
 Identities = 37/105 (35%), Positives = 62/105 (58%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           VSNS KDL+R+ML  +P  RI++A  L+H W++      S    I+S  LS+++ F + NKL
Sbjct: 201 VSNSAKDLVRKMLTQDPRRRITSAQVLDHPWMREGGEASDKPIDSAVLSRMKQFRAMNKL 260

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L  +F  +D +  GT++ +EL  G+
Sbjct: 261 KQLALKVI-AESLSEEEIKGLKSMFANMDTDKSGTITYEELKSGL 304
```

```
>gi|12313676|dbj|BAB21081.1|  (AP002819) putative calcium-dependent protein kinase [Oryza
sativa (japonica cultivar-group)]
          Length = 518

 Score = 67.4 bits (163), Expect = 3e-11
 Identities = 31/105 (29%), Positives = 65/105 (61%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ--INSLTLSKLQDFDSKNKL 178
           +SN  KDL+++ML  +P  R+++A+ LNH WI+ +       ++    +S+++ F + NKL
Sbjct: 293 ISNGAKDLVKKMLRQDPKERLTSAEILNHPWIREDGEAPDKPLDITVISRMKQFRAMNKL 352

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + +  + ++++E T L ++F+ +D +  GT++ +EL  G+
Sbjct: 353 KKVALK-VVAENLSDEEITGLKEMFRSLDTDNSGTITLEELRSGL 396



 Score = 33.9 bits (76), Expect = 0.32
 Identities = 21/73 (28%), Positives = 36/73 (48%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + YA F   T+ +   +++  +LK F+  DK+  G ++ DEL + ++
Sbjct: 411 LMEAADVDGNGTIDYAEFISATMHMNRLEKEDHILKAFEYFDKDHSGYITVDELEEALKK 470

Query: 320 NKQQPIQKFIKEI 358
                  K IKEI
Sbjct: 471 YDMGD-DKTIKEI 482
```

```
>gi|4336426|gb|AAD17800.1|  (AF090835) Ca2+-dependent protein kinase [Mesembryanthemum
crystallinum]
          Length = 534

 Score = 67.4 bits (163), Expect = 3e-11
 Identities = 34/104 (32%), Positives = 64/104 (60%), Gaps = 1/104 (0%)
 Frame = +2
```

```
Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +SN  KDL+R+ML  +P  RI+AA  L H W+++ ++  + I+S  L +++ F   NKLK
Sbjct: 313 ISNGAKDLVRKMLTQDPKKRITAAQVLEHPWLRDGEASDKPIDSAVLLRMKQFRVMNKLK 372

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
               + I  + + ++E   L ++F  +D +G G+++ +EL +G+
Sbjct: 373 KLALKVI-AENLPDEEIQGLKQMFANMDTDGSGSITYEELREGL 415



>gi|15234441|ref|NP_192383.1|  (NM_116712) putative calcium dependent protein kinase
[Arabidopsis thaliana]
 gi|7267232|emb|CAB80839.1|  (AL161501) putative calcium dependent protein kinase
[Arabidopsis thaliana]
 gi|16648907|gb|AAL24305.1|  (AY059823) Unknown protein [Arabidopsis thaliana]
 gi|20148457|gb|AAM10119.1|  (AY081557) unknown protein [Arabidopsis thaliana]
          Length = 520

 Score = 67.4 bits (163), Expect = 3e-11
 Identities = 36/119 (30%), Positives = 68/119 (56%), Gaps = 1/119 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKS-KGQINSLTLSKLQDFDSKNKLK 181
           +S+S KDL+ +ML  +P  RI+AA  L H WIK  ++ +  I+S  LS+++ F + NKLK
Sbjct: 296 ISDSAKDLVEKMLTEDPKRRITAAQVLEHPWIKGGEAPEKPIDSTVLSRMKQFRAMNKLK 355

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
               +   V  ++ +E   L  +F  +D N  GT++ ++L  G+   + +  +  ++++
Sbjct: 356 KLALKVSAVS-LSEEEIKGLKTLFANMDTNRSGTITYEQLQTGLSRLRSRLSETEVQQL 413




 Score = 29.3 bits (64), Expect = 7.9
 Identities = 18/77 (23%), Positives = 35/77 (45%), Gaps = 3/77 (3%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+          + K FQ +DK+ +G +++DEL   ++
Sbjct: 413 LVEASDVDGNGTIDYYEFISATMHRYKLHHDEHVHKAFQHLDKDKNGHITRDELESAMKE 472

Query: 320 ---NKQQPIQKFIKEIN 361
                +  I++ I E++
Sbjct: 473 YGMGDEASIKEVISEVD 489



>gi|13872969|dbj|BAB44074.1|  (AP003073) putative calcium-dependent protein kinase [Oryza
sativa (japonica cultivar-group)]
 gi|20160680|dbj|BAB89623.1|  (AP003260) putative calcium-dependent protein kinase [Oryza
sativa (japonica cultivar-group)]
          Length = 501

 Score = 67.0 bits (162), Expect = 3e-11
 Identities = 35/106 (33%), Positives = 64/106 (60%), Gaps = 2/106 (1%)
 Frame = +2
```

```
Query: 2   KVSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ--INSLTLSKLQDFDSKNK 175
           ++S S KDL+R+ML  +P  RISA + LNH WIK +       +++  +++L+ F + N+
Sbjct: 273 RISASAKDLVRKMLNSDPKKRISAYEVLNHPWIKEDGEAPDTPLDNAVMNRLKQFRAMNQ 332

Query: 176 LKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
            K A  + I    ++ +E   L ++F+ +D +  GT++ DEL +G+
Sbjct: 333 FKKAALRVI-AGCLSEEEIRGLKEMFKSMDSDNSGTITVDELRKGL 377




 Score = 31.6 bits (70), Expect = 1.6
 Identities = 17/59 (28%), Positives = 27/59 (44%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ 316
           L +  D D    + Y  F   T+ +     +  L   FQ  DK+  G +SK+EL Q ++
Sbjct: 392 LMEAADADGNGTIDYDEFITATMHMNRMDREEHLYTAFQYFDKDNSGCISKEELEQALR 450



>gi|7428003|pir||S71770  calcium-dependent protein kinase (EC 2.7.1.-) - mung bean
 gi|967125|gb|AAC49405.1|  (U08140) calcium dependent protein kinase [Vigna radiata]
          Length = 487

 Score = 66.6 bits (161), Expect = 4e-11
 Identities = 38/106 (35%), Positives = 61/106 (56%), Gaps = 2/106 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+SGKDLIR+ML   P+ R++A   L H WI  N       I+   LS+L+ F + NKL
Sbjct: 251 ISDSGKDLIRKMLCSQPSERLTAHQVLCHPWICENGVAPDRAIDPAVLSRLKQFSAMNKL 310

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ 316
           K    + I  + ++ +E   L ++FQ +D +  G ++ DEL  G++
Sbjct: 311 KKMALRVI-AESLSEEEIAGLREMFQAMDTDNSGAITFDELKAGLR 355




 Score = 37.0 bits (84), Expect = 0.038
 Identities = 22/77 (28%), Positives = 36/77 (46%), Gaps = 2/77 (2%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI-- 313
           L +  D D    + Y  F   TV +   + +  L+  FQ  DK+G G ++ DEL Q
Sbjct: 369 LMEAADVDKSGTIDYGEFIAATVHLNKLEREEHLIAAFQYFDKDGSGYITVDELQQACAE 428

Query: 314 QNNKQQPIQKFIKEINQ 364
            N     ++  I+E++Q
Sbjct: 429 HNMTDAFLEDIIREVDQ 445



>gi|1705734|sp|P53683|CDP2_ORYSA  Calcium-dependent protein kinase, isoform 2 (CDPK 2)
 gi|1362176|pir||S56652  calcium-dependent protein kinase (EC 2.7.1.-) 2 - rice
```

```
 gi|587498|emb|CAA57157.1|  (X81394) calcium-dependent protein kinase [Oryza sativa
(japonica cultivar-group)]
          Length = 533

 Score = 66.6 bits (161), Expect = 4e-11
 Identities = 34/104 (32%), Positives = 62/104 (58%), Gaps = 1/104 (0%)
 Frame = +2


Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLK 181
           +S S KDL+R+ML  +P  RI++A  L H W+++ ++  + I+S  LS+++ F + NKLK
Sbjct: 312 ISESAKDLVRKMLTQDPKKRITSAQVLQHPWLRDGEASDKPIDSAVLSRMKQFRAMNKLK 371

Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
               + I  +    +E   L ++F  +D +  GT++ +EL  G+
Sbjct: 372 KMALKVIASN-LNEEEIKGLKQMFTNMDTDNSGTITYEELKAGL 414




 Score = 30.8 bits (68), Expect = 2.7
 Identities = 19/77 (24%), Positives = 33/77 (42%), Gaps = 3/77 (3%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+     +    L K FQ  DK+  G +++DEL   +
Sbjct: 429 LMEAADVDGNGSIDYVEFITATMHRHKLERDEHLFKAFQYFDKDNSGFITRDELESALIE 488

Query: 320 NKQ---QPIQKFIKEIN 361
           ++      I+  I E++
Sbjct: 489 HEMGDTSTIKDIISEVD 505



>gi|15289760|dbj|BAB63464.1|  (AB051809) calcium dependent protein kinase [Solanum
tuberosum]
          Length = 496

 Score = 66.2 bits (160), Expect = 6e-11
 Identities = 37/120 (30%), Positives = 72/120 (59%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDLIR++L  NP  R++A + L H WI ++       ++S  LS+L+ F + NKL
Sbjct: 255 ISDSAKDLIRKILDRNPKRRLTAHEVLCHPWIVDDTVAPDKPLDSAVLSRLKQFSAMNKL 314

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L ++F+ +D +  GT++ +EL +G++    + ++  IK++
Sbjct: 315 KKMALRVI-AERLSEEEIGGLKELFKMLDTDNSGTITFEELKEGLRRVGSELMESEIKDL 373




 Score = 34.7 bits (78), Expect = 0.19
 Identities = 20/72 (27%), Positives = 36/72 (49%), Gaps = 2/72 (2%)
 Frame = +2
```

```
Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ--NNKQ 328
           D D+   + Y  F   TV +   + + +LL  F   DK+G G ++ +EL Q  +     +
Sbjct: 378 DIDNNGTIDYGEFIAATVHLNKLEREENLLSAFSYFDKDGSGYITIEELQQACKEFGLSE 437

Query: 329 QPIQKFIKEINQ 364
             + + IK+I+Q
Sbjct: 438 LNLDEIIKDIDQ 449




>gi|15236966|ref|NP_195257.1|  (NM_119697) calmodulin-domain protein kinase CDPK isoform 5
(CPK5) [Arabidopsis thaliana]
 gi|7434359|pir||T06126  calcium-dependent protein kinase (EC 2.7.1.-) CPK5 - Arabidopsis
thaliana
 gi|1399273|gb|AAB03245.1|  (U31834) calmodulin-domain protein kinase CDPK isoform 5
[Arabidopsis thaliana]
 gi|3080419|emb|CAA18738.1|  (AL022604) calmodulin-domain protein kinase CDPK isoform 5
(CPK5) [Arabidopsis thaliana]
 gi|7270483|emb|CAB80248.1|  (AL161587) calmodulin-domain protein kinase CDPK isoform 5
(CPK5) [Arabidopsis thaliana]
          Length = 556

 Score = 66.2 bits (160), Expect = 6e-11
 Identities = 37/106 (34%), Positives = 60/106 (55%), Gaps = 2/106 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDLIRRML   P  R++A + L H WI  N       ++   LS+L+ F + NKL
Sbjct: 324 ISDSAKDLIRRMLSSKPAERLTAHEVLRHPWICENGVAPDRALDPAVLSRLKQFSAMNKL 383

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ 316
           K    + I  + ++ +E   L ++FQ +D +  G ++ DEL  G++
Sbjct: 384 KKMALKVI-AESLSEEEIAGLREMFQAMDTDNSGAITFDELKAGLR 428




 Score = 36.2 bits (82), Expect = 0.065
 Identities = 20/72 (27%), Positives = 36/72 (49%), Gaps = 2/72 (2%)
 Frame = +2

Query: 155 DFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQP 334
           D D+   + Y+ F   T+ +   + +  L+  FQ  DK+G G ++ DEL Q    +
Sbjct: 447 DVDNSGTIDYSEFIAATIHLNKLEREEHLVAAFQYFDKDGSGFITIDELQQACVEHGMAD 506

Query: 335 I--QKFIKEINQ 364
           +  +  IKE++Q
Sbjct: 507 VFLEDIIKEVDQ 518




>gi|5162877|dbj|BAA81748.1|  (AB017515) calcium-dependent protein kinase [Marchantia
polymorpha]
 gi|5162880|dbj|BAA81750.1|  (AB017516) calcium-dependent protein kinase [Marchantia
polymorpha]
          Length = 548

 Score = 66.2 bits (160), Expect = 6e-11
```

```
 Identities = 34/120 (28%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S   KDL++RML  NP  R++A + L+H WI          +++  LS+L+ F + NKL
Sbjct: 308 ISQKAKDLVKRMLKQNPKERLTAHEVLSHPWISVDGEAPDKPLDNAVLSRLKQFTAMNKL 367

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L ++F+ +D +  GT++ +EL  G+Q       +  ++++
Sbjct: 368 KKLALKVI-AESLSEEEIMGLKEMFKSMDTDNSGTITFEELKDGLQKQGSNLAESEVRQL 426



>gi|5162878|dbj|BAA81749.1|  (AB017515) calcium-dependent protein kinase [Marchantia
polymorpha]
 gi|5162882|dbj|BAA81751.1|  (AB017517) calcium-dependent protein kinase [Marchantia
polymorpha]
          Length = 548

 Score = 66.2 bits (160), Expect = 6e-11
 Identities = 34/120 (28%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S   KDL++RML  NP  R++A + L+H WI          +++  LS+L+ F + NKL
Sbjct: 308 ISQKAKDLVKRMLKQNPKERLTAHEVLSHPWISVDGEAPDKPLDNAVLSRLKQFTAMNKL 367

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L ++F+ +D +  GT++ +EL  G+Q       +  ++++
Sbjct: 368 KKLALKVI-AESLSEEEIMGLKEMFKSMDTDNSGTITFEELKDGLQKQGSNLAESEVRQL 426



>gi|15234432|ref|NP_192380.1|  (NM_116709) putative calcium dependent protein kinase
[Arabidopsis thaliana]
 gi|4115942|gb|AAD03452.1|  (AF118223) contains similarity to eukaryotic protein kinase
domains (Pfam: PF00069, score=238.4, E= 1e-67, N=1) and EF hand domains (Pfam: PF00036,
score=109.0, E=8.9e-29, N=5) [Arabidopsis thaliana]
 gi|7267229|emb|CAB80836.1|  (AL161501) putative calcium dependent protein kinase
[Arabidopsis thaliana]
          Length = 575

 Score = 65.9 bits (159), Expect = 8e-11
 Identities = 35/118 (29%), Positives = 66/118 (55%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQINSLTLSKLQDFDSKNKLKY 184
           +S   K LI +ML   P  RISAAD L H W+K+      I+++ LS+++ F + NKLK
Sbjct: 260 ISFKAKHLIGKMLTKKPKERISAADVLEHPWMKSEAPDKPIDNVVLSRMKQFRAMNKLKK 319

Query: 185 AIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
              + I  + ++ +E   L  +F+ +D +  G+++ +EL  G+  +  +  +  +K++
Sbjct: 320 LALKVI-AEGLSEEEIKGLKTMFENMDMDKSGSITYEELKMGLNRHGSKLSETEVKQL 376
```

```
 Score = 30.8 bits (68), Expect = 2.7
 Identities = 17/63 (26%), Positives = 29/63 (45%)
 Frame = +2

Query: 113 SKGQINSLTLSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSK 292
           S+ ++  L  +   D D    + Y  F   T+     +    L K FQ  DK+G G ++K
Sbjct: 369 SETEVKQLMEAVSADVDGNGTIDYIEFISATMHRHRLERDEHLYKAFQYFDKDGSGHITK 428

Query: 293 DEL 301
           +E+
Sbjct: 429 EEV 431




 Score = 29.6 bits (65), Expect = 6.1
 Identities = 14/45 (31%), Positives = 22/45 (48%)
 Frame = +2

Query: 248 IFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEINQQLRLLQ 382
           +  E DKN DG +  +E    ++N   QP  K +K +   L  L+
Sbjct: 449 LISEFDKNNDGKIDYEEFCTMMRNGILQPQGKLLKRLYMNLEELK 493



>gi|7428004|pir||T03271  calcium-dependent protein kinase (EC 2.7.1.-) 1 - maize
 gi|1632768|dbj|BAA12338.1|  (D84408) calcium dependent protein kinase [Zea mays]
          Length = 492

 Score = 65.5 bits (158), Expect = 1e-10
 Identities = 35/106 (33%), Positives = 63/106 (59%), Gaps = 2/106 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+S KDLIRRML   P  R++A + L H WI+++       ++   LS+++ F + NKL
Sbjct: 254 ISDSAKDLIRRMLNPRPAERLTAHEVLCHPWIRDHGVAPDRPLDPAVLSRIKQFSAMNKL 313

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQ 316
           K   + I  + ++ +E   L ++FQ +D +  G ++ DEL +G++
Sbjct: 314 KKMALRVI-AESLSEEEIAGLKEMFQTMDTDNSGAITYDELKEGLR 358



>gi|14970920|emb|CAC44471.1|  (AJ320242) calcium dependent calmodulin independent protein
kinase [Cucumis sativus]
 gi|15011837|gb|AAK26164.2|  (AY027885) calcium-dependent calmodulin-independent protein
kinase 5 [Cucumis sativus]
          Length = 514

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 34/105 (32%), Positives = 64/105 (60%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+  KDLIR++L  +P  R+SA +ALNH W+K +   S   ++   L++++ F + NKL
Sbjct: 290 ISSGAKDLIRKLLHSDPKERLSAIEALNHPWMKEDGEASDKPLDIAVLTRMKQFRAMNKL 349
```

```
Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L ++F+ +D +  GT++ +EL  G+
Sbjct: 350 KKVALKVI-AENLSEEEIVGLKEMFKSMDTDNSGTITFEELKAGL 393



>gi|5732059|gb|AAD48958.1|AF149414_7  (AF149414) similar to Pfam families PF00069
(Eukaryotic protein kinase domain; score=180.8, E=2.2e-50, N=2) and PF00036 (EF hand;
score=123.5, E=4e-33, N=1) [Arabidopsis thaliana]
          Length = 453

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 36/99 (36%), Positives = 57/99 (57%), Gaps = 1/99 (1%)
 Frame = +2

Query: 20  KDLIRRMLMYNPTNRISAADALNHEWIKNNKSKGQ-INSLTLSKLQDFDSKNKLKYAIFQ 196
           K L+ RML  NP  RISAA+ L H W+K+ ++  + I+ + LS+L+ F   NKLK    +
Sbjct: 233 KHLVNRMLNRNPKERISAAEVLGHPWMKDGEASDKPIDGVVLSRLKQFRDMNKLKKVALK 292

Query: 197 FITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
            I    ++ +E   L  +F  ID +  GT++ +EL  G+
Sbjct: 293 VIAAN-LSEEEIKGLKTLFTNIDTDKSGTITLEELKTGL 330



>gi|16754824|emb|CAC83060.1|  (AJ298072) calcium dependent calmodulin independent protein
kinase [Cucumis sativus]
          Length = 413

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 34/105 (32%), Positives = 64/105 (60%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           +S+  KDLIR++L  +P  R+SA +ALNH W+K +   S   ++   L++++ F + NKL
Sbjct: 189 ISSGAKDLIRKLLHSDPKERLSAIEALNHPWMKEDGEASDKPLDIAVLTRMKQFRAMNKL 248

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I  + ++ +E   L ++F+ +D +  GT++ +EL  G+
Sbjct: 249 KKVALKVI-AENLSEEEIVGLKEMFKSMDTDNSGTITFEELKAGL 292


>gi|1345719|sp|P49101|CDP2_MAIZE  Calcium-dependent protein kinase 2 (CDPK 2)
 gi|7428010|pir||T02259  calcium-dependent protein kinase (EC 2.7.1.-) 2 - maize
 gi|886821|gb|AAA69507.1|  (U28376) calcium-dependent protein kinase [Zea mays]
          Length = 513

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 34/105 (32%), Positives = 60/105 (56%), Gaps = 2/105 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +S S KDL+R+ML  +P  R+++A  L H+W++     S   I+S  LS+++ F + NKL
Sbjct: 292 ISESAKDLVRKMLTRDPKKRLTSAQVLQHQWLREGGEASDKPIDSAVLSRMKQFRAMNKL 351

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGI 313
           K    + I    +  +E   L ++F  +D +  GT++ +EL  G+
Sbjct: 352 KKMALKVIASN-LNEEEIKGLKQMFMNMDTDNSGTITYEELKAGL 395
```

```
 Score = 31.6 bits (70), Expect = 1.6
 Identities = 19/77 (24%), Positives = 34/77 (43%), Gaps = 3/77 (3%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+     +    L K FQ  DK+  G +++DEL    +
Sbjct: 410 LMEAADVDGNGSIDYVEFITATMHRHKLERDEHLFKAFQYFDKDNSGFITRDELESALIE 469

Query: 320 NKQ---QPIQKFIKEIN 361
           ++      I++ I E++
Sbjct: 470 HEMGDTSTIREIISEVD 486



>gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falciparum]
          Length = 513

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 42/131 (32%), Positives = 70/131 (53%), Gaps = 3/131 (2%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWI---KNNKSKGQINSLTLSKLQDFDSKNK 175
           +S+  K+LI ++L YNP  R +  +ALNH WI     +    +++S  L  L++F  +N+
Sbjct: 295 ISSDAKNLITKLLTYNPNERCTIEEALNHPWITQMTKSHEHVELSSTLLKNLKNFKKENE 354

Query: 176 LKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKE 355
           LK  I   I  + + + E  +L  IF  +D +  GT+S  E+  G+   K+   QK   +
Sbjct: 355 LK-KIALTIIAKHLCDVEINNLRNIFIALDVDNSGTLSSQEILDGL---KKIGYQKIPPD 410

Query: 356 INQQLRLLQKN 388
           I+Q LR +  N
Sbjct: 411 IHQVLRDIDSN 421



>gi|10568116|gb|AAD28192.2|  (AF115406) calcium-dependent protein kinase [Solanum tuberosum]
          Length = 532

 Score = 64.7 bits (156), Expect = 2e-10
 Identities = 37/120 (30%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIK--NNKSKGQINSLTLSKLQDFDSKNKL 178
           +++S KDL+R+ML  +P  RISAA  L H W+K     S   +++  LS+++ F + NKL
Sbjct: 311 ITSSAKDLVRKMLNKDPKQRISAAQVLEHPWLKVGGVASDKPLDNAVLSRMKQFRAMNKL 370

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++  E   L  +F  ID +  GT++ +EL  G+     +  +  +K++
Sbjct: 371 KRLALKVI-AENLSADEIQGLKSMFHNIDTDNSGTITYEELKSGLARLGSKLTEAEVKQL 429
```

```
 Score = 33.5 bits (75), Expect = 0.42
 Identities = 17/61 (27%), Positives = 29/61 (46%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+     +   +L K FQ  DK+G G +++DEL   ++
Sbjct: 429 LMEAADVDGNGSIDYIEFITATMHKHRLERDENLYKAFQYFDKDGSGFITRDELETSMEE 488

Query: 320 N 322
           +
Sbjct: 489 H 489




>gi|14029712|gb|AAK52801.1|AF363784_1  (AF363784) calcium-dependent protein kinase CDPK1
[Lycopersicon esculentum]
          Length = 521

 Score = 64.7 bits (156), Expect = 2e-10
 Identities = 37/120 (30%), Positives = 66/120 (54%), Gaps = 2/120 (1%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNN--KSKGQINSLTLSKLQDFDSKNKL 178
           VS+S KDL+++ML  +P  RI+AA  L H W+K     S   ++   LS+++ F + NKL
Sbjct: 300 VSSSAKDLVQKMLTKDPKKRITAAQVLEHPWLKEGGVASDKPLDGAVLSRMKQFRAMNKL 359

Query: 179 KYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEI 358
           K    + I  + ++ +E   L  +F  ID +  GT++ +EL  G+     +  +  +K++
Sbjct: 360 KKLALKVI-AENLSAEEIHGLKAMFHNIDTDNSGTITYEELKSGLAKLGSKLTEAEVKQL 418




 Score = 29.3 bits (64), Expect = 7.9
 Identities = 15/61 (24%), Positives = 27/61 (43%)
 Frame = +2

Query: 140 LSKLQDFDSKNKLKYAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQN 319
           L +  D D    + Y  F   T+     +   +L   FQ  DK+  G +++DEL   ++
Sbjct: 418 LMEAADVDGNGSIDYTEFITATMHKHRLERDENLYTAFQYFDKDSSGFITRDELEAAMEE 477

Query: 320 N 322
           +
Sbjct: 478 H 478




>gi|18076255|emb|CAC79947.1|  (AJ278957) protein kinase [Nyctotherus ovalis]
          Length = 465

 Score = 64.7 bits (156), Expect = 2e-10
 Identities = 36/123 (29%), Positives = 70/123 (56%), Gaps = 1/123 (0%)
 Frame = +2

Query: 5   VSNSGKDLIRRMLMYNPTNRISAADALNHEWIKNNKS-KGQINSLTLSKLQDFDSKNKLK 181
           VS+  KDL+R+ML  +   RI+AA+AL H W  +NKS K   N  ++  L+++   +KLK
Sbjct: 258 VSSEAKDLVRKMLESDYERRITAAEALKHPWFTSNKSLKTNFNEKSIEALRNYSQFSKLK 317
```

```
Query: 182 YAIFQFITVQVVTNQEKTDLLKIFQEIDKNGDGTVSKDELYQGIQNNKQQPIQKFIKEIN 361
             I   + V+ +++ E  +  + F ++DKN  G ++  +L + +     +   K ++E+
Sbjct: 318 KDILT-VLVKNISDTELKEFQEAFLQLDKNKTGMITCKDLEENLNKIGSKATAKELEELT 376

Query: 362 QQL 370
           +++
Sbjct: 377 RRI 379




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 187,862,706
Number of Sequences: 919285
Number of extensions: 3598008
Number of successful extensions: 18249
Number of sequences better than 10.0: 1697
Number of HSP's better than 10.0 without gapping: 16800
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 18162
length of database: 288,558,979
effective HSP length: 106
effective length of database: 191,114,769
effective search space used: 4586754456
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020952642-022889-23746
```

7.1.6 Query= **hy-6_11b250005**

```
         (387 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 209 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                        (bits) Value

gi|15292559|gb|AAK93548.1|  (AY052124) SD07613p [Drosophila ...     70   4e-17
gi|20129761|ref|NP_610310.1|  (NM_136466) CG2064 gene produc...     70   4e-17
gi|12861668|dbj|BAB32258.1|  (AK020927) data source:SPTR, so...     60   2e-15
gi|7290709|gb|AAF46156.1|  (AE003437) CG3842 gene product [D...     63   4e-14
gi|19921754|ref|NP_610309.1|  (NM_136465) CG2065 gene produc...     61   4e-14
gi|10190746|ref|NP_065956.1|  (NM_020905) PAN2 protein [Homo...     60   2e-13
gi|17229214|ref|NP_485762.1|  (NC_003272) probable oxidoredu...     74   4e-13
gi|12963791|ref|NP_076186.1|  (NM_023697) alcohol dehydrogen...     60   5e-13
gi|20260302|gb|AAM13049.1|  (AY093050) putative protein [Ara...     71   2e-12
gi|18413950|ref|NP_568102.1|  (NM_120332) putative protein [...     71   2e-12
gi|15827080|ref|NP_301343.1|  (NC_002677) putative oxidoredu...     61   3e-12
gi|6630624|dbj|BAA88521.1|  (AB035959) M42C60 [Mus musculus]        54   4e-12
gi|20070798|gb|AAH26274.1|  (BC026274) androgen-regulated sh...     51   4e-12
gi|14784226|ref|XP_031073.1|  (XM_031073) CGI-82 protein [Ho...     51   4e-12
gi|20070272|ref|NP_057110.2|  (NM_016026) androgen-regulated...     51   4e-12
gi|19482172|ref|NP_067532.2|  (NM_021557) short-chain dehydr...     54   4e-12
gi|5668735|dbj|BAA82657.1|  (AB030504) UBE-1b [Mus musculus]        54   4e-12
gi|5668733|dbj|BAA82656.1|  (AB030503) UBE-1a [Mus musculus]        54   4e-12
gi|10727763|gb|AAF59216.2|  (AE003840) CG17986 gene product ...     61   7e-12
gi|629592|pir||S42651  hypothetical protein - rape >gi|60657...     62   1e-11
gi|15807577|ref|NP_296314.1|  (NC_001263) oxidoreductase, sh...     62   2e-11
gi|14424501|gb|AAH09269.1|AAH09269  (BC009269) Unknown (prot...     68   2e-11
gi|17826714|emb|CAC82170.1|  (AJ293620) putative oxidoreduct...     68   2e-11
gi|18043627|gb|AAH19696.1|AAH19696  (BC019696) Unknown (prot...     68   2e-11
```

```
gi|15607210|ref|NP_214582.1|  (NC_000962) hypothetical prote...   59   3e-11
gi|15839447|ref|NP_334484.1|  (NC_002755) oxidoreductase, sh...   59   3e-11
gi|17549344|ref|NP_522684.1|  (NC_003296) PUTATIVE OXIDOREDU...   57   6e-11
gi|15609400|ref|NP_216779.1|  (NC_000962) hypothetical prote...   60   8e-11
gi|16740649|gb|AAH16204.1|AAH16204  (BC016204) Similar to RI...   44   1e-10
gi|18416145|ref|NP_567681.1|  (NM_118472) putative protein [...   65   1e-10
gi|9663977|dbj|BAB03618.1|  (AP002522) contains ESTs AU06428...   61   1e-10
gi|15224306|ref|NP_181290.1|  (NM_129309) putative oxidoredu...   65   1e-10
gi|12848558|dbj|BAB27997.1|  (AK012056) cell line MC/9.IL4 d...   54   2e-10
gi|15607580|ref|NP_214953.1|  (NC_000962) hypothetical prote...   64   3e-10
gi|15839827|ref|NP_334864.1|  (NC_002755) oxidoreductase, sh...   64   3e-10
gi|19075284|ref|NP_587784.1|  (NC_003421) hypothetical short...   63   5e-10
gi|15233359|ref|NP_192880.1|  (NM_117212) putative protein [...   63   5e-10
gi|13876535|gb|AAK43511.1|AC020666_21  (AC020666) putative W...   63   7e-10
gi|81946|pir||S20941  protochlorophyllide reductase (EC 1.3....   62   1e-09
gi|266742|sp|Q01289|POR_PEA  Protochlorophyllide reductase, ...   62   1e-09
gi|17558006|ref|NP_506570.1|  (NM_074169) dehydrogenase [Cae...   57   1e-09
gi|17826716|emb|CAC82539.1|  (AJ296079) SCAD family protein ...   52   2e-09
gi|13876532|gb|AAK43508.1|AC020666_18  (AC020666) putative W...   61   2e-09
gi|15217684|ref|NP_176640.1|  (NM_105134) oxidoreductase, pu...   60   3e-09
gi|15233666|ref|NP_194136.1|  (NM_118537) putative protein [...   60   4e-09
gi|15218860|ref|NP_171860.1|  (NM_100243) putative protochlo...   60   6e-09
gi|13676404|dbj|BAB41191.1|  (AB050124) NADPH-protochlorophy...   59   9e-09
gi|11250521|pir||T48275  hypothetical protein T22P11.130 - A...   59   9e-09
gi|15789712|ref|NP_279536.1|  (NC_002607) probable oxidoredu...   45   1e-08
gi|15239574|ref|NP_200230.1|  (NM_124799) NADPH:protochlorop...   59   1e-08
gi|10720221|sp|Q42536|PORA_ARATH  Protochlorophyllide reduct...   59   1e-08
gi|15600224|ref|NP_253718.1|  (NC_002516) probable short cha...   58   2e-08
gi|15305404|ref|XP_053194.1|  (XM_053194) hypothetical prote...   58   2e-08
gi|9049763|gb|AAF82471.1|  (AF243521) light dependent NADH:p...   58   2e-08
gi|9049769|gb|AAF82474.1|  (AF243523) light dependent NADH:p...   58   2e-08
gi|10720233|sp|Q9SDT1|POR_DAUCA  Protochlorophyllide reducta...   58   2e-08
gi|10720220|sp|Q41249|PORA_CUCSA  Protochlorophyllide reduct...   57   5e-08
gi|15146202|gb|AAK83584.1|  (AY049242) AT5g50130/MPF21_15 [A...   57   5e-08
gi|18423110|ref|NP_568721.1|  (NM_124392) ribitol dehydrogen...   57   5e-08
gi|16124653|ref|NP_419217.1|  (NC_002696) oxidoreductase, sh...   57   5e-08
gi|10177224|dbj|BAB10299.1|  (AB026650) ribitol dehydrogenas...   57   5e-08
gi|13676400|dbj|BAB41189.1|  (AB050122) NADPH-protochlorophy...   56   6e-08
gi|17433971|ref|XP_059348.1|  (XM_059348) similar to PAN2 pr...   56   6e-08
gi|15234129|ref|NP_194474.1|  (NM_118879) protochlorophyllid...   56   6e-08
gi|16119519|ref|NP_396225.1|  (NC_003064) AGR_pAT_417p [Agro...   56   6e-08
gi|99752|pir||S14990  protochlorophyllide reductase (EC 1.3....   56   6e-08
gi|17938875|ref|NP_535663.1|  (NC_003306) dehydrogenase [Agr...   56   6e-08
gi|10720232|sp|Q39617|POR_CHLRE  Protochlorophyllide reducta...   56   8e-08
gi|9587209|gb|AAF89208.1|AF279251_1  (AF279251) NADPH-protoc...   56   8e-08
gi|13699822|gb|AAD20020.2|  (AF126871) NADPH-protochlorophyl...   56   8e-08
gi|20347241|ref|XP_109906.1|  (XM_109906) similar to androge...   42   9e-08
gi|129707|sp|P15904|POR_AVESA  Protochlorophyllide reductase...   55   1e-07
gi|10720231|sp|O80333|POR_MARPA  Protochlorophyllide reducta...   55   1e-07
gi|19343615|gb|AAH25724.1|  (BC025724) similar to unnamed pr...   55   1e-07
gi|16549449|dbj|BAB70811.1|  (AK054835) unnamed protein prod...   55   1e-07
gi|7485551|pir||T05381  hypothetical protein F16G20.130 - Ar...   55   1e-07
gi|129708|sp|P13653|PORA_HORVU  Protochlorophyllide reductas...   55   1e-07
gi|10720235|sp|Q41578|PORA_WHEAT  Protochlorophyllide reduct...   55   1e-07
gi|2598163|gb|AAB86734.1|  (AF027356) NADPH:protochlorophyll...   55   1e-07
gi|7481135|pir||T37150  probable oxidoreductase - Streptomyc...   55   2e-07
gi|7330644|gb|AAC60560.2|  (S63824) NADPH-protochlorophyllid...   54   2e-07
gi|419742|pir||S30167  protochlorophyllide reductase (EC 1.3...   54   2e-07
gi|7300387|gb|AAF55546.1|  (AE003721) CG7675 gene product [a...   54   2e-07
gi|7300388|gb|AAF55547.1|  (AE003721) CG7675 gene product [a...   54   2e-07
gi|9049771|gb|AAF82475.1|  (AF243522) light dependent NADH:p...   54   2e-07
```

```
gi|19922628|ref|NP_611471.1|  (NM_137627) CG11200 gene produ...    54   3e-07
gi|7304178|gb|AAF59214.1|  (AE003840) CG2070 gene product [D...    54   3e-07
gi|17229235|ref|NP_485783.1|  (NC_003272) protochlorophyllid...    49   3e-07
gi|6759586|emb|CAB69779.1|  (AL137187) putative oxidoreducta...    54   4e-07
gi|8927391|gb|AAF82054.1|AF227527_1  (AF227527) FOR II prote...    45   5e-07
gi|7706523|ref|NP_057457.1|  (NM_016373) WW domain-containin...    45   5e-07
gi|13699258|ref|NP_061030.2|  (NM_018560) WW domain-containi...    45   5e-07
gi|20129347|ref|NP_609171.1|  (NM_135327) CG7221 gene produc...    45   7e-07
gi|18071421|gb|AAL58280.1|AC068923_22  (AC068923) putative d...    52   1e-06
gi|10720236|sp|Q42850|PORB_HORVU  Protochlorophyllide reduct...    52   1e-06
gi|17532805|ref|NP_495501.1|  (NM_063100) E04F6.15.p [Caenor...    52   1e-06
gi|9625012|ref|NP_062519.1|  (NM_019573) WW-domain oxidoredu...    44   1e-06
gi|15928476|gb|AAH14716.1|AAH14716  (BC014716) WW domain-con...    44   1e-06
gi|12860307|dbj|BAB31911.1|  (AK019911) WW-DOMAIN OXIDOREDUC...    44   1e-06
gi|5734629|dbj|BAA83360.1|  (AP000391) Similar to ribitol de...    51   3e-06
```

Alignments

```
>gi|15292559|gb|AAK93548.1|  (AY052124) SD07613p [Drosophila melanogaster]
          Length = 330

 Score = 69.7 bits (169), Expect(2) = 4e-17
 Identities = 38/67 (56%), Positives = 50/67 (73%), Gaps = 1/67 (1%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLGHF 339
           +LDLS  +SI++ V+ FKK + P++ +LINNAGVM  P+T   TK  YELQ G NH+GHF
Sbjct: 100 ELDLSSLDSIRKFVDGFKKEQ-PKLHVLINNAGVMRCPKTL--TKDGYELQLGVNHIGHF 156

Query: 340 LLTELLV 360
           LLT LL+
Sbjct: 157 LLTNLLL 163




 Score = 37.4 bits (85), Expect(2) = 4e-17
 Identities = 18/51 (35%), Positives = 29/51 (56%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIEQNNQSL 155
           NTGIGKE AL++     T+ +ACRD  + ++    ++      E NNQ++
Sbjct: 52  NTGIGKETALEIARRGGTVYLACRDMNRCEKARKDII------KETNNQNI 96



>gi|20129761|ref|NP_610310.1| ☐ (NM_136466) CG2064 gene product [Drosophila melanogaster]
 gi|10727765|gb|AAF59212.2| ☐ (AE003840) CG2064 gene product [Drosophila melanogaster]
          Length = 309

 Score = 69.7 bits (169), Expect(2) = 4e-17
 Identities = 38/67 (56%), Positives = 50/67 (73%), Gaps = 1/67 (1%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLGHF 339
           +LDLS  +SI++ V+ FKK + P++ +LINNAGVM  P+T   TK  YELQ G NH+GHF
Sbjct: 100 ELDLSSLDSIRKFVDGFKKEQ-PKLHVLINNAGVMRCPKTL--TKDGYELQLGVNHIGHF 156
```

```
Query: 340 LLTELLV 360
           LLT LL+
Sbjct: 157 LLTNLLL 163




 Score = 37.4 bits (85), Expect(2) = 4e-17
 Identities = 18/51 (35%), Positives = 29/51 (56%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIEQNNQSL 155
           NTGIGKE AL++     T+ +ACRD  + ++     ++      E NNQ++
Sbjct: 52  NTGIGKETALEIARRGGTVYLACRDMNRCEKARKDII------KETNNQNI 96



>gi|12861668|dbj|BAB32258.1| ☐ (AK020927) data source:SPTR, source key:Q9NRW0,
evidence:ISS~homolog to ANDROGEN-REGULATED SHORT-CHAIN DEHYDROGENASE/REDUCTASE 1~putative
[Mus musculus]
          Length = 316

 Score = 60.1 bits (144), Expect(2) = 2e-15
 Identities = 36/81 (44%), Positives = 50/81 (61%)
 Frame = +1

Query: 118 NQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQ 297
           ++++   K ++  + +LDLS   SI+   E F   +  ++DILINNAGVM     KTT
Sbjct: 81  SEIRADTKNSQVLVRKLDLSDTKSIRAFAERFLAEE-KKLDILINNAGVMMCPYSKTT-D 138

Query: 298 SYELQFGTNHLGHFLLTELLV 360
            +E  FG NHLGHFLLT LL+
Sbjct: 139 GFETHFGVNHLGHFLLTYLLL 159




 Score = 41.2 bits (95), Expect(2) = 2e-15
 Identities = 17/36 (47%), Positives = 23/36 (63%)
 Frame = +3


Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           NTGIGKE A +L    A + IACRD +KG+    ++
Sbjct: 48  NTGIGKETARELARRGARVYIACRDVLKGESAASEI 83



>gi|7290709|gb|AAF46156.1| ☐ (AE003437) CG3842 gene product [Drosophila melanogaster]
          Length = 406

 Score = 63.2 bits (152), Expect(2) = 4e-14
 Identities = 37/66 (56%), Positives = 45/66 (68%), Gaps = 1/66 (1%)
 Frame = +1
```

```
Query: 166 LDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMA-PQTYKTTKQSYELQFGTNHLGHFL 342
           LDL    S++  VE FK  +  ++DILINNAGVMA P+T   T   +E QFG NHLGHFL
Sbjct: 132 LDLGSLQSVRNFVERFKAEE-SRLDILINNAGVMACPRTL--TADGFEQQFGVNHLGHFL 188

Query: 343 LTELLV 360
           LT LL+
Sbjct: 189 LTNLLL 194



 Score = 33.9 bits (76), Expect(2) = 4e-14
 Identities = 14/25 (56%), Positives = 17/25 (68%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRD 77
           NTGIGKE  L+L    A + +ACRD
Sbjct: 83  NTGIGKETVLELAKRGARVYMACRD 107



>gi|19921754|ref|NP_610309.1| ☐ (NM_136465) CG2065 gene product [Drosophila melanogaster]
 gi|7304177|gb|AAF59213.1| ☐ (AE003840) CG2065 gene product [Drosophila melanogaster]
 gi|17946603|gb|AAL49332.1| ☐ (AY071710) RH23455p [Drosophila melanogaster]
          Length = 300

 Score = 61.2 bits (147), Expect(2) = 4e-14
 Identities = 34/67 (50%), Positives = 46/67 (67%), Gaps = 1/67 (1%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLGHF 339
           +LDLS   SI++    FKK +  ++ +LINNAGVM  P+T   TK  +E+Q G NH+GHF
Sbjct: 71  ELDLSSLESIRKFAAGFKKEQ-DKLHVLINNAGVMHCPRTL--TKDGFEMQLGVNHMGHF 127

Query: 340 LLTELLV 360
           LLT LL+
Sbjct: 128 LLTHLLL 134



 Score = 35.8 bits (81), Expect(2) = 4e-14
 Identities = 17/51 (33%), Positives = 28/51 (54%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIEQNNQSL 155
           NTGIGKE  L++     T+ +ACRD  + ++    ++       E NNQ++
Sbjct: 23  NTGIGKETVLEIAKRGGTVYMACRDMNRCEKARQDII------RETNNQNI 67



>gi|10190746|ref|NP_065956.1| ☐ (NM_020905) PAN2 protein [Homo sapiens]
 gi|10039619|gb|AAG12190.1|AF237952_1 ☐ (AF237952) PAN2 [Homo sapiens]
 gi|14602625|gb|AAH09830.1|AAH09830 ☐ (BC009830) PAN2 protein [Homo sapiens]
```

```
          Length = 336

 Score = 60.1 bits (144), Expect(2) = 2e-13
 Identities = 30/77 (38%), Positives = 50/77 (63%)
 Frame = +1

Query: 130 GQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYEL 309
           G + + +   + +LDL+   S++    ++   + + P++D+LINNAG+      Y   T+  +E+
Sbjct: 102 GVSGVGELIVRELDLASLRSVRAFCQEMLQEE-PRLDVLINNAGIFQCP-YMKTEDGFEM 159

Query: 310 QFGTNHLGHFLLTELLV 360
           QFG NHLGHFLLT LL+
Sbjct: 160 QFGVNHLGHFLLTNLLL 176




 Score = 34.7 bits (78), Expect(2) = 2e-13
 Identities = 15/45 (33%), Positives = 26/45 (57%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIE 137
           N+G+G+   A +L   + A +I+ CRD   + ++   QL   +L Q   E
Sbjct: 52  NSGLGRATAAELLRLGARVIMGCRDRARAEEAAGQLRRELRQAAE 96
```

```
>gi|17229214|ref|NP_485762.1|  (NC_003272) probable oxidoreductase [Nostoc sp. PCC 7120]
 gi|17130812|dbj|BAB73421.1|  (AP003587) ORF_ID:alr1722~probable oxidoreductase [Nostoc sp.
PCC 7120]
          Length = 311

 Score = 73.6 bits (179), Expect = 4e-13
 Identities = 41/85 (48%), Positives = 56/85 (65%)
 Frame = +1

Query: 106 NQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYK 285
           N+   ++  QNK     L++LDL+   S+K   E+F+K   + ++D+LINNAGVM P   K
Sbjct: 54  NKALAKILQQNKDADVKLMELDLANLASVKNFAENFRKNYL-RLDLLINNAGVMIPPYSK 112

Query: 286 TTKQSYELQFGTNHLGHFLLTELLV 360
           TT   +ELQFGTNHLGHF LT  L+
Sbjct: 113 TT-DGFELQFGTNHLGHFALTGQLL 136




 Score = 32.0 bits (71), Expect = 1.2
 Identities = 16/39 (41%), Positives = 25/39 (64%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQ 119
           ++GIG E A  L N  A++IIA R+  KG + L +++ Q
Sbjct: 25  SSGIGYETARVLANKQASVIIAVRNLDKGNKALAKILQQ 63
```

```
>gi|12963791|ref|NP_076186.1|  (NM_023697) alcohol dehydrogenase PAN2; expressed sequence
AW215439 [Mus musculus]
 gi|11120506|gb|AAG30904.1|AF303831_1  (AF303831) alcohol dehydrogenase PAN2 [Mus
musculus]
 gi|18043332|gb|AAH20094.1|AAH20094  (BC020094) RIKEN cDNA 3110030G19 gene [Mus musculus]
          Length = 334

 Score = 59.7 bits (143), Expect(2) = 5e-13
 Identities = 30/66 (45%), Positives = 45/66 (67%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFL 342
           +LDL+   S++   ++  + + P++D+LINNAGV     Y  T+  +E+QFG NHLGHFL
Sbjct: 111 ELDLASLRSVRAFCQELLQEE-PRLDVLINNAGVFHCP-YTKTEDGFEMQFGVNHLGHFL 168

Query: 343 LTELLV 360
           LT LL+
Sbjct: 169 LTNLLL 174




 Score = 33.5 bits (75), Expect(2) = 5e-13
 Identities = 14/42 (33%), Positives = 25/42 (59%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQ 128
           N+G+G+  A +L  + A +I+ CRD  + ++   QL  +L Q
Sbjct: 53  NSGLGRATAAELLRLGARVIMGCRDRARAEEAAGQLRQELCQ 94



>gi|20260302|gb|AAM13049.1|  (AY093050) putative protein [Arabidopsis thaliana]
          Length = 331

 Score = 71.2 bits (173), Expect = 2e-12
 Identities = 39/80 (48%), Positives = 53/80 (65%)
 Frame = +1

Query: 121 QVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQS 300
           ++  QN   +  L+QLDLS   SIK  V +F  L +P +++LINNAGVM    Y+ ++
Sbjct: 76  EILRQNANARVTLLQLDLSSIKSIKAFVREFHALHLP-LNLLINNAGVMFCP-YQLSEDG 133

Query: 301 YELQFGTNHLGHFLLTELLV 360
            ELQF TNH+GHFLLT LL+
Sbjct: 134 IELQFATNHIGHFLLTNLLL 153



>gi|18413950|ref|NP_568102.1|  (NM_120332) putative protein [Arabidopsis thaliana]
          Length = 331

 Score = 71.2 bits (173), Expect = 2e-12
 Identities = 39/80 (48%), Positives = 53/80 (65%)
 Frame = +1
```

```
Query: 121 QVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQS 300
           ++  QN   +  L+QLDLS   SIK  V +F  L +P +++LINNAGVM    Y+ ++
Sbjct: 76  EILRQNANARVTLLQLDLSSIKSIKAFVREFHALHLP-LNLLINNAGVMFCP-YQLSEDG 133

Query: 301 YELQFGTNHLGHFLLTELLV 360
            ELQF TNH+GHFLLT LL+
Sbjct: 134 IELQFATNHIGHFLLTNLLL 153



>gi|15827080|ref|NP_301343.1|  (NC_002677) putative oxidoreductase [Mycobacterium leprae]
 gi|11250835|pir||T44727  probable oxidoreductase [imported] - Mycobacterium leprae
 gi|4154042|emb|CAA22691.1|  (AL035159) putative oxidoreductase [Mycobacterium leprae]
 gi|13092628|emb|CAC29823.1|  (AL583918) putative oxidoreductase [Mycobacterium leprae]
          Length = 304

 Score = 60.8 bits (146), Expect(2) = 3e-12
 Identities = 34/69 (49%), Positives = 45/69 (64%), Gaps = 1/69 (1%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLG 333
           L +LDL+   S++   +  +      ID+LINNAGVM  P++  TTK  +ELQFGTNHLG
Sbjct: 69  LQELDLASLESVRAAAKQLRS-DYDHIDLLINNAGVMWTPKS--TTKDGFELQFGTNHLG 125

Query: 334 HFLLTELLV 360
           HF  T LL+
Sbjct: 126 HFAFTGLLL 134



 Score = 29.6 bits (65), Expect(2) = 3e-12
 Identities = 16/54 (29%), Positives = 29/54 (53%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIEQNNQSLLDI 164
           NTG+G + AL L    A +++A R+  KG+        +++    QNN +L ++
Sbjct: 23  NTGLGYQTALALAEHGAHVVLAVRNLDKGKDA----AARITATSAQNNVALQEL 72



>gi|6630624|dbj|BAA88521.1|   (AB035959) M42C60 [Mus musculus]
          Length = 355

 Score = 53.5 bits (127), Expect(2) = 4e-12
 Identities = 30/72 (41%), Positives = 44/72 (60%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           ++ ++ +LDL+   SI+   +DF   +   + +LINNAGVM    Y  T   +E+  G N
Sbjct: 89  SQVFVRKLDLADTKSIRAFAKDFLAEE-KHLHLLINNAGVMMCP-YSKTADGFEMHIGVN 146

Query: 325 HLGHFLLTELLV 360
           HLGHFLLT LL+
Sbjct: 147 HLGHFLLTHLLL 158
```

```
 Score = 36.6 bits (83), Expect(2) = 4e-12
 Identities = 16/30 (53%), Positives = 19/30 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQ 92
            NTGIGKE A  L    A + +ACRD  KG+
Sbjct: 47  NTGIGKETAKDLAQRGARVYLACRDVDKGE 76



>gi|20070798|gb|AAH26274.1| ☐ (BC026274) androgen-regulated short-chain
dehydrogenase/reductase 1 [Homo sapiens]
          Length = 318

 Score = 51.2 bits (121), Expect(2) = 4e-12
 Identities = 30/66 (45%), Positives = 39/66 (58%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFL 342
            +LDLS   SI+    + F    +    + +LINNAGVM     Y  T    +E+   G NHLGHFL
Sbjct: 98  KLDLSDTKSIRAFAKGFLAEE-KHLHVLINNAGVMMCP-YSKTADGFEMHIGVNHLGHFL 155

Query: 343 LTELLV 360
            LT LL+
Sbjct: 156 LTHLLL 161



 Score = 38.9 bits (89), Expect(2) = 4e-12
 Identities = 17/36 (47%), Positives = 23/36 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
            NTGIGKE A +L    A + +ACRD  KG+ V  ++
Sbjct: 50  NTGIGKETAKELAQRGARVYLACRDVEKGELVAKEI 85



>gi|14784226|ref|XP_031073.1| ☐ (XM_031073) CGI-82 protein [Homo sapiens]
 gi|4929633|gb|AAD34077.1|AF151840_1 ☐ (AF151840) CGI-82 protein [Homo sapiens]
 gi|12652725|gb|AAH00112.1|AAH00112 ☐ (BC000112) CGI-82 protein [Homo sapiens]
 gi|14669795|gb|AAK72049.1|AF395068_1 ☐ (AF395068) HCV core-binding protein HCBP12 [Homo
sapiens]
 gi|15079855|gb|AAH11727.1|AAH11727 ☐ (BC011727) Similar to CGI-82 protein [Homo sapiens]
          Length = 318

 Score = 51.2 bits (121), Expect(2) = 4e-12
 Identities = 30/66 (45%), Positives = 39/66 (58%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFL 342
            +LDLS   SI+    + F    +    + +LINNAGVM     Y  T    +E+   G NHLGHFL
Sbjct: 98  KLDLSDTKSIRAFAKGFLAEE-KHLHVLINNAGVMMCP-YSKTADGFEMHIGVNHLGHFL 155
```

```
Query: 343 LTELLV 360
           LT LL+
Sbjct: 156 LTHLLL 161




 Score = 38.9 bits (89), Expect(2) = 4e-12
 Identities = 17/36 (47%), Positives = 23/36 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           NTGIGKE A +L    A + +ACRD  KG+ V  ++
Sbjct: 50  NTGIGKETAKELAQRGARVYLACRDVEKGELVAKEI 85



>gi|20070272|ref|NP_057110.2| (NM_016026) androgen-regulated short-chain
dehydrogenase/reductase 1; prostate short-chain dehydrogenase reductase 1; likely ortholog
of mouse cell line MC/9.IL4 derived transcript 1; CGI-82 protein; HCV core-binding protein
[Homo sapiens]
 gi|9622124|gb|AAF89632.1|AF167438_1 (AF167438) androgen-regulated short-chain
dehydrogenase/reductase 1 [Homo sapiens]
          Length = 318

 Score = 51.2 bits (121), Expect(2) = 4e-12
 Identities = 30/66 (45%), Positives = 39/66 (58%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFL 342
           +LDLS   SI+   + F   +   + +LINNAGVM    Y  T   +E+  G NHLGHFL
Sbjct: 98  KLDLSDTKSIRAFAKGFLAEE-KHLHVLINNAGVMMCP-YSKTADGFEMHIGVNHLGHFL 155

Query: 343 LTELLV 360
           LT LL+
Sbjct: 156 LTHLLL 161




 Score = 38.9 bits (89), Expect(2) = 4e-12
 Identities = 17/36 (47%), Positives = 23/36 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           NTGIGKE A +L    A + +ACRD  KG+ V  ++
Sbjct: 50  NTGIGKETAKELAQRGARVYLACRDVEKGELVAKEI 85



>gi|19482172|ref|NP_067532.2| (NM_021557) short-chain dehydrogenase/reductase 1; short-
chain aldehyde dehydrogenase [Mus musculus]
 gi|12835589|dbj|BAB23296.1| (AK004413) cell line MC/9.IL4 derived transcript 1~data
source:MGD, source key:MGI:102581, evidence:ISS~putative [Mus musculus]
 gi|17223687|gb|AAK91516.1| (AY039032) short-chain dehydrogenase/reductase [Mus musculus]
```

```
 gi|17390609|gb|AAH18261.1|AAH18261  (BC018261) cell line MC/9.IL4 derived transcript 1 [Mus
musculus]
 gi|18874690|gb|AAL79910.1|AF474027_1  (AF474027) short-chain aldehyde dehydrogenase
SCALD [Mus musculus]
          Length = 316

 Score = 53.5 bits (127), Expect(2) = 4e-12
 Identities = 30/72 (41%), Positives = 44/72 (60%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           ++ ++ +LDL+   SI+   +DF   +   + +LINNAGVM    Y  T   +E+  G N
Sbjct: 89  SQVFVRKLDLADTKSIRAFAKDFLAEE-KHLHLLINNAGVMMCP-YSKTADGFEMHIGVN 146

Query: 325 HLGHFLLTELLV 360
           HLGHFLLT LL+
Sbjct: 147 HLGHFLLTHLLL 158



 Score = 36.6 bits (83), Expect(2) = 4e-12
 Identities = 16/30 (53%), Positives = 19/30 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQ 92
           NTGIGKE A  L    A + +ACRD  KG+
Sbjct: 47  NTGIGKETAKDLAQRGARVYLACRDVDKGE 76



>gi|5668735|dbj|BAA82657.1|  (AB030504) UBE-1b [Mus musculus]
          Length = 300

 Score = 53.5 bits (127), Expect(2) = 4e-12
 Identities = 30/72 (41%), Positives = 44/72 (60%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           ++ ++ +LDL+   SI+   +DF   +   + +LINNAGVM    Y  T   +E+  G N
Sbjct: 73  SQVFVRKLDLADTKSIRAFAKDFLAEE-KHLHLLINNAGVMMCP-YSKTADGFEMHIGVN 130

Query: 325 HLGHFLLTELLV 360
           HLGHFLLT LL+
Sbjct: 131 HLGHFLLTHLLL 142



 Score = 36.6 bits (83), Expect(2) = 4e-12
 Identities = 16/30 (53%), Positives = 19/30 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQ 92
           NTGIGKE A  L    A + +ACRD  KG+
Sbjct: 31  NTGIGKETAKDLAQRGARVYLACRDVDKGE 60
```

```
>gi|5668733|dbj|BAA82656.1|  (AB030503) UBE-1a [Mus musculus]
          Length = 293

 Score = 53.5 bits (127), Expect(2) = 4e-12
 Identities = 30/72 (41%), Positives = 44/72 (60%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           ++ ++ +LDL+   SI+   +DF   +   + +LINNAGVM    Y  T   +E+  G N
Sbjct: 66  SQVFVRKLDLADTKSIRAFAKDFLAEE-KHLHLLINNAGVMMCP-YSKTADGFEMHIGVN 123

Query: 325 HLGHFLLTELLV 360
           HLGHFLLT LL+
Sbjct: 124 HLGHFLLTHLLL 135



 Score = 36.6 bits (83), Expect(2) = 4e-12
 Identities = 16/30 (53%), Positives = 19/30 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQ 92
           NTGIGKE A  L    A + +ACRD  KG+
Sbjct: 24  NTGIGKETAKDLAQRGARVYLACRDVDKGE 53
```

```
>gi|10727763|gb|AAF59216.2|  (AE003840) CG17986 gene product [Drosophila melanogaster]
          Length = 594

 Score = 60.8 bits (146), Expect = 2e-09
 Identities = 33/90 (36%), Positives = 55/90 (60%)
 Frame = +1

Query: 91  NKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMA 270
           NK  + +   ++  +   +  +  + DLS  +SI++  E+FKK +   + ILINNAGV
Sbjct: 322 NKEKVERARREIVKETGNSNVFSRECDLSSLDSIRKFAENFKKEQ-RVLHILINNAGVFW 380

Query: 271 PQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
            + ++ TK+ +E+  G NH+GHFLLT LL+
Sbjct: 381 -EPHRLTKEGFEMHLGVNHIGHFLLTNLLL 409



 Score = 54.7 bits (130), Expect(2) = 7e-12
 Identities = 34/77 (44%), Positives = 45/77 (58%), Gaps = 1/77 (1%)
 Frame = +1

Query: 133 QNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYEL 309
           + K    Y  Q DL+   SI+  V  FK+ +   + +LINNAGVM  P++   T    EL
Sbjct: 92  ETKNKYVYCRQCDLASQESIRHFVAAFKREQ-EHLHVLINNAGVMRCPRSL--TSDGIEL 148

Query: 310 QFGTNHLGHFLLTELLV 360
           Q G NH+GHFLLT LL+
Sbjct: 149 QLGVNHMGHFLLTNLLL 165
```

```
 Score = 34.7 bits (78), Expect = 0.19
 Identities = 14/37 (37%), Positives = 24/37 (64%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLV 113
            NTG+GKE  ++L    AT+ +ACR+  K ++   ++V
Sbjct: 298 NTGLGKETVMELARRGATVYMACRNKEKVERARREIV 334




 Score = 34.7 bits (78), Expect(2) = 7e-12
 Identities = 13/43 (30%), Positives = 25/43 (57%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQR 131
            NTGIGKE   ++      T+ +ACR+  K ++   +++V +   +
Sbjct: 54  NTGIGKETVREIAKRGGTVYMACRNLKKCEEAREEIVLETKNK 96



>gi|629592|pir||S42651  hypothetical protein - rape
 gi|6065752|emb|CAB58175.1|  (X74225) putative pod-specific dehydrogenase SAC25 [Brassica
napus]
          Length = 320

 Score = 61.6 bits (148), Expect(2) = 1e-11
 Identities = 32/67 (47%), Positives = 45/67 (66%)
 Frame = +1

Query: 160 IQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHF 339
            +QLDLS   S++  +  F  L +P ++ILINNAGVM    ++ ++   E QF TNH+GHF
Sbjct: 89  LQLDLSSIKSVRSFIHQFLALNVP-LNILINNAGVMFCP-FQLSEDGIESQFATNHIGHF 146

Query: 340 LLTELLV 360
            LLT LL+
Sbjct: 147 LLTNLLL 153




 Score = 26.9 bits (58), Expect(2) = 1e-11
 Identities = 14/25 (56%), Positives = 17/25 (68%)
 Frame = +3

Query: 6   TGIGKEAALQLGNMNATIIIACRDT 80
            +GIG EAA  LG   A +IIA R+T
Sbjct: 43  SGIGLEAARVLGMRGAHVIIASRNT 67



>gi|15807577|ref|NP_296314.1|  (NC_001263) oxidoreductase, short-chain
dehydrogenase/reductase family [Deinococcus radiodurans]
```

```
 gi|7473258|pir||H75255  oxidoreductase, short-chain dehydrogenase/reductase family -
Deinococcus radiodurans (strain R1)
 gi|6460419|gb|AAF12130.1|AE002088_7  (AE002088) oxidoreductase, short-chain
dehydrogenase/reductase family [Deinococcus radiodurans]
          Length = 336

 Score = 62.4 bits (150), Expect(2) = 2e-11
 Identities = 34/68 (50%), Positives = 45/68 (66%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGH 336
           L+ LDL    S+++   + ++L  P+I ILINNAGVMA    +T    +E QFGTNHLGH
Sbjct: 97  LVDLDLGSLASVRRGAAEIRQLA-PRIHILINNAGVMATPQSRTV-DGFETQFGTNHLGH 154

Query: 337 FLLTELLV 360
           FLLT  L+
Sbjct: 155 FLLTRELL 162



 Score = 25.8 bits (55), Expect(2) = 2e-11
 Identities = 13/35 (37%), Positives = 20/35 (57%)
 Frame = +3

Query: 6   TGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           +G+G E A  L    A +I+  RD  KG++V  +L
Sbjct: 53  SGLGTETARALLLAGAHVILPVRDRAKGERVAAEL 87



>gi|14424501|gb|AAH09269.1|AAH09269  (BC009269) Unknown (protein for IMAGE:3140944) [Homo
sapiens]
          Length = 288

 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 40/92 (43%), Positives = 56/92 (60%), Gaps = 1/92 (1%)
 Frame = +1

Query: 88  DNKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGV- 264
           +N S   Q  +++K +    K   +  DL+   SI+Q V+ FK  KIP + +LINNAGV
Sbjct: 33  NNDSKAKQVVSKIKEETLNDKVEFLYCDLASMTSIRQFVQKFKMKKIP-LHVLINNAGVM 91

Query: 265 MAPQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
           M PQ  + T+  +E  FG N+LGHFLLT LL+
Sbjct: 92  MVPQ--RKTRDGFEEHFGLNYLGHFLLTNLLL 121



>gi|17826714|emb|CAC82170.1|  (AJ293620) putative oxidoreductase [Homo sapiens]
          Length = 330

 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 40/92 (43%), Positives = 56/92 (60%), Gaps = 1/92 (1%)
 Frame = +1

Query: 88  DNKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGV- 264
           +N S   Q  +++K +    K   +  DL+   SI+Q V+ FK  KIP + +LINNAGV
Sbjct: 75  NNDSKAKQVVSKIKEETLNDKVEFLYCDLASMTSIRQFVQKFKMKKIP-LHVLINNAGVM 133
```

```
Query: 265 MAPQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
           M PQ  + T+  +E  FG N+LGHFLLT LL+
Sbjct: 134 MVPQ--RKTRDGFEEHFGLNYLGHFLLTNLLL 163




>gi|18043627|gb|AAH19696.1|AAH19696  (BC019696) Unknown (protein for IMAGE:4398682) [Homo
sapiens]
          Length = 334

 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 40/92 (43%), Positives = 56/92 (60%), Gaps = 1/92 (1%)
 Frame = +1

Query: 88  DNKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGV- 264
           +N S   Q  +++K +    K   +  DL+   SI+Q V+ FK  KIP + +LINNAGV
Sbjct: 79  NNDSKAKQVVSKIKEETLNDKVEFLYCDLASMTSIRQFVQKFKMKKIP-LHVLINNAGVM 137

Query: 265 MAPQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
           M PQ  + T+  +E  FG N+LGHFLLT LL+
Sbjct: 138 MVPQ--RKTRDGFEEHFGLNYLGHFLLTNLLL 167




>gi|15607210|ref|NP_214582.1|  (NC_000962) hypothetical protein Rv0068 [Mycobacterium
tuberculosis H37Rv]
 gi|7478509|pir||E70848  probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
 gi|2808725|emb|CAA16249.1|  (AL021428) hypothetical protein Rv0068 [Mycobacterium
tuberculosis H37Rv]
          Length = 303

 Score = 59.3 bits (142), Expect(2) = 3e-11
 Identities = 33/68 (48%), Positives = 43/68 (62%)
 Frame = +1


Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGH 336
           L +LDL+   S++      K     +ID+LINNAGVM     +TT   +E+QFGTNHLGH
Sbjct: 69  LQELDLTSLASVRAAAAQLKS-DHQRIDLLINNAGVMYTPR-QTTADGFEMQFGTNHLGH 126

Query: 337 FLLTELLV 360
           F LT LL+
Sbjct: 127 FALTGLLI 134





 Score = 28.1 bits (61), Expect(2) = 3e-11
 Identities = 13/31 (41%), Positives = 19/31 (60%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQ 95
           NTG+G E A  L    A +++A R+  KG+Q
Sbjct: 23  NTGLGFETAAALAAHGAHVVLAVRNLDKGKQ 53
```

```
>gi|15839447|ref|NP_334484.1|  (NC_002755) oxidoreductase, short-chain
dehydrogenase/reductase family [Mycobacterium tuberculosis CDC1551]
 gi|13879120|gb|AAK44298.1|  (AE006919) oxidoreductase, short-chain dehydrogenase/reductase
family [Mycobacterium tuberculosis CDC1551]
          Length = 303

 Score = 59.3 bits (142), Expect(2) = 3e-11
 Identities = 33/68 (48%), Positives = 43/68 (62%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGH 336
           L +LDL+   S++      K     +ID+LINNAGVM     +TT   +E+QFGTNHLGH
Sbjct: 69  LQELDLTSLASVRAAAAQLKS-DHQRIDLLINNAGVMYTPR-QTTADGFEMQFGTNHLGH 126

Query: 337 FLLTELLV 360
           F LT LL+
Sbjct: 127 FALTGLLI 134




 Score = 28.1 bits (61), Expect(2) = 3e-11
 Identities = 13/31 (41%), Positives = 19/31 (60%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQ 95
           NTG+G E A  L    A +++A R+  KG+Q
Sbjct: 23  NTGLGFETAAALAAHGAHVVLAVRNLDKGKQ 53




>gi|17549344|ref|NP_522684.1|  (NC_003296) PUTATIVE OXIDOREDUCTASE PROTEIN [Ralstonia
solanacearum]
 gi|17431597|emb|CAD18274.1|  (AL646083) PUTATIVE OXIDOREDUCTASE PROTEIN [Ralstonia
solanacearum]
          Length = 306

 Score = 56.6 bits (135), Expect(2) = 6e-11
 Identities = 32/67 (47%), Positives = 42/67 (61%)
 Frame = +1

Query: 160 IQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHF 339
           + LDL+   S++      + +  P+ID LINNAGVM P   +TT Q +ELQ G NHLG F
Sbjct: 73  LPLDLADLASVRAAAG--QAMTAPRIDALINNAGVMMPPLMRTT-QGFELQLGVNHLGGF 129

Query: 340 LLTELLV 360
            LT LL+
Sbjct: 130 ALTALLL 136




 Score = 29.6 bits (65), Expect(2) = 6e-11
 Identities = 12/33 (36%), Positives = 19/33 (57%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVL 101
           NTG+G E +  L    A +++ACRD  K +  +
Sbjct: 26  NTGLGLETSRVLAARGARVLLACRDRSKAEAAM 58
```

```
>gi|15609400|ref|NP_216779.1|  (NC_000962) hypothetical protein Rv2263 [Mycobacterium
tuberculosis H37Rv]
 gi|15841754|ref|NP_336791.1|  (NC_002755) oxidoreductase, short-chain
dehydrogenase/reductase family [Mycobacterium tuberculosis CDC1551]
 gi|7477134|pir||C70863  hypothetical protein Rv2263 - Mycobacterium tuberculosis  (strain
H37RV)
 gi|2909569|emb|CAA17300.1|  (AL021925) hypothetical protein Rv2263 [Mycobacterium
tuberculosis H37Rv]
 gi|13882013|gb|AAK46605.1|  (AE007075) oxidoreductase, short-chain dehydrogenase/reductase
family [Mycobacterium tuberculosis CDC1551]
          Length = 317

 Score = 59.7 bits (143), Expect(2) = 8e-11
 Identities = 34/65 (52%), Positives = 39/65 (59%)
 Frame = +1

Query: 166 LDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFLL 345
           LDLS   S+    E       P ID+LINNAGVM P    TT   +ELQFG+NHLGHF L
Sbjct: 72  LDLSSLASVAALGEQLMADGRP-IDLLINNAGVMTPPERVTTADGFELQFGSNHLGHFAL 130

Query: 346 TELLV 360
           T  L+
Sbjct: 131 TAHLL 135




 Score = 26.2 bits (56), Expect(2) = 8e-11
 Identities = 11/36 (30%), Positives = 22/36 (60%)
 Frame = +3


Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           N+G+G   A +L    A +I+A R+  KG+  ++++
Sbjct: 23  NSGLGFGLARRLSAAGADVIMAIRNRAKGEAAVEEI 58



>gi|16740649|gb|AAH16204.1|AAH16204  (BC016204) Similar to RIKEN cDNA A930033N07 gene [Mus
musculus]
          Length = 304

 Score = 44.3 bits (103), Expect(2) = 1e-10
 Identities = 30/81 (37%), Positives = 41/81 (50%)
 Frame = +1


Query: 118 NQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQ 297
           ++++   K ++  + +LDLS   SI+   E F              AGVM     KTT
Sbjct: 81  SEIRADTKNSQVLVRKLDLSDTKSIRAFAERFL-------------AGVMMCPYSKTT-D 126

Query: 298 SYELQFGTNHLGHFLLTELLV 360
            +E  FG NHLGHFLLT LL+
Sbjct: 127 GFETHFGVNHLGHFLLTYLLL 147
```

```
 Score = 41.2 bits (95), Expect(2) = 1e-10
 Identities = 17/36 (47%), Positives = 23/36 (63%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           NTGIGKE A +L    A + IACRD +KG+    ++
Sbjct: 48  NTGIGKETARELARRGARVYIACRDVLKGESAASEI 83



>gi|18416145|ref|NP_567681.1|  (NM_118472) putative protein [Arabidopsis thaliana]
          Length = 320

 Score = 65.5 bits (158), Expect = 1e-10
 Identities = 33/71 (46%), Positives = 51/71 (71%)
 Frame = +1

Query: 148 KAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNH 327
           K  +++LDLS   S+++   ++K   +P +++LINNAG+MA   +  +K + ELQF TNH
Sbjct: 81  KLDVMELDLSSMQSVRKFASEYKSTGLP-LNLLINNAGIMACP-FMLSKDNIELQFATNH 138

Query: 328 LGHFLLTELLV 360
           LGHFLLT+LL+
Sbjct: 139 LGHFLLTKLLL 149



>gi|9663977|dbj|BAB03618.1|  (AP002522) contains ESTs
AU064282(E20459),AU097686(S6509)~unknown protein [Oryza sativa (japonica cultivar-group)]
 gi|13872934|dbj|BAB44039.1|  (AP003045) contains ESTs
AU064282(E20459),AU097686(S6509)~unknown protein [Oryza sativa (japonica cultivar-group)]
          Length = 321

 Score = 60.8 bits (146), Expect(2) = 1e-10
 Identities = 31/68 (45%), Positives = 46/68 (67%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGH 336
           +++LDLS   S++   + F   + +P ++ILINNAGVM    +  ++   E+QF TNHLGH
Sbjct: 88  VLKLDLSSLKSVRAFADQFNSMNLP-LNILINNAGVMFCP-FGLSEDGVEMQFATNHLGH 145

Query: 337 FLLTELLV 360
           FLLT LL+
Sbjct: 146 FLLTNLLL 153



 Score = 24.3 bits (51), Expect(2) = 1e-10
 Identities = 16/56 (28%), Positives = 28/56 (49%)
 Frame = +3

Query: 6   TGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQLSQRIEQNNQSLLDIIGL 173
           +GIG E +       A +IIA R+T     V  ++       IE+N ++ +D++ L
Sbjct: 43  SGIGLETSRVFAMRGAHVIIAARNTEAASVVRKKI-------IEENPKAHIDVLKL 91
```

```
>gi|15224306|ref|NP_181290.1|  (NM_129309) putative oxidoreductase [Arabidopsis thaliana]
 gi|7485438|pir||T02520  hypothetical protein F13M22.4 - Arabidopsis thaliana
 gi|3236237|gb|AAC23625.1|  (AC004684) putative oxidoreductase [Arabidopsis thaliana]
 gi|20466185|gb|AAM20410.1|  (AY099558) putative oxidoreductase [Arabidopsis thaliana]
          Length = 321

 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 36/90 (40%), Positives = 52/90 (57%)
 Frame = +1

Query: 91  NKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMA 270
           N    N+    +   N   +   +Q+D+S   S++  V+ F  L +P ++ILINNAGVM
Sbjct: 66  NPKAANESKEMILQMNPNARVDYLQIDVSSIKSVRSFVDQFLALNVP-LNILINNAGVMF 124

Query: 271 PQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
              +K T+   E QF TNH+GHFLLT LL+
Sbjct: 125 CP-FKLTEDGIESQFATNHIGHFLLTNLLL 153



>gi|12848558|dbj|BAB27997.1|  (AK012056) cell line MC/9.IL4 derived transcript 1~data
source:MGD, source key:MGI:102581, evidence:ISS~putative [Mus musculus]
          Length = 316

 Score = 53.9 bits (128), Expect(2) = 2e-10
 Identities = 30/72 (41%), Positives = 44/72 (60%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           ++ ++ +LDL+   SI+   +DF   +   + +LINNAGVM    Y  T   +E+  G N
Sbjct: 89  SQVFVRKLDLADTKSIRAFAKDFLAEE-KHLHLLINNAGVMMCP-YSNTADGFEMHIGVN 146

Query: 325 HLGHFLLTELLV 360
           HLGHFLLT LL+
Sbjct: 147 HLGHFLLTHLLL 158



 Score = 30.8 bits (68), Expect(2) = 2e-10
 Identities = 15/29 (51%), Positives = 17/29 (57%)
 Frame = +3

Query: 6   TGIGKEAALQLGNMNATIIIACRDTIKGQ 92
           TGIGKE A  L    A + IA RD  KG+
Sbjct: 48  TGIGKETAKDLAQRGARVYIAFRDVDKGE 76


>gi|15607580|ref|NP_214953.1|  (NC_000962) hypothetical protein Rv0439c [Mycobacterium
tuberculosis H37Rv]
 gi|7476356|pir||H70829  hypothetical protein Rv0439c - Mycobacterium tuberculosis  (strain
H37RV)
 gi|2909514|emb|CAA17396.1|  (AL021932) hypothetical protein Rv0439c [Mycobacterium
tuberculosis H37Rv]
          Length = 311
```

```
 Score = 63.9 bits (154), Expect = 3e-10
 Identities = 36/69 (52%), Positives = 47/69 (67%), Gaps = 1/69 (1%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLG 333
           L QLDL   +S++    D  +   P+ID+LINNAGVM  P+  + TK  +ELQFGTNHLG
Sbjct: 76  LQQLDLCSLDSVRAAA-DALRTAYPRIDVLINNAGVMWTPK--QVTKDGFELQFGTNHLG 132

Query: 334 HFLLTELLV 360
           HF LT L++
Sbjct: 133 HFALTGLVL 141



>gi|15839827|ref|NP_334864.1|  (NC_002755) oxidoreductase, short-chain
dehydrogenase/reductase family [Mycobacterium tuberculosis CDC1551]
 gi|13879959|gb|AAK44678.1|  (AE006948) oxidoreductase, short-chain dehydrogenase/reductase
family [Mycobacterium tuberculosis CDC1551]
          Length = 338

 Score = 63.9 bits (154), Expect = 3e-10
 Identities = 36/69 (52%), Positives = 47/69 (67%), Gaps = 1/69 (1%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVM-APQTYKTTKQSYELQFGTNHLG 333
           L QLDL   +S++    D  +   P+ID+LINNAGVM  P+  + TK  +ELQFGTNHLG
Sbjct: 103 LQQLDLCSLDSVRAAA-DALRTAYPRIDVLINNAGVMWTPK--QVTKDGFELQFGTNHLG 159

Query: 334 HFLLTELLV 360
           HF LT L++
Sbjct: 160 HFALTGLVL 168



>gi|19075284|ref|NP_587784.1|  (NC_003421) hypothetical short chain dehydrogenase.
[Schizosaccharomyces pombe]
 gi|7491977|pir||T41570  hypothetical protein SPCC736.13 - fission yeast
(Schizosaccharomyces pombe)
 gi|3169083|emb|CAA19277.1|  (AL023705) hypothetical short chain dehydrogenase.
[Schizosaccharomyces pombe]
          Length = 339

 Score = 63.2 bits (152), Expect = 5e-10
 Identities = 35/90 (38%), Positives = 53/90 (58%)
 Frame = +1

Query: 91  NKSWINQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMA 270
           N+    +   Q+  + + +K   ++LDL  F S+ Q  E F   K  ++ IL+NNAG+M
Sbjct: 75  NEEKYQKVMKQIHDEVRHSKIRFLRLDLLDFESVYQAAESFIA-KEEKLHILVNNAGIMN 133

Query: 271 PQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
           P  ++ TK  YELQ  TN+L H+L TELL+
Sbjct: 134 PP-FELTKDGYELQIQTNYLSHYLFTELLL 162



>gi|15233359|ref|NP_192880.1|  (NM_117212) putative protein [Arabidopsis thaliana]
 gi|7486117|pir||T10561  hypothetical protein F25E4.30 - Arabidopsis thaliana
```

```
 gi|7267840|emb|CAB81242.1|  (AL161531) putative protein [Arabidopsis thaliana]
 gi|7321038|emb|CAB82146.1|  (AL050399) putative protein [Arabidopsis thaliana]
          Length = 317

 Score = 63.2 bits (152), Expect = 5e-10
 Identities = 34/85 (40%), Positives = 55/85 (64%)
 Frame = +1

Query: 106 NQYFNQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYK 285
           NQ  +++  +    K  +++LDLS   S++    +++ L +P +++LINNAG+MA   +
Sbjct: 67  NQVRDKILKEIPQAKIDVMKLDLSSMASVRSFASEYQSLDLP-LNLLINNAGIMACP-FL 124

Query: 286 TTKQSYELQFGTNHLGHFLLTELLV 360
            +  + ELQF TNHLGHFLLT LL+
Sbjct: 125 LSSDNIELQFATNHLGHFLLTNLLL 149




>gi|13876535|gb|AAK43511.1|AC020666_21  (AC020666) putative WW-domain oxidoreductase [Oryza
sativa (japonica cultivar-group)]
          Length = 320

 Score = 62.8 bits (151), Expect = 7e-10
 Identities = 30/69 (43%), Positives = 49/69 (70%)
 Frame = +1

Query: 154 YLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLG 333
           +++++DLS  +S+++   +F+ L +P ++ILINNAG+M     ++     ELQF TNH+G
Sbjct: 83  HVLEMDLSSMDSVRRFATEFEALNLP-LNILINNAGIMTRNCTRSI-DGLELQFATNHIG 140

Query: 334 HFLLTELLV 360
           HFLLT LL+
Sbjct: 141 HFLLTNLLL 149



>gi|81946|pir||S20941  protochlorophyllide reductase (EC 1.3.1.33) precursor - garden pea
          Length = 400

 Score = 62.0 bits (149), Expect = 1e-09
 Identities = 30/70 (42%), Positives = 49/70 (69%), Gaps = 2/70 (2%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKT--TKQSYELQFGTNHL 330
           ++ LDL+  +S++Q V++F++ ++P +D+LINNA V  P   +   T   +E+  GTNHL
Sbjct: 141 IMHLDLASLDSVRQFVDNFRRSEMP-LDVLINNAAVYFPTAKEPSFTADGFEISVGTNHL 199

Query: 331 GHFLLTELLV 360
           GHFLL+ LL+
Sbjct: 200 GHFLLSRLLL 209




>gi|266742|sp|Q01289|POR_PEA  Protochlorophyllide reductase, chloroplast precursor (PCR)
(NADPH-protochlorophyllide oxidoreductase) (POR)
 gi|20830|emb|CAA44786.1|  (X63060) protochlorophyllide reductase [Pisum sativum]
          Length = 399

 Score = 62.0 bits (149), Expect = 1e-09
```

```
 Identities = 30/70 (42%), Positives = 49/70 (69%), Gaps = 2/70 (2%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKT--TKQSYELQFGTNHL 330
           ++ LDL+  +S++Q V++F++ ++P +D+LINNA V  P   +   T   +E+  GTNHL
Sbjct: 140 IMHLDLASLDSVRQFVDNFRRSEMP-LDVLINNAAVYFPTAKEPSFTADGFEISVGTNHL 198

Query: 331 GHFLLTELLV 360
           GHFLL+ LL+
Sbjct: 199 GHFLLSRLLL 208




>gi|17558006|ref|NP_506570.1|   (NM_074169) dehydrogenase [Caenorhabditis elegans]
 gi|7496064|pir||T19314  hypothetical protein C15H11.4 - Caenorhabditis elegans
 gi|3874345|emb|CAB02732.1|  (Z81035) predicted using Genefinder~Similarity to
dehydrogenases, contains similarity to Pfam domain: PF00106 (short chain dehydrogenase),
Score=137.8, E-value=6.4e-38, N=1~cDNA EST yk68e1.3 comes from this gene~cDNA EST yk115c8.5
comes from this gene~>
          Length = 333

 Score = 57.4 bits (137), Expect(2) = 1e-09
 Identities = 29/72 (40%), Positives = 46/72 (63%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           T+ + I+ DL+ F S+++  ++  +     IDILINNAG+M    ++ TK  +E  + +N
Sbjct: 98  TRLHFIECDLTDFESVRRAAKETLE-STDTIDILINNAGIMFQSKHEQTKDGHEKTWQSN 156

Query: 325 HLGHFLLTELLV 360
           HLG FLLTEL +
Sbjct: 157 HLGPFLLTELFL 168




 Score = 24.3 bits (51), Expect(2) = 1e-09
 Identities = 13/39 (33%), Positives = 17/39 (43%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQ 119
           N GIG E   +L    A + + CR   K  +    LV Q
Sbjct: 55  NCGIGFETVRELNLRKADVYMLCRSEEKANEAKRMLVRQ 93




>gi|17826716|emb|CAC82539.1|  (AJ296079) SCAD family protein [Mus musculus]
          Length = 280

 Score = 52.0 bits (123), Expect(2) = 2e-09
 Identities = 28/81 (34%), Positives = 47/81 (57%)
 Frame = +1

Query: 118 NQVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQ 297
           + ++ +    +A+ + LDL+   S++    DF+ L +P + +L+NNA VM  +    T+
Sbjct: 85  SSIRAEMGSDRAHFLPLDLASLASVRGFARDFQALGLP-LHLLVNNAAVML-EPRAETED 142
```

```
Query: 298 SYELQFGTNHLGHFLLTELLV 360
            +E   G N LGHFLLT LL+
Sbjct: 143 GFERHLGVNFLGHFLLTLLLL 163




 Score = 28.9 bits (63), Expect(2) = 2e-09
 Identities = 11/34 (32%), Positives = 20/34 (58%)
 Frame = +3

Query: 9   GIGKEAALQLGNMNATIIIACRDTIKGQQVLDQL 110
           GIG+  A QL  +   +++A  D  +GQ+V+  +
Sbjct: 54  GIGRSTARQLARLGMCVVVAGNDEHRGQEVVSSI 87



>gi|13876532|gb|AAK43508.1|AC020666_18  (AC020666) putative WW-domain oxidoreductase [Oryza
sativa (japonica cultivar-group)]
          Length = 319

 Score = 60.8 bits (146), Expect = 2e-09
 Identities = 28/69 (40%), Positives = 48/69 (68%)
 Frame = +1

Query: 154 YLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLG 333
           +++++DLS  +S+++   +F  L +P ++ILINNAG+++    ++     EL F TNH+G
Sbjct: 85  HVLEMDLSSMDSVRRFASEFDSLNLP-LNILINNAGILSKDCIRSI-DGLELHFATNHIG 142

Query: 334 HFLLTELLV 360
           HFLLT LL+
Sbjct: 143 HFLLTNLLL 151



>gi|15217684|ref|NP_176640.1|  (NM_105134) oxidoreductase, putative [Arabidopsis thaliana]
 gi|6633817|gb|AAF19676.1|AC009519_10  (AC009519) F1N19.16 [Arabidopsis thaliana]
 gi|15081624|gb|AAK82467.1|  (AY048204) At1g64590/F1N19_15 [Arabidopsis thaliana]
          Length = 334

 Score = 60.5 bits (145), Expect = 3e-09
 Identities = 31/68 (45%), Positives = 49/68 (71%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGH 336
           ++ LDLS   S++ V+DF+ L +P ++ILINNAG  A + +  ++   E+ F TN+LGH
Sbjct: 89  VMHLDLSSLTSVRRFVDDFESLNLP-LNILINNAGKYAHK-HALSEDGVEMTFATNYLGH 146

Query: 337 FLLTELLV 360
           FLLT+LL+
Sbjct: 147 FLLTKLLL 154



>gi|15233666|ref|NP_194136.1|  (NM_118537) putative protein [Arabidopsis thaliana]
 gi|7487174|pir||T13447  hypothetical protein T19F6.40 - Arabidopsis thaliana
 gi|2262111|gb|AAB63619.1|  (AC002343) ribitol dehydrogenase isolog [Arabidopsis thaliana]
```

```
 gi|5668633|emb|CAB51648.1|  (AL109619) putative protein [Arabidopsis thaliana]
 gi|7269254|emb|CAB81323.1|  (AL161560) putative protein [Arabidopsis thaliana]
 gi|16612264|gb|AAL27501.1|AF439829_1  (AF439829) AT4g24050/T19F6_40 [Arabidopsis thaliana]
          Length = 332

 Score = 60.1 bits (144), Expect = 4e-09
 Identities = 31/72 (43%), Positives = 50/72 (69%)
 Frame = +1

Query: 145 TKAYLIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTN 324
           T+  +++LDLS   S++  V DF+ L +P +++LINNAG +A + +  ++   E+ F TN
Sbjct: 85  TEIVVMKLDLSSIASVRNFVADFESLDLP-LNLLINNAGKLAHE-HAISEDGIEMTFATN 142

Query: 325 HLGHFLLTELLV 360
           +LGHFLLT LL+
Sbjct: 143 YLGHFLLTNLLL 154



>gi|15218860|ref|NP_171860.1|  (NM_100243) putative protochlorophyllide reductase
[Arabidopsis thaliana]
 gi|10720234|sp|O48741|PORC_ARATH  Protochlorophyllide reductase C, chloroplast precursor
(PCR C) (NADPH-protochlorophyllide oxidoreductase C) (POR C)
 gi|7488284|pir||T00897  protochlorophyllide reductase (EC 1.3.1.33) precursor F21B7.11 -
Arabidopsis thaliana
 gi|8467964|dbj|BAA96654.1|  (AB035746) NADPH:protochlorophyllide oxidoreductase
[Arabidopsis thaliana]
 gi|9280649|gb|AAF86518.1|AC002560_11  (AC002560) F21B7.24 [Arabidopsis thaliana]
 gi|15081741|gb|AAK82525.1|  (AY048263) At1g03630/F21B7_11 [Arabidopsis thaliana]
          Length = 401

 Score = 59.7 bits (143), Expect = 6e-09
 Identities = 30/70 (42%), Positives = 47/70 (66%), Gaps = 2/70 (2%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKT--TKQSYELQFGTNHL 330
           ++ LDL+   S+KQ VE+F++ + P +D+L+ NA V  P   +   T + +E+  GTNHL
Sbjct: 143 VMHLDLASLESVKQFVENFRRTEQP-LDVLVCNAAVYQPTAKEPSFTAEGFEISVGTNHL 201

Query: 331 GHFLLTELLV 360
           GHFLL+ LL+
Sbjct: 202 GHFLLSRLLL 211



>gi|13676404|dbj|BAB41191.1|  (AB050124) NADPH-protochlorophyllide oxidoreductase 2
[Amaranthus tricolor]
          Length = 224

 Score = 58.9 bits (141), Expect = 9e-09
 Identities = 30/70 (42%), Positives = 47/70 (66%), Gaps = 2/70 (2%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTT--KQSYELQFGTNHL 330
           ++ LDL+   S++Q VE+F++  +P +D+L+ NA V  P   + T   + +EL  GTNHL
Sbjct: 30  VMHLDLASLESVRQFVENFRRSGMP-LDVLVCNAAVYLPTDKEPTFTAEGFELSVGTNHL 88

Query: 331 GHFLLTELLV 360
           GHFLL+ LL+
Sbjct: 89  GHFLLSRLLL 98
```

```
>gi|11250521|pir||T48275  hypothetical protein T22P11.130 - Arabidopsis thaliana
 gi|7413643|emb|CAB85991.1|  (AL162971) putative protein [Arabidopsis thaliana]
          Length = 350

 Score = 58.9 bits (141), Expect = 9e-09
 Identities = 38/98 (38%), Positives = 49/98 (49%), Gaps = 18/98 (18%)
 Frame = +1

Query: 121 QVKGQNKITKAYLIQLDLSCFNSIKQCVEDFKKLKIP-----------------QIDIL 246
           ++  QN   +  L+QLDLS   SIK  V +F  L +P                     L
Sbjct: 76  EILRQNANARVTLLQLDLSSIKSIKAFVREFHALHLPLNLLMYTFSLSLIQFKAFAPPFL 135

Query: 247 INNAGVMAPQTYKTTKQSYELQFGTNHLGHFLLTELLV 360
            NNAGVM    Y+ ++   ELQF TNH+GHFLLT LL+
Sbjct: 136 ANNAGVMFCP-YQLSEDGIELQFATNHIGHFLLTNLLL 172



>gi|15789712|ref|NP_279536.1|  (NC_002607) probable oxidoreductase; YajO1 [Halobacterium sp.
NRC-1]
 gi|10580084|gb|AAG19016.1|  (AE005001) probable oxidoreductase; YajO1 [Halobacterium sp.
NRC-1]
          Length = 316

 Score = 44.7 bits (104), Expect(2) = 1e-08
 Identities = 25/62 (40%), Positives = 35/62 (56%)
 Frame = +1

Query: 163 QLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTTKQSYELQFGTNHLGHFL 342
           +LDL+  +S+     D+   +   + +L NNAGVMA    +T    +E QFG NHLGH
Sbjct: 74  ELDLAALDSVA-AFADWFTAEFDSLHVLANNAGVMAIPRSETA-DGFETQFGVNHLGHVA 131

Query: 343 LT 348
           LT
Sbjct: 132 LT 133



 Score = 33.9 bits (76), Expect(2) = 1e-08
 Identities = 14/40 (35%), Positives = 23/40 (57%)
 Frame = +3

Query: 3   NTGIGKEAALQLGNMNATIIIACRDTIKGQQVLDQLVFQL 122
           N+G+G EA        A +++ACR T +G+   D +V +L
Sbjct: 26  NSGLGFEATRAFARAGAHVVMACRSTERGEDARDDIVAEL 65



>gi|15239574|ref|NP_200230.1|  (NM_124799) NADPH:protochlorophyllide oxidoreductase A
            (gb|AAC49043.1) [Arabidopsis thaliana]
 gi|10178168|dbj|BAB11581.1|  (AB013387) NADPH:protochlorophyllide oxidoreductase A
[Arabidopsis thaliana]
          Length = 405
```

```
 Score = 58.5 bits (140), Expect = 1e-08
 Identities = 28/70 (40%), Positives = 48/70 (68%), Gaps = 2/70 (2%)
 Frame = +1

Query: 157 LIQLDLSCFNSIKQCVEDFKKLKIPQIDILINNAGVMAPQTYKTT--KQSYELQFGTNHL 330
           ++ LDL+  +S++Q V++F++ ++P +D+L+ NA V  P   + T   + +EL  G NHL
Sbjct: 146 VMHLDLASLDSVRQFVDNFRRAEMP-LDVLVCNAAVYQPTANQPTFTAEGFELSVGINHL 204

Query: 331 GHFLLTELLV 360
           GHFLL+ LL+
Sbjct: 205 GHFLLSRLLI 214




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 198,613,104
Number of Sequences: 919285
Number of extensions: 3670974
Number of successful extensions: 13813
Number of sequences better than 10.0: 506
Number of HSP's better than 10.0 without gapping: 13376
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 13717
length of database: 288,558,979
effective HSP length: 104
effective length of database: 192,953,339
effective search space used: 4630880136
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

RID: 1020953202-02653-30479

7.1.7 Query= **hy-7_11b235002**

(275 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
919,285 sequences; 288,558,979 total letters

If you have any problems or questions with the results of this search
please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 150 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|4028547|gb|AAC97114.1|  (AF093689) MEK kinase alpha [Dict...    57   6e-08
gi|9857521|gb|AAG00876.1|AC064840_7  (AC064840) NPK1-related...    57   6e-08
gi|2342425|dbj|BAA21856.1|  (AB000798) NPK1-related protein ...    57   6e-08
gi|456309|dbj|BAA05648.1|  (D26601) protein kinase [Nicotian...    55   1e-07
gi|477094|pir||A48084  STE11 protein kinase homolog NPK1 - c...    55   1e-07
gi|15230612|ref|NP_187254.1|  (NM_111477) NPK1-related prote...    54   5e-07
gi|2342421|dbj|BAA21854.1|  (AB000796) NPK1-related protein ...    54   5e-07
gi|18390967|ref|NP_563832.1|  (NM_100771) putative NPK1-rela...    54   5e-07
gi|15240456|ref|NP_200320.1|  (NM_124891) putative protein [...    53   6e-07
gi|15236903|ref|NP_194419.1|  (NM_118823) putative NPK1-rela...    52   1e-06
gi|15225692|ref|NP_180810.1|  (NM_128810) putative protein k...    52   1e-06
gi|19423926|gb|AAL87297.1|  (AY080818) unknown protein [Arab...    52   1e-06
gi|5803088|ref|NP_005913.1|  (NM_005922) MAP/ERK kinase kina...    51   3e-06
gi|7959712|gb|AAF71029.1|AF116721_4  (AF116604) PRO0412 [Hom...    51   3e-06
gi|1504010|dbj|BAA13204.1|  (D86968) Similar to Mouse TFIIi-...    51   3e-06
gi|6031180|ref|NP_006715.1|  (NM_006724) MAP/ERK kinase kina...    51   3e-06
gi|4505599|ref|NP_002568.1|  (NM_002577) p21 (CDKN1A)-activa...    50   5e-06
gi|16758002|ref|NP_445758.1|  (NM_053306) p21 (CDKN1A)-activ...    50   5e-06
gi|2499647|sp|Q29502|PAK2_RABIT  Serine/threonine-protein ki...    50   5e-06
gi|3041712|sp|Q13177|PAK2_HUMAN  Serine/threonine-protein ki...    50   5e-06
gi|6288680|gb|AAF06695.1|U19967_1  (U19967) PAK2 [Rattus nor...    50   5e-06
gi|5138914|gb|AAD40374.1|  (AF092132) PAK2 [Homo sapiens]          50   5e-06
gi|984305|gb|AAA75468.1|  (U25975) hPAK65 [Homo sapiens]           50   5e-06
gi|15220416|ref|NP_172003.1|  (NM_100389) putative NPK1-rela...    50   7e-06
```

```
gi|5824350|emb|CAB54520.1|  (AJ238845) MAP3K epsilon 1 prote...   50   7e-06
gi|15240194|ref|NP_201509.1|  (NM_126108) protein kinase-lik...   49   9e-06
gi|16930437|gb|AAL31904.1|AF419572_1  (AF419572) At1g53570/F...   49   1e-05
gi|4204912|gb|AAD10848.1|  (U58918) MEK kinase [Arabidopsis ...   49   1e-05
gi|18404528|ref|NP_564635.1|  (NM_104235) MAP3K alpha protei...   49   1e-05
gi|8671870|gb|AAF78433.1|AC018748_12  (AC018748) Identical t...   49   1e-05
gi|6531993|gb|AAF15541.1|  (AF011756) septation [Aspergillus...   49   1e-05
gi|11358501|pir||T51625  MAP3K alpha protein kinase (EC 2.7....   49   1e-05
gi|15408764|dbj|BAB64164.1|  (AP003254) NPK1-related protein...   49   1e-05
gi|15231270|ref|NP_187962.1|  (NM_112199) MAP3K epsilon prot...   49   1e-05
gi|3549652|emb|CAA12272.1|  (AJ224982) MAP3K epsilon protein...   49   1e-05
gi|15231910|ref|NP_187455.1|  (NM_111677) putative MAP3K eps...   49   1e-05
gi|9588137|emb|CAC00588.1|  (AL138761) bA16H23.1.1 (protein ...   48   2e-05
gi|7661994|ref|NP_055535.1|  (NM_014720) Ste20-related serin...   48   2e-05
gi|1256422|gb|AAC50590.1|  (U51120) PAK1 [Homo sapiens]           48   2e-05
gi|9588136|emb|CAC00587.1|  (AL138761) bA16H23.1.2 (protein ...   48   2e-05
gi|543451|pir||S40482  serine/threonine-specific protein kin...   48   2e-05
gi|8393901|ref|NP_058894.1|  (NM_017198) p21 (CDKN1A)-activa...   48   2e-05
gi|7514128|pir||T18532  serine/threoine protein kinase - gui...   48   2e-05
gi|1523800|emb|CAA68958.1|  (Y07694) MAP kinase kinase alpha...   48   2e-05
gi|15232491|ref|NP_188759.1|  (NM_113017) MAP kinase kinase ...   48   2e-05
gi|11385654|gb|AAG34908.1|AF273048_1  (AF273048) CTCL tumor ...   48   2e-05
gi|1944185|dbj|BAA19655.1|  (AB002804) hSLK [Homo sapiens]        48   2e-05
gi|19698851|gb|AAL91161.1|  (AY081272) MAP kinase kinase 5 [...   48   2e-05
gi|4741823|gb|AAD28717.1|AF112855_1  (AF112855) Ste20-relate...   48   3e-05
gi|9507155|ref|NP_062222.1|  (NM_019349) Serine/threonine ki...   48   3e-05
gi|3688193|emb|CAA08995.1|  (AJ010091) MAP3K alpha 1 protein...   48   3e-05
gi|7106427|ref|NP_033315.1|  (NM_009289) serine/threonine ki...   48   3e-05
gi|7382496|ref|NP_002567.2|  (NM_002576) p21-activated kinas...   47   3e-05
gi|9256908|pdb|1F3M|C  Chain C, Crystal Structure Of Human S...   47   3e-05
gi|7677056|gb|AAF67008.1|AF155651_1  (AF155651) P21-activate...   47   3e-05
gi|4505601|ref|NP_002569.1|  (NM_002578) p21-activated kinas...   47   3e-05
gi|9506947|ref|NP_062083.1|  (NM_019210) p21 (CDKN1A)-activa...   47   3e-05
gi|6754630|ref|NP_036078.1|  (NM_011948) mitogen activated p...   47   4e-05
gi|15222311|ref|NP_172195.1|  (NM_100589) MAP3K gamma protei...   47   4e-05
gi|7300435|gb|AAF55592.1|  (AE003723) CG7717 gene product [D...   47   4e-05
gi|15222512|ref|NP_176557.1|  (NM_105047) putative protein k...   47   4e-05
gi|1168176|gb|AAB35358.1|  mPAK-3=p21Cdc42/Rac activated kin...   47   4e-05
gi|6679205|ref|NP_032804.1|  (NM_008778) p21 (CDKN1A)-activa...   47   4e-05
gi|15186736|dbj|BAB62892.1|  (AB069962) mekk1b [Drosophila m...   47   4e-05
gi|6685609|sp|O08648|M3K4_MOUSE  Mitogen-activated protein k...   47   4e-05
gi|15231352|ref|NP_190202.1|  (NM_114485) protein kinase -li...   47   4e-05
gi|15186734|dbj|BAB62891.1|  (AB069961) mekk1a [Drosophila m...   47   4e-05
gi|2654103|gb|AAC21676.1|  (AF034090) MAPKK kinase [Neurospo...   47   6e-05
gi|3420949|gb|AAC31969.1|  (AF082297) Cdc42/Rac effector kin...   47   6e-05
gi|6979343|gb|AAF34436.1|AF172282_25  (AF172282) similar to ...   46   1e-04
gi|20342260|ref|XP_110494.1|  (XM_110494) similar to Mitogen...   45   1e-04
gi|15227689|ref|NP_180565.1|  (NM_128559) putative protein k...   45   1e-04
gi|15528439|emb|CAC69137.1|  (AJ293274) MEK map kinase kinsa...   45   1e-04
gi|19113776|ref|NP_592864.1|  (NC_003424) serine/threonine-p...   45   1e-04
gi|6822245|emb|CAB70978.1|  (AJ242726) p21 activated kinase ...   45   1e-04
gi|10764163|gb|AAG22590.1|  (AF193343) GCN2beta [Mus musculus]    45   2e-04
gi|7305017|ref|NP_038747.1|  (NM_013719) GCN2 eIF2alpha kina...   45   2e-04
gi|19114476|ref|NP_593564.1|  (NC_003424) putative Ste20-lik...   45   2e-04
gi|10764165|gb|AAG22591.1|  (AF193344) GCN2gamma [Mus musculus]   45   2e-04
gi|15221948|ref|NP_175894.1|  (NM_104370) NPK1-related prote...   45   2e-04
gi|10764161|gb|AAG22589.1|  (AF193342) GCN2alpha [Mus musculus]   45   2e-04
gi|7243057|dbj|BAA92576.1|  (AB037759) KIAA1338 protein [Hom...   45   2e-04
gi|7504728|pir||T16511  hypothetical protein F59A6.1 - Caeno...   45   2e-04
gi|3688209|emb|CAA08997.1|  (AJ010093) MAP3K beta 1 protein ...   45   2e-04
gi|19110337|gb|AAL82716.1|  (AY074935) NPK1-related protein ...   45   2e-04
```

```
gi|1405366|gb|AAB03507.1|  (U60169) MkcA [Dictyostelium disc...   45   2e-04
gi|17535343|ref|NP_494925.1|  (NM_062524) Protein kinase [Ca...   45   2e-04
gi|1359632|emb|CAA66660.1|  (X98048) p21 activated kinase [E...   44   3e-04
gi|2131084|emb|CAA89388.1|  (Z49369) ORF YJL095w [Saccharomy...   44   4e-04
gi|6322366|ref|NP_012440.1|  (NC_001142) bypass requirement ...   44   4e-04
gi|3415|emb|CAA42788.1|  (X60227) protein kinase [Saccharomy...   44   4e-04
gi|218490|dbj|BAA01226.1|  (D10389) Ssp31 protein kinase [Sa...   44   4e-04
gi|3858883|gb|AAD09141.1|  (U67056) myosin I heavy chain kin...   44   5e-04
gi|7300149|gb|AAF55316.1|  (AE003713) CG14895 gene product [...   43   6e-04
gi|6754978|ref|NP_035165.1|  (NM_011035) p21 (CDKN1A)-activa...   43   6e-04
gi|11096132|gb|AAG30205.1|  (AF294841) Ste11alpha protein [F...   43   6e-04
gi|2660696|gb|AAB88118.1|  (AF000239) p21-activated kinase [...   43   8e-04
gi|4006878|emb|CAB16796.1|  (Z99707) MAP3K-like protein kina...   43   8e-04
gi|17979279|gb|AAL49865.1|  (AY070367) unknown protein [Arab...   43   8e-04
gi|6503190|gb|AAF14631.1|  (AF200688) tyrosine kinase ZAK1 [...   43   8e-04

                                Alignments


>gi|4028547|gb|AAC97114.1|  (AF093689) MEK kinase alpha [Dictyostelium discoideum]
          Length = 942

 Score = 56.6 bits (135), Expect = 6e-08
 Identities = 26/63 (41%), Positives = 37/63 (58%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PPW         +  I+S    P IP+++SQ   DFLN C + D K+R  A QLL+HPFI
Sbjct: 370 PPWSNITELAAVMYHIASSNSIPNIPSHMSQEAFDFLNLCFKRDPKERPDANQLLKHPFI 429

Query: 183 ISI 191
            +++
Sbjct: 430 MNL 432



>gi|9857521|gb|AAG00876.1|AC064840_7  (AC064840) NPK1-related protein kinase 2 [Arabidopsis
thaliana]
          Length = 651

 Score = 56.6 bits (135), Expect = 6e-08
 Identities = 29/78 (37%), Positives = 41/78 (52%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
            PW +    + A+  I + +  PPIP+NIS +  DFL KCLQ +   R  A +LL+HPF+
Sbjct: 272 PWSQQYKEIAAIFHIGTTKSHPPIPDNISSDANDFLLKCLQQEPNLRPTASELLKHPFVT 331

Query: 186 SIPKKLLKTDQLIFQRKS 239
               K+     D    F    S
Sbjct: 332 GKQKESASKDLTSFMDNS 349



>gi|2342425|dbj|BAA21856.1|  (AB000798) NPK1-related protein kinase 2 [Arabidopsis thaliana]
          Length = 642

 Score = 56.6 bits (135), Expect = 6e-08
 Identities = 29/78 (37%), Positives = 41/78 (52%)
```

```
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
           PW +    + A+  I + +  PPIP+NIS +  DFL KCLQ +   R  A +LL+HPF+
Sbjct: 263 PWSQQYKEIAAIFHIGTTKSHPPIPDNISSDANDFLLKCLQQEPNLRPTASELLKHPFVT 322

Query: 186 SIPKKLLKTDQLIFQRKS 239
              K+    D   F   S
Sbjct: 323 GKQKESASKDLTSFMDNS 340



>gi|456309|dbj|BAA05648.1|  (D26601) protein kinase [Nicotiana tabacum]
          Length = 690

 Score = 55.5 bits (132), Expect = 1e-07
 Identities = 26/60 (43%), Positives = 34/60 (56%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +    + AL  I + +  PPIP ++S    DFL KCLQ +   R  A  LLQHPF+
Sbjct: 283 PPWSQQYQEVAALFHIGTTKSHPPIPEHLSAESKDFLLKCLQKEPHLRHSASNLLQHPFV 342



>gi|477094|pir||A48084  STE11 protein kinase homolog NPK1 - common tobacco
          Length = 706

 Score = 55.5 bits (132), Expect = 1e-07
 Identities = 26/60 (43%), Positives = 34/60 (56%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +    + AL  I + +  PPIP ++S    DFL KCLQ +   R  A  LLQHPF+
Sbjct: 299 PPWSQQYQEVAALFHIGTTKSHPPIPEHLSAESKDFLLKCLQKEPHLRHSASNLLQHPFV 358



>gi|15230612|ref|NP_187254.1|  (NM_111477) NPK1-related protein kinase 3 [Arabidopsis
thaliana]
 gi|2342427|dbj|BAA21857.1|  (AB000799) NPK1-related protein kinase 3 [Arabidopsis thaliana]
 gi|7658341|gb|AAF66131.1|  (AC068073) NPK1-related protein kinase 3; 8286-4476 [Arabidopsis
thaliana]
 gi|17979012|gb|AAL47465.1|  (AY069917) AT3g06030/F24F17_1 [Arabidopsis thaliana]
          Length = 651

 Score = 53.5 bits (127), Expect = 5e-07
 Identities = 24/60 (40%), Positives = 32/60 (53%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW        A++ I   +  PPIP ++S    DFL KCL  +   R  A +LLQHPF+
Sbjct: 271 PPWSEQYQQFAAVLHIGRTKAHPPIPEDLSPEAKDFLMKCLHKEPSLRLSATELLQHPFV 330
```

```
>gi|2342421|dbj|BAA21854.1|  (AB000796) NPK1-related protein kinase 1L [Arabidopsis
thaliana]
          Length = 661

 Score = 53.5 bits (127), Expect = 5e-07
 Identities = 26/70 (37%), Positives = 39/70 (55%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
           PW +    + A+  I + +  PPIP+ +S +  DFL KCLQ     R  A +LL+HPF++
Sbjct: 268 PWSQQYKEVAAIFFIGTTKSHPPIPDTLSSDAKDFLLKCLQEVPNLRPTASELLKHPFVM 327

Query: 186 SIPKKLLKTD 215
              K+   TD
Sbjct: 328 GKHKESASTD 337




>gi|18390967|ref|NP_563832.1|  (NM_100771) putative NPK1-related protein kinase 2
[Arabidopsis thaliana]
 gi|2342423|dbj|BAA21855.1|  (AB000797) NPK1-related protein kinase 1S [Arabidopsis
thaliana]
          Length = 376

 Score = 53.5 bits (127), Expect = 5e-07
 Identities = 26/70 (37%), Positives = 39/70 (55%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
           PW +    + A+  I + +  PPIP+ +S +  DFL KCLQ     R  A +LL+HPF++
Sbjct: 273 PWSQQYKEVAAIFFIGTTKSHPPIPDTLSSDAKDFLLKCLQEVPNLRPTASELLKHPFVM 332

Query: 186 SIPKKLLKTD 215
              K+   TD
Sbjct: 333 GKHKESASTD 342




>gi|15240456|ref|NP_200320.1|  (NM_124891) putative protein [Arabidopsis thaliana]
 gi|9758106|dbj|BAB08578.1|  (AB010071) contains similarity to NPK1-related protein
kinase~gene_id:MCO15.4 [Arabidopsis thaliana]
          Length = 448

 Score = 53.1 bits (126), Expect = 6e-07
 Identities = 25/59 (42%), Positives = 34/59 (57%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PW      + A+  I     SP IP  +S+   DFL KCL+ D K+RW  ++LLQHPF+
Sbjct: 198 PWPELNDVVAAIYKIGFTGESPVIPVWLSEKGQDFLRKCLRKDPKQRWTVEELLQHPFL 256
```

```
>gi|15236903|ref|NP_194419.1|  (NM_118823) putative NPK1-related protein kinase [Arabidopsis
thaliana]
 gi|7487975|pir||T04812  NPK1-related protein kinase homolog F10M23.230 - Arabidopsis
thaliana
 gi|4455212|emb|CAB36535.1|  (AL035440) putative NPK1-related protein kinase [Arabidopsis
thaliana]
 gi|7269542|emb|CAB79544.1|  (AL161566) putative NPK1-related protein kinase [Arabidopsis
thaliana]
          Length = 444

 Score = 52.4 bits (124), Expect = 1e-06
 Identities = 24/59 (40%), Positives = 32/59 (53%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PW      + A+  I     SP IP  IS    DFL  CL+ D K+RW  ++LL+HPF+
Sbjct: 195 PWPELNDVVAAMYKIGFSGESPAIPAWISDKAKDFLKNCLKEDQKQRWTVEELLKHPFL 253




>gi|15225692|ref|NP_180810.1|  (NM_128810) putative protein kinase [Arabidopsis thaliana]
 gi|7487976|pir||T02550  NPK1-related protein kinase homolog T26B15.7 - Arabidopsis thaliana
 gi|3298539|gb|AAC25933.1|  (AC004681) putative protein kinase [Arabidopsis thaliana]
          Length = 372

 Score = 52.4 bits (124), Expect = 1e-06
 Identities = 24/67 (35%), Positives = 37/67 (54%), Gaps = 3/67 (4%)
 Frame = +3

Query: 3   PPWGRFQS---PMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQH 173
            PPW +  S   P+  L  +     +P +P  +++   DFL KCL+ +  +RW A QLL H
Sbjct: 197 PPWTKADSREDPVSVLYRVGYSSETPELPCLLAEEAKDFLEKCLKREANERWTATQLLNH 256

Query: 174 PFIISIP 194
            PF+ + P
Sbjct: 257 PFLTTKP 263




>gi|19423926|gb|AAL87297.1|  (AY080818) unknown protein [Arabidopsis thaliana]
          Length = 394

 Score = 52.4 bits (124), Expect = 1e-06
 Identities = 24/59 (40%), Positives = 32/59 (53%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PW      + A+  I     SP IP  IS    DFL  CL+ D K+RW  ++LL+HPF+
Sbjct: 195 PWPELNDVVAAMYKIGFSGESPAIPAWISDKAKDFLKNCLKEDQKQRWTVEELLKHPFL 253



>gi|5803088|ref|NP_005913.1|  (NM_005922) MAP/ERK kinase kinase 4, isoform a; SSK2/SSK22
MAP kinase kinase kinase, yeast, homolog of [Homo sapiens]
 gi|6685621|sp|Q9Y6R4|M3K4_HUMAN  Mitogen-activated protein kinase kinase kinase 4 (MAPK/ERK
kinase kinase 4) (MEK kinase 4) (MEKK 4) (MAP three kinase 1)
 gi|7513141|pir||T03022  MAP kinase kinase kinase - human
 gi|2352277|gb|AAB68804.1|  (AF002715) MAP kinase kinase kinase [Homo sapiens]
          Length = 1607
```

```
 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 23/59 (38%), Positives = 31/59 (51%)
 Frame = +3

Query: 6    PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PW  ++   Q +  +      PPIP  +S    DFL+ CL+ D K RW A QLL H F+
Sbjct: 1543 PWHEYEHNFQIMYKVGMGH-KPPIPERLSPEGKDFLSHCLESDPKMRWTASQLLDHSFV 1600



>gi|7959712|gb|AAF71029.1|AF116721_4  (AF116604) PRO0412 [Homo sapiens]
          Length = 544

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 23/59 (38%), Positives = 31/59 (51%)
 Frame = +3

Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PW  ++   Q +  +      PPIP  +S    DFL+ CL+ D K RW A QLL H F+
Sbjct: 480 PWHEYEHNFQIMYKVGMGH-KPPIPERLSPEGKDFLSHCLESDPKMRWTASQLLDHSFV 537



>gi|1504010|dbj|BAA13204.1|  (D86968) Similar to Mouse TFIIi-associated transactivator
factor p17(GB_RO:MMU11548): Containing protein kinase motif [Homo sapiens]
          Length = 1491

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 23/59 (38%), Positives = 31/59 (51%)
 Frame = +3

Query: 6    PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PW  ++   Q +  +      PPIP  +S    DFL+ CL+ D K RW A QLL H F+
Sbjct: 1427 PWHEYEHNFQIMYKVGMGH-KPPIPERLSPEGKDFLSHCLESDPKMRWTASQLLDHSFV 1484



>gi|6031180|ref|NP_006715.1|  (NM_006724) MAP/ERK kinase kinase 4, isoform b; SSK2/SSK22
MAP kinase kinase kinase, yeast, homolog of [Homo sapiens]
          Length = 1558

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 23/59 (38%), Positives = 31/59 (51%)
 Frame = +3

Query: 6    PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PW  ++   Q +  +      PPIP  +S    DFL+ CL+ D K RW A QLL H F+
Sbjct: 1494 PWHEYEHNFQIMYKVGMGH-KPPIPERLSPEGKDFLSHCLESDPKMRWTASQLLDHSFV 1551



>gi|4505599|ref|NP_002568.1|  (NM_002577) p21 (CDKN1A)-activated kinase 2; novel serine
kinase; hPAK65 [Homo sapiens]
 gi|2136032|pir||S58682  protein kinase, p21-activated (EC 2.7.1.-) - human
 gi|780808|gb|AAA65442.1|  (U24153) p21-activated protein kinase [Homo sapiens]
```

```
          Length = 525

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++    +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 447 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 504

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L      LI   K    + N
Sbjct: 505 KPLSSLTPLIMAAKEAMKSN 524




>gi|16758002|ref|NP_445758.1|  (NM_053306) p21 (CDKN1A)-activated kinase 2 [Rattus
norvegicus]
 gi|2499648|sp|Q64303|PAK2_RAT  Serine/threonine-protein kinase PAK 2 (p21-activated kinase
2) (PAK-2) (Gamma-PAK) (P62-PAK)
 gi|1016005|gb|AAA79064.1|  (U35345) serine/threonine kinase [Rattus norvegicus]
 gi|1245844|gb|AAB35608.1|  (S80221) gamma-PAK [Rattus sp.]
          Length = 524

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++    +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 446 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 503

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L      LI   K    + N
Sbjct: 504 KPLSSLTPLILAAKEAMKSN 523




>gi|2499647|sp|Q29502|PAK2_RABIT  Serine/threonine-protein kinase PAK 2 (p21-activated
kinase 2) (PAK-2) (Gamma-PAK) (p21-activated protein kinase I) (PAKI)
 gi|1236054|gb|AAC48537.1|  (U46915) p21-activated protein kinase I [Oryctolagus cuniculus]
 gi|1589148|prf||2210322A  cytostatic G protein-activated protein kinase PAK I [Oryctolagus
cuniculus]
          Length = 524

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++    +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 446 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 503

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L      LI   K    + N
Sbjct: 504 KPLSSLTPLIMAAKEAMKSN 523
```

```
>gi|3041712|sp|Q13177|PAK2_HUMAN  Serine/threonine-protein kinase PAK 2 (p21-activated
kinase 2) (PAK-2) (PAK65) (Gamma-PAK) (S6/H4 kinase)
          Length = 524

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 446 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 503

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L     LI   K   + N
Sbjct: 504 KPLSSLTPLIMAAKEAMKSN 523



>gi|6288680|gb|AAF06695.1|U19967_1    (U19967) PAK2 [Rattus norvegicus]
          Length = 524

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 446 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 503

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L     LI   K   + N
Sbjct: 504 KPLSSLTPLILAAKEAMKSN 523



>gi|5138914|gb|AAD40374.1|  (AF092132) PAK2 [Homo sapiens]
          Length = 540

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 462 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 519

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
           K L     LI   K   + N
Sbjct: 520 KPLSSLTPLIMAAKEAMKSN 539



>gi|984305|gb|AAA75468.1|    (U25975) hPAK65 [Homo sapiens]
          Length = 493
```

```
 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/80 (37%), Positives = 46/80 (57%), Gaps = 2/80 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
            ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQHPF + +
Sbjct: 415 ENPLRALYLIATNG-TPELQNPEKLSPIFRDFLNRCLEMDVEKRGSAKELLQHPF-LKLA 472

Query: 195 KKLLKTDQLIFQRKSTDQYN 254
            K L     LI   K   + N
Sbjct: 473 KPLSSLTPLIMAAKEAMKSN 492




>gi|15220416|ref|NP_172003.1|  (NM_100389) putative NPK1-related MAP kinase [Arabidopsis
thaliana]
 gi|4056416|gb|AAC97990.1|AAC97990  (AC005322) Strong similarity to Dsor1 protein kinase
gb|D13782 from Drosophila melanogaster. [Arabidopsis thaliana]
 gi|17381166|gb|AAL36395.1|  (AY064039) putative NPK1-related MAP kinase [Arabidopsis
thaliana]
 gi|20259425|gb|AAM14033.1|  (AY091011) putative NPK1-related MAP kinase [Arabidopsis
thaliana]
          Length = 339

 Score = 49.7 bits (117), Expect = 7e-06
 Identities = 23/69 (33%), Positives = 38/69 (54%), Gaps = 2/69 (2%)
 Frame = +3

Query: 6   PW--GRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPF 179
            PW    F  P+  L  +      P +P ++++   DFL KCL+ +  +RW A QLL HPF
Sbjct: 203 PWIGADFTDPVSVLYRVGYLGELPELPCSLTEQAKDFLGKCLKKEATERWTASQLLNHPF 262

Query: 180 IISIPKKLL 206
            +++   +L+
Sbjct: 263 LVNKEPELV 271




>gi|5824350|emb|CAB54520.1|  (AJ238845) MAP3K epsilon 1 protein kinase [Brassica napus]
          Length = 1299

 Score = 49.7 bits (117), Expect = 7e-06
 Identities = 27/60 (45%), Positives = 38/60 (63%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
            PP+   Q PM AL   I     SPPIP+++S ++ DFL +C + D ++R  AK LL HP+I
Sbjct: 217 PPYYDLQ-PMPALFRIVQDD-SPPIPDSLSPDITDFLRQCFKKDSRQRPDAKTLLSHPWI 274




>gi|15240194|ref|NP_201509.1|  (NM_126108) protein kinase-like protein [Arabidopsis
thaliana]
 gi|10177595|dbj|BAB10942.1|  (AB020742) protein kinase-like protein [Arabidopsis thaliana]
          Length = 344

 Score = 49.3 bits (116), Expect = 9e-06
 Identities = 27/74 (36%), Positives = 39/74 (52%), Gaps = 2/74 (2%)
```

```
 Frame = +3

Query: 30  MQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI-ISIPKK-L 203
           M  L+ I      P IP  +S+   DFL+KC   D KKRW A+ LL HPF+ + +    L
Sbjct: 219 MSLLLRIGVGDEVPMIPEELSEQGRDFLSKCFVKDPKKRWTAEMLLNHPFVTVDVDHDVL 278

Query: 204 LKTDQLIFQRKSTD 245
           +K +  +   K+ D
Sbjct: 279 VKEEDFVVNMKTED 292



>gi|16930437|gb|AAL31904.1|AF419572_1  (AF419572) At1g53570/F22G10_18 [Arabidopsis thaliana]
          Length = 609

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 22/60 (36%), Positives = 38/60 (62%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +F+  + A+  I + + +P IP+++S +  +F+  CLQ +   R  A QLL+HPF+
Sbjct: 412 PPWSQFEG-VAAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL 470



>gi|4204912|gb|AAD10848.1|  (U58918) MEK kinase [Arabidopsis thaliana]
          Length = 608

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 22/60 (36%), Positives = 38/60 (62%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +F+  + A+  I + + +P IP+++S +  +F+  CLQ +   R  A QLL+HPF+
Sbjct: 411 PPWSQFEG-VAAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL 469



>gi|18404528|ref|NP_564635.1|  (NM_104235) MAP3K alpha protein kinase, putative [Arabidopsis
thaliana]
          Length = 609

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 22/60 (36%), Positives = 38/60 (62%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +F+  + A+  I + + +P IP+++S +  +F+  CLQ +   R  A QLL+HPF+
Sbjct: 412 PPWSQFEG-VAAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL 470



>gi|8671870|gb|AAF78433.1|AC018748_12  (AC018748) Identical to MEK kinase from Arabidopsis
thaliana gb|U58918 and contains protein kinase PF|00069 domain. ESTs gb|Z33980, gb|T20498,
gb|AA650775 come from this gene
 gi|12324016|gb|AAG51965.1|AC024260_3  (AC024260) MEK kinase MAP3Ka, putative; 84794-81452
[Arabidopsis thaliana]
```

```
          Length = 608

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 22/60 (36%), Positives = 38/60 (62%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +F+   + A+  I + + +P IP+++S +  +F+  CLQ +   R  A QLL+HPF+
Sbjct: 412 PPWSQFEG-VAAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL 470



>gi|6531993|gb|AAF15541.1|  (AF011756) septation [Aspergillus nidulans]
          Length = 1320

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 31/91 (34%), Positives = 49/91 (53%), Gaps = 7/91 (7%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PP+   Q PM AL  I +    PP+P   S  + DFL +C Q D   R  A++LL+HP+I
Sbjct: 208 PPYYNLQ-PMPALFRIVNDD-HPPLPQGASPAVKDFLMQCFQKDPNLRVSARKLLKHPWI 265

Query: 183 IS-------IPKKLLKTDQLIFQRKSTDQYN 254
           ++       +PKK  + ++ +   KS  ++N
Sbjct: 266 VNARRSDSVVPKKSTEYEEAV---KSVQEWN 293



>gi|11358501|pir||T51625  MAP3K alpha protein kinase (EC 2.7.1.-) [imported] - Arabidopsis
thaliana
 gi|3688191|emb|CAA08994.1|  (AJ010090) MAP3K alpha protein kinase [Arabidopsis thaliana]
          Length = 582

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 22/60 (36%), Positives = 38/60 (62%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PPW +F+   + A+  I + + +P IP+++S +  +F+  CLQ +   R  A QLL+HPF+
Sbjct: 385 PPWSQFEG-VAAIFKIGNSKDTPEIPDHLSNDAKNFIRLCLQRNPTVRPTASQLLEHPFL 443



>gi|15408764|dbj|BAB64164.1|  (AP003254) NPK1-related protein kinase-like protein [Oryza
sativa (japonica cultivar-group)]
          Length = 387

 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 20/61 (32%), Positives = 32/61 (51%)
 Frame = +3


Query: 6   PWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
           PW   +  + A+  I      P +P  +S    DFL +C   + ++RW + QLL+HPF+
Sbjct: 200 PWSDMEDILSAVRRIGYTDAVPEVPEWLSAEAKDFLARCFARNPRERWTSSQLLEHPFLA 259
```

```
Query: 186 S 188
           S
Sbjct: 260 S 260




>gi|15231270|ref|NP_187962.1|  (NM_112199) MAP3K epsilon protein kinase [Arabidopsis
thaliana]
 gi|9280305|dbj|BAB01760.1|  (AP000603) MAP3K epsilon protein kinase [Arabidopsis thaliana]
          Length = 1368

 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 26/60 (43%), Positives = 38/60 (63%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PP+   Q PM AL  I     +PPIP+++S ++ DFL +C + D ++R  AK LL HP+I
Sbjct: 217 PPYYDLQ-PMPALFRIVQDD-NPPIPDSLSPDITDFLRQCFKKDSRQRPDAKTLLSHPWI 274




>gi|3549652|emb|CAA12272.1|  (AJ224982) MAP3K epsilon protein kinase [Arabidopsis thaliana]
          Length = 1368

 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 26/60 (43%), Positives = 38/60 (63%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PP+   Q PM AL  I     +PPIP+++S ++ DFL +C + D ++R  AK LL HP+I
Sbjct: 217 PPYYDLQ-PMPALFRIVQDD-NPPIPDSLSPDITDFLRQCFKKDSRQRPDAKTLLSHPWI 274




>gi|15231910|ref|NP_187455.1|  (NM_111677) putative MAP3K epsilon protein kinase
[Arabidopsis thaliana]
 gi|6648210|gb|AAF21208.1|AC013483_32  (AC013483) putative MAP3K epsilon protein kinase
[Arabidopsis thaliana]
          Length = 1367

 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 33/90 (36%), Positives = 48/90 (52%), Gaps = 3/90 (3%)
 Frame = +3

Query: 3   PPWGRFQSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           PP+   Q PM AL  I     +PPIP+++S ++ DFL  C + D ++R  AK LL HP+I
Sbjct: 217 PPYYDLQ-PMPALYRIVQDD-TPPIPDSLSPDITDFLRLCFKKDSRQRPDAKTLLSHPWI 274

Query: 183 ISIPKKL---LKTDQLIFQRKSTDQYNNLD 263
            +  + L   L+    I   K TD  +  D
Sbjct: 275 RNSRRALRSSLRHSGTIRYMKETDSSSEKD 304




>gi|9588137|emb|CAC00588.1|  (AL138761) bA16H23.1.1 (protein kinase KIAA0204 (HSLK) (isoform
1)) [Homo sapiens]
          Length = 1154
```

```
 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           +PM+ L+ I+   +  PP     P+  S N  DFL KCL+ +    RW    QLLQHPF+
Sbjct: 189 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVDS 246

Query: 195 KKLLKTDQLIFQRKS 239
            K ++   +LI + K+
Sbjct: 247 NKPIR--ELIAEAKA 259



>gi|7661994|ref|NP_055535.1|  (NM_014720) Ste20-related serine/threonine kinase; Ste20-
like kinase [Homo sapiens]
 gi|1503992|dbj|BAA13195.1|  (D86959) KIAA0204 protein [Homo sapiens]
          Length = 1152

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           +PM+ L+ I+   +  PP     P+  S N  DFL KCL+ +    RW    QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVDS 296

Query: 195 KKLLKTDQLIFQRKS 239
            K ++   +LI + K+
Sbjct: 297 NKPIR--ELIAEAKA 309



>gi|1256422|gb|AAC50590.1|  (U51120) PAK1 [Homo sapiens]
          Length = 545

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 30/81 (37%), Positives = 46/81 (56%), Gaps = 2/81 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +   P  +S    DFLN+CL+ D +KR  AK+LLQH F + I
Sbjct: 467 ENPLRALYLIATNG-TPELQNPEKLSAIFRDFLNRCLEMDVEKRGSAKELLQHQF-LKIA 524

Query: 195 KKLLKTDQLIFQRKSTDQYNN 257
           K L     LI    K    + N+
Sbjct: 525 KPLSSLTPLIAAAKEATKNNH 545



>gi|9588136|emb|CAC00587.1|  (AL138761) bA16H23.1.2 (protein kinase KIAA0204 (HSLK) (isoform
2)) [Homo sapiens]
          Length = 1185
```

```
 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           +PM+ L+ I+  +  PP    P+  S N  DFL KCL+ +   RW   QLLQHPF+
Sbjct: 189 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVDS 246

Query: 195 KKLLKTDQLIFQRKS 239
            K ++   +LI + K+
Sbjct: 247 NKPIR--ELIAEAKA 259



>gi|543451|pir||S40482  serine/threonine-specific protein kinase (EC 2.7.1.-) - rat
 gi|739718|prf||2003404A  Ser/Thr protein kinase [Rattus norvegicus]
          Length = 544

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 30/81 (37%), Positives = 46/81 (56%), Gaps = 2/81 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQH F + I
Sbjct: 466 ENPLRALYLIATNG-TPELQNPEKLSAIFRDFLNRCLEMDVEKRGSAKELLQHQF-LKIA 523

Query: 195 KKLLKTDQLIFQRKSTDQYNN 257
           K L     LI   K   + N+
Sbjct: 524 KPLSSLTPLIAAAKEATKNNH 544



>gi|8393901|ref|NP_058894.1|  (NM_017198) p21 (CDKN1A)-activated kinase 1 [Rattus
norvegicus]
 gi|3915789|sp|P35465|PAK1_RAT  Serine/threonine-protein kinase PAK 1 (p21-activated kinase
1) (PAK-1) (P68-PAK) (Alpha-PAK) (Protein kinase MUK2)
 gi|1399508|gb|AAB61533.1|  (U49953) protein kinase MUK2 [Rattus norvegicus]
 gi|2772514|gb|AAB95646.1|  (U23443) serine/threonine protein kinase [Rattus norvegicus]
          Length = 544

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 30/81 (37%), Positives = 46/81 (56%), Gaps = 2/81 (2%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPI--PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           ++P++AL  I++   +P +  P  +S    DFLN+CL+ D +KR  AK+LLQH F + I
Sbjct: 466 ENPLRALYLIATNG-TPELQNPEKLSAIFRDFLNRCLEMDVEKRGSAKELLQHQF-LKIA 523

Query: 195 KKLLKTDQLIFQRKSTDQYNN 257
           K L     LI   K   + N+
Sbjct: 524 KPLSSLTPLIAAAKEATKNNH 544



>gi|7514128|pir||T18532  serine/threoine protein kinase - guinea pig
 gi|2911114|dbj|BAA24930.1|  (D88425) serine/threoine kinase [Cavia porcellus]
          Length = 1231
```

```
 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
            +PM+ L+ I+  +  PP    P+  S N  DFL KCL+ +   RW   QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTIDS 296

Query: 195 KKLLKTDQLIFQRKS 239
             K ++   +LI + K+
Sbjct: 297 NKPIR--ELIAEAKA 309



>gi|1523800|emb|CAA68958.1|  (Y07694) MAP kinase kinase alpha protein kinase [Arabidopsis
thaliana]
          Length = 348

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 24/55 (43%), Positives = 29/55 (52%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
            Q    +LM        P  P   SQ    F++ CLQ D  KRW A+QLLQHPFI+
Sbjct: 272 QGDWASLMCAICMSQPPEAPATASQEFRHFVSCCLQSDPPKRWSAQQLLQHPFIL 326



>gi|15232491|ref|NP_188759.1|  (NM_113017) MAP kinase kinase 5 [Arabidopsis thaliana]
 gi|11259770|pir||T51340  mitogen-activated protein kinase kinase (EC 2.7.1.-) 5 [imported]
- Arabidopsis thaliana
 gi|3219273|dbj|BAA28831.1|  (AB015316) MAP kinase kinase 5 [Arabidopsis thaliana]
 gi|9280224|dbj|BAB01714.1|  (AB023045) MAP kinase kinase 5 [Arabidopsis thaliana]
          Length = 348

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 24/55 (43%), Positives = 29/55 (52%)
 Frame = +3


Query: 21  QSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
            Q    +LM        P  P   SQ    F++ CLQ D  KRW A+QLLQHPFI+
Sbjct: 272 QGDWASLMCAICMSQPPEAPATASQEFRHFVSCCLQSDPPKRWSAQQLLQHPFIL 326



>gi|11385654|gb|AAG34908.1|AF273048_1  (AF273048) CTCL tumor antigen se20-9 [Homo sapiens]
          Length = 1235

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
            +PM+ L+ I+  +  PP    P+  S N  DFL KCL+ +   RW   QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVDS 296
```

```
Query: 195 KKLLKTDQLIFQRKS 239
            K ++  +LI + K+
Sbjct: 297 NKPIR--ELIAEAKA 309



>gi|1944185|dbj|BAA19655.1| ☐ (AB002804) hSLK [Homo sapiens]
          Length = 1204

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 28/75 (37%), Positives = 42/75 (55%), Gaps = 3/75 (4%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFIISIP 194
           +PM+ L+ I+  +  PP    P+  S N  DFL KCL+ +   RW   QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVDS 296

Query: 195 KKLLKTDQLIFQRKS 239
            K ++  +LI + K+
Sbjct: 297 NKPIR--ELIAEAKA 309



>gi|19698851|gb|AAL91161.1|  (AY081272) MAP kinase kinase 5 [Arabidopsis thaliana]
          Length = 335

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 24/55 (43%), Positives = 29/55 (52%)
 Frame = +3

Query: 21  QSPMQALMTISSKQCSPPIPNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFII 185
           Q    +LM        P  P   SQ    F++ CLQ D  KRW A+QLLQHPFI+
Sbjct: 259 QGDWASLMCAICMSQPPEAPATASQEFRHFVSCCLQSDPPKRWSAQQLLQHPFIL 313



>gi|4741823|gb|AAD28717.1|AF112855_1 ☐ (AF112855) Ste20-related kinase SMAK [Mus musculus]
          Length = 1202

 Score = 47.8 bits (112), Expect = 3e-05
 Identities = 24/56 (42%), Positives = 33/56 (58%), Gaps = 3/56 (5%)
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           +PM+ L+ I+  +  PP    P+  S N  DFL KCL+ +   RW   QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSKWSSNFKDFLRKCLEKNVDARWTTSQLLQHPFV 292



>gi|9507155|ref|NP_062222.1| ☐ (NM_019349) Serine/threonine kinase 2 [Rattus norvegicus]
 gi|7514055|pir||T34021  protein kinase SK2 - rat
 gi|2114074|dbj|BAA20077.1| ☐ (AB003357) protein kinase [Rattus norvegicus]
          Length = 1206

 Score = 47.8 bits (112), Expect = 3e-05
 Identities = 24/56 (42%), Positives = 33/56 (58%), Gaps = 3/56 (5%)
```

```
 Frame = +3

Query: 24  SPMQALMTISSKQCSPPI---PNNISQNLMDFLNKCLQFDHKKRWKAKQLLQHPFI 182
           +PM+ L+ I+   +  PP     P+  S N  DFL KCL+ +    RW   QLLQHPF+
Sbjct: 239 NPMRVLLKIAKSE--PPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFV 292





  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 154,446,950
Number of Sequences: 919285
Number of extensions: 2884131
Number of successful extensions: 11425
Number of sequences better than 10.0: 1028
Number of HSP's better than 10.0 without gapping: 11266
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 11418
length of database: 288,558,979
effective HSP length: 67
effective length of database: 226,966,884
effective search space used: 5447205216
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

RID: 1020953576-09357-2027

7.1.8 Query= **hy-8_13a250005**

(536 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
919,285 sequences; 288,558,979 total letters

If you have any problems or questions with the results of this search
please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 22 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|7291278|gb|AAF46709.1|  (AE003453) CG9752 gene product [D...    52   3e-06
gi|19114965|ref|NP_594053.1|  (NC_003424) conserved hypothet...    46   2e-04
gi|11359514|pir||T51082  hypothetical protein B2A19.160 [imp...    37   0.13
gi|17538666|ref|NP_500924.1|  (NM_068523) C11D2.4.p [Caenorh...    36   0.17
gi|80469|pir||B29838  parasporal crystal protein - Bacillus ...    33   1.5
gi|15925464|ref|NP_372998.1|  (NC_002758) conserved hypothet...    32   2.5
gi|111664|pir||A40467  glucuronosyltransferase (EC 2.4.1.17)...    32   3.3
gi|13928718|ref|NP_113721.1|  (NM_031533) androsterone UDP-g...    32   3.3
gi|57453|emb|CAA27198.1|  (X03478) UDPGT (aa 1-500) [Rattus ...    32   3.3
gi|15896731|ref|NP_350080.1|  (NC_003030) Transcription regu...    32   3.3
gi|57449|emb|CAA68351.1|  (Y00156) precursor polypeptide (AA...    32   4.3
gi|19527110|ref|NP_598655.1|  (NM_133894) expressed sequence...    32   4.3
gi|6678501|ref|NP_033493.1|  (NM_009467) UDP-glucuronosyltra...    32   4.3
gi|20381430|gb|AAH28262.1|  (BC028262) UDP-glucuronosyltrans...    32   4.3
gi|92287|pir||S07390  glucuronosyltransferase (EC 2.4.1.17) ...    32   4.3
gi|7499184|pir||T20978  hypothetical protein F15D3.9 - Caeno...    31   5.6
gi|3892157|emb|CAA10033.1|  (AJ012469) DYS-1 protein [Caenor...    31   5.6
gi|17506447|ref|NP_492946.1|  (NM_060545) dystrophin [Caenor...    31   5.6
gi|7499176|pir||T20971  hypothetical protein F15D3.1 - Caeno...    31   5.6
gi|19703788|ref|NP_603350.1|  (NC_003454) Xaa-Pro aminopepti...    31   7.3
gi|15792817|ref|NP_282640.1|  (NC_002163) putative proline d...    31   7.3
gi|14578290|gb|AAF99456.1|  (AY003872) PV1H14050_P [Plasmodi...    30   9.5
```

Alignments

```
>gi|7291278|gb|AAF46709.1|  (AE003453) CG9752 gene product [Drosophila melanogaster]
          Length = 343

 Score = 52.0 bits (123), Expect = 3e-06
 Identities = 26/80 (32%), Positives = 45/80 (55%), Gaps = 7/80 (8%)
 Frame = +2

Query: 206 RXSCQFVSENSKHVKI---NKEELQKQVNEWFINNKHKYKEFDEYECH----LAADEEQV 364
           R S  ++ EN+KHV+I     E+L +++ +  I  +   + F ++E H       D  +
Sbjct: 9   RESGAYIVENAKHVRILPAGIEKLTQEIVKGLIEKRIAVENFSQHELHPNPNQPEDYSKA 68

Query: 365 VDFLFILDSLNFCFWPQKDY 424
            D++F+LD+LNFCFW   +Y
Sbjct: 69  ADWVFVLDTLNFCFWTPNNY 88



>gi|19114965|ref|NP_594053.1|  (NC_003424) conserved hypothetical protein; putative
glycoprotein [Schizosaccharomyces pombe]
 gi|12043549|emb|CAC19762.1|  (AL512496) conserved hypothetical protein; putative
glycoprotein [Schizosaccharomyces pombe]
          Length = 346

 Score = 46.2 bits (108), Expect = 2e-04
 Identities = 25/81 (30%), Positives = 41/81 (49%), Gaps = 3/81 (3%)
 Frame = +2

Query: 194 MSLVRXSCQFVSENSKHVKINK---EELQKQVNEWFINNKHKYKEFDEYECHLAADEEQV 364
           MS V    +F+S NS  VK+NK     +   + E   +   ++ E+  +E H    +
Sbjct: 1   MSRVLQDAEFISLNSNDVKVNKGGCAAVATWIKEKLDSLGPQFAEWQNHELHPKTRDVST 60

Query: 365 VDFLFILDSLNFCFWPQKDYE 427
           +D++F++D LNF FW   D E
Sbjct: 61  LDWIFLVDILNFSFWSDVDVE 81



>gi|11359514|pir||T51082  hypothetical protein B2A19.160 [imported] - Neurospora crassa
 gi|9368586|emb|CAB98246.1|  (AL390092) conserved hypothetical protein [Neurospora crassa]
          Length = 372

 Score = 36.6 bits (83), Expect = 0.13
 Identities = 22/78 (28%), Positives = 36/78 (45%), Gaps = 3/78 (3%)
 Frame = +2

Query: 203 VRXSCQFVSENSKHVKINKEELQKQVNE-WFINNKHKYKE--FDEYECHLAADEEQVVDF 373
           V  S +FV  ++  V ++    +K  N  W    + +Y    +  +  H     E+ V F
Sbjct: 36  VLESAEFVYNHAIDVALDMRGCKKAANTIWDQMQQREYSPATWSTHALHPKEKNEETVKF 95

Query: 374 LFILDSLNFCFWPQKDYE 427
           +F +D LNF FW +K  E
Sbjct: 96  IFTMDLLNFSFWSEKSEE 113
```

```
>gi|17538666|ref|NP_500924.1|  (NM_068523) C11D2.4.p [Caenorhabditis elegans]
 gi|7495851|pir||T32961  hypothetical protein C11D2.4 - Caenorhabditis elegans
 gi|2854153|gb|AAC02573.1|  (AF045640) Hypothetical protein C11D2.4 [Caenorhabditis elegans]
          Length = 543

 Score = 36.2 bits (82), Expect = 0.17
 Identities = 22/78 (28%), Positives = 41/78 (52%), Gaps = 4/78 (5%)
 Frame = +2

Query: 212 SCQFVSENSKHVKINKEELQKQVNEWFI----NNKHKYKEFDEYECHLAADEEQVVDFLF 379
           S +F+ E+   VKIN+  ++K V E  +    +   K   F   E H  + ++  V ++
Sbjct: 12  SGKFIVEHGNLVKINQAGVRK-VAEHILGAAKDGSLKEALFLSPELHPKSGDKDAVQWVL 70

Query: 380 ILDSLNFCFWPQKDYEYE 433
           ++D++NF FWP +   Y+
Sbjct: 71  LVDTINFSFWPDEGDHYD 88




>gi|80469|pir||B29838  parasporal crystal protein - Bacillus thuringiensis subsp.
israelensis (fragment)
          Length = 934

 Score = 33.1 bits (74), Expect = 1.5
 Identities = 28/112 (25%), Positives = 50/112 (44%)
 Frame = +2

Query: 167 HLYQYNIQQMSLVRXSCQFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYECHLA 346
           H Y   I ++  +  +   +    K    N E+ QK VN+ F+NN    K   + +
Sbjct: 635 HSYNIYIDKIEFIPITQSVLDYTEKQ---NIEKTQKIVNDLFVNN----KVLTKIDSMAV 687

Query: 347 ADEEQVVDFLFILDSLNFCFWPQKDYEYENLSSAIKECFQKSPQQFKAESIL 502
               +VV+ +F   + N       DYE + +++AI EC   S +Q+  E ++
Sbjct: 688 KKARKVVNPMFTSGAKNRLKLETTDYEIDQVANAI-ECM--SDEQYSKEKLM 736




>gi|15925464|ref|NP_372998.1|  (NC_002758) conserved hypothetical protein [Staphylococcus
aureus subsp. aureus Mu50]
 gi|15928053|ref|NP_375586.1|  (NC_002745) conserved hypothetical protein [Staphylococcus
aureus subsp. aureus N315]
 gi|13702424|dbj|BAB43565.1|  (AP003137) conserved hypothetical protein [Staphylococcus
aureus subsp. aureus N315]
 gi|14248248|dbj|BAB58636.1|  (AP003365) conserved hypothetical protein [Staphylococcus
aureus subsp. aureus Mu50]
          Length = 140

 Score = 32.3 bits (72), Expect = 2.5
 Identities = 26/136 (19%), Positives = 64/136 (46%), Gaps = 4/136 (2%)
 Frame = +2

Query: 2   YFIKQDNGQANALEEKNQPFSIYNIDNQTIDDSLHLSVLFLFGK*TLPYLCAFLLHLYQY 181
           Y  K+ NG+  ++  + + F    I    ++D L   +  L  K      C + + +++
Sbjct: 3   YNYKKQNGELMSVMSQGEKF----IHQSPVNDELSALIKLLISKING---CHYCVDIHKK 55

Query: 182 NIQQMSLVRXSCQFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYEC----HLAA 349
            ++++ + +     V  + +H+ +  +  Q++V   F    +  K+F ++E
Sbjct: 56  ELKELGVTQMKIDEVL-SFRHLDLFTD--QEKVTLEFAEMLNSIKDFKKFEIIDRLKSFY 112
```

```
Query: 350 DEEQVVDFLFILDSLN 397
           DEEQ++D +F+++ +N
Sbjct: 113 DEEQIIDLVFVVNQIN 128




>gi|111664|pir||A40467  glucuronosyltransferase (EC 2.4.1.17) precursor - rat
          Length = 530

 Score = 32.0 bits (71), Expect = 3.3
 Identities = 18/49 (36%), Positives = 26/49 (52%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
            S D   + FI  L +W        C S++ IL Q L+ EFS  +LS+C+
Sbjct: 82  SKDELQNHFIKLLDVWTYELPRDTCLSYSPIL-QNLVYEFSYFYLSICK 129




>gi|13928718|ref|NP_113721.1|  (NM_031533) androsterone UDP-glucuronosyltransferase
[Rattus norvegicus]
 gi|136732|sp|P08541|UDB2_RAT  UDP-glucuronosyltransferase 2B2 precursor, microsomal (UDPGT)
(3-hydroxyandrogen specific) (UDPGTR-4) (RLUG23)
 gi|92284|pir||A26064  glucuronosyltransferase (EC 2.4.1.17) 4 - rat
 gi|207583|gb|AAA42314.1|  (J02589) UDP glucuronosyltransferase [Rattus norvegicus]
          Length = 530

 Score = 32.0 bits (71), Expect = 3.3
 Identities = 18/49 (36%), Positives = 26/49 (52%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
            S D   + FI  L +W        C S++ IL Q L+ EFS  +LS+C+
Sbjct: 82  SKDELQNHFIKLLDVWTYELPRDTCLSYSPIL-QNLVYEFSYFYLSICK 129




>gi|57453|emb|CAA27198.1|  (X03478) UDPGT (aa 1-500) [Rattus norvegicus]
          Length = 500

 Score = 32.0 bits (71), Expect = 3.3
 Identities = 18/49 (36%), Positives = 26/49 (52%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
            S D   + FI  L +W        C S++ IL Q L+ EFS  +LS+C+
Sbjct: 52  SKDELQNHFIKLLDVWTYELPRDTCLSYSPIL-QNLVYEFSYFYLSICK 99




>gi|15896731|ref|NP_350080.1|  (NC_003030) Transcription regulator, YOBV B.subtilis homolog
[Clostridium acetobutylicum]
 gi|15026585|gb|AAK81420.1|AE007846_11  (AE007846) Transcription regulator, YOBV B.subtilis
homolog [Clostridium acetobutylicum]
          Length = 300

 Score = 32.0 bits (71), Expect = 3.3
```

```
 Identities = 26/100 (26%), Positives = 54/100 (54%), Gaps = 9/100 (9%)
 Frame = +2

Query: 221 FVSENSKHVKINKEELQKQVNEWF---------INNKHKYKEFDEYECHLAADEEQVVDF 373
           +VSE S ++K  K+ + K+ N ++         I++K  Y+ +DE+E     D +++ D
Sbjct: 198 YVSEESFNIKAPKK-IMKEDNVFYDEFITLKMKISSKMAYRIYDEFE-----DFKKLDDG 251

Query: 374 LFILDSLNFCFWPQKDYEYENLSSAIKECFQKSPQQFKAE 493
            FI+ +     +P+ ++ +  +SS  +EC   +PQ+ + E
Sbjct: 252 SFIVTAE----YPKGEWLFSYISSFGEECEVLAPQEIREE 287



>gi|57449|emb|CAA68351.1|  (Y00156) precursor polypeptide (AA -23 to 507) [Rattus
norvegicus]
          Length = 530

 Score = 31.6 bits (70), Expect = 4.3
 Identities = 20/59 (33%), Positives = 31/59 (51%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE*EFQTCQNQQ 258
           S D   + FI  + +W        C S+ S L+Q +I+EFS  +LS+C+    T  N+Q
Sbjct: 82  SKDELENYFIKLVDVWTYELQRDTCLSY-SPLLQNMIDEFSDYYLSLCK---DTVSNKQ 136



>gi|19527110|ref|NP_598655.1|  (NM_133894) expressed sequence AA986709 [Mus musculus]
 gi|15929692|gb|AAH15272.1|AAH15272  (BC015272) Similar to UDP-glucuronosyltransferase 2
family, member 5 [Mus musculus]
          Length = 530

 Score = 31.6 bits (70), Expect = 4.3
 Identities = 18/49 (36%), Positives = 27/49 (54%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
           S D   + FI  + +W        C S+ S L+Q +I+EFS  FLS+C+
Sbjct: 82  SKDNLENFFIKFVDVWTYEMPRDTCLSY-SPLLQNMIDEFSDYFLSLCK 129



>gi|6678501|ref|NP_033493.1|  (NM_009467) UDP-glucuronosyltransferase 2 family, member 5
[Mus musculus]
 gi|136725|sp|P17717|UDB5_MOUSE  UDP-glucuronosyltransferase 2B5 precursor, microsomal
(UDPGT) (M-1)
 gi|90521|pir||S00163  glucuronosyltransferase (EC 2.4.1.17) precursor - mouse
 gi|55120|emb|CAA29657.1|  (X06358) UDP-glucuronosyltransferase precursor (530 AA) [Mus
musculus]
          Length = 530

 Score = 31.6 bits (70), Expect = 4.3
 Identities = 18/49 (36%), Positives = 27/49 (54%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
           S D   + FI  + +W        C S+ S L+Q +I+EFS  FLS+C+
Sbjct: 82  SKDNLENFFIKFVDVWTYEMPRDTCLSY-SPLLQNMIDEFSDYFLSLCK 129
```

```
>gi|20381430|gb|AAH28262.1|  (BC028262) UDP-glucuronosyltransferase 2 family, member 5 [Mus
musculus]
          Length = 530

 Score = 31.6 bits (70), Expect = 4.3
 Identities = 18/49 (36%), Positives = 27/49 (54%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE 228
           S D   + FI   + +W        C S+  S L+Q +I+EFS   FLS+C+
Sbjct: 82  SKDNLENFFIKFVDVWTYEMPRDTCLSY-SPLLQNMIDEFSDYFLSLCK 129



>gi|92287|pir||S07390  glucuronosyltransferase (EC 2.4.1.17) 3 precursor - rat
          Length = 530

 Score = 31.6 bits (70), Expect = 4.3
 Identities = 20/59 (33%), Positives = 31/59 (51%)
 Frame = +1

Query: 82  SNDR*LSAFICTLSIWQMNSSIFMCFSFTSILIQYLINEFSKXFLSVCE*EFQTCQNQQ 258
           S D   + FI   + +W        C S+  S L+Q +I+EFS   +LS+C+    T  N+Q
Sbjct: 82  SKDELENYFIKLVDVWTYELQRDTCLSY-SPLLQNMIDEFSDYYLSLCK---DTVSNKQ 136



>gi|7499184|pir||T20978  hypothetical protein F15D3.9 - Caenorhabditis elegans
          Length = 2261

 Score = 31.2 bits (69), Expect = 5.6
 Identities = 12/39 (30%), Positives = 23/39 (58%)
 Frame = +2

Query: 218 QFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYE 334
           QF   +N+   V++ +E+L  +  + +W I      Y+E D Y+
Sbjct: 115 QFFWKNNAPVQLREEKLSEAIKQWCIEVMKSYEEIDVYD 153



>gi|3892157|emb|CAA10033.1|  (AJ012469) DYS-1 protein [Caenorhabditis elegans]
          Length = 3674

 Score = 31.2 bits (69), Expect = 5.6
 Identities = 12/39 (30%), Positives = 23/39 (58%)
 Frame = +2

Query: 218 QFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYE 334
           QF   +N+   V++ +E+L  +  + +W I      Y+E D Y+
Sbjct: 115 QFFWKNNAPVQLREEKLSEAIKQWCIEVMKSYEEIDVYD 153
```

```
>gi|17506447|ref|NP_492946.1|  (NM_060545) dystrophin [Caenorhabditis elegans]
 gi|14530423|emb|CAB61012.2|  (Z81063) contains similarity to Pfam domain: PF00307 (Calponin
homology (CH) domain), Score=35.0, E-value=5.6e-07, N=1; PF00397 (WW domain), Score=29.5, E-
value=2.4e-05, N=1; PF00435 (Spectrin repeat), Score=66.1, E-value=2.5e-17, N=10; PF00569
(Zinc f>
 gi|14530444|emb|CAB61005.2|  (Z81522) contains similarity to Pfam domain: PF00307 (Calponin
homology (CH) domain), Score=35.0, E-value=5.6e-07, N=1; PF00397 (WW domain), Score=29.5, E-
value=2.4e-05, N=1; PF00435 (Spectrin repeat), Score=66.1, E-value=2.5e-17, N=10; PF00569
(Zinc f>
          Length = 3674

 Score = 31.2 bits (69), Expect = 5.6
 Identities = 12/39 (30%), Positives = 23/39 (58%)
 Frame = +2

Query: 218 QFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYE 334
           QF  +N+  V++ +E+L + + +W I     Y+E D Y+
Sbjct: 115 QFFWKNNAPVQLREEKLSEAIKQWCIEVMKSYEEIDVYD 153



>gi|7499176|pir||T20971  hypothetical protein F15D3.1 - Caenorhabditis elegans
          Length = 2241

 Score = 31.2 bits (69), Expect = 5.6
 Identities = 12/39 (30%), Positives = 23/39 (58%)
 Frame = +2

Query: 218 QFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEYE 334
           QF  +N+  V++ +E+L + + +W I     Y+E D Y+
Sbjct: 115 QFFWKNNAPVQLREEKLSEAIKQWCIEVMKSYEEIDVYD 153



>gi|19703788|ref|NP_603350.1|  (NC_003454) Xaa-Pro aminopeptidase [Fusobacterium nucleatum
subsp. nucleatum ATCC 25586]
 gi|19713932|gb|AAL94649.1|  (AE010557) Xaa-Pro aminopeptidase [Fusobacterium nucleatum
subsp. nucleatum ATCC 25586]
          Length = 584

 Score = 30.8 bits (68), Expect = 7.3
 Identities = 14/43 (32%), Positives = 26/43 (59%)
 Frame = +2

Query: 203 VRXSCQFVSENSKHVKINKEELQKQVNEWFINNKHKYKEFDEY 331
           V  S   +SE    + IN+++L K+  ++F +NK + KE+ E+
Sbjct: 209 VALSFTIISEKKSTLYINEKKLDKKAQKYFKDNKVEIKEYFEF 251



>gi|15792817|ref|NP_282640.1|  (NC_002163) putative proline dehydrogenase/delta-1-pyrroline-
5-carboxylate dehydrogenase [Campylobacter jejuni]
 gi|11347200|pir||C81297  proline dehydrogenase (EC 1.5.99.8) / 1-pyrroline-5-carboxylate
dehydrogenase (EC 1.5.1.12) Cj1503c [similarity] - Campylobacter jejuni (strain NCTC 11168)
 gi|6968928|emb|CAB73924.1|  (AL139078) putative proline dehydrogenase/delta-1-pyrroline-5-
carboxylate dehydrogenase [Campylobacter jejuni]
          Length = 1162
```

```
 Score = 30.8 bits (68), Expect = 7.3
 Identities = 17/64 (26%), Positives = 34/64 (52%), Gaps = 4/64 (6%)
 Frame = -1

Query: 308 CVYYQ*TTH*LIFVIPPC*FQHVWNSHS----QTDKNXLLNSFIKYYINIDVKEKHINME 141
           C Y     H LI   P C   H  N+ +    + D+N   ++F++Y+ N+ V +K+ N++
Sbjct: 373 CSYELSKMHDLILYTPVCDEAHFNNAIAYLVRRLDENTSEDNFMRYFFNLKVGDKNWNIQ 432

Query: 140 EFIY 129
           + ++
Sbjct: 433 KELF 436




>gi|14578290|gb|AAF99456.1|  (AY003872) PV1H14050_P [Plasmodium vivax]
          Length = 1364

 Score = 30.4 bits (67), Expect = 9.5
 Identities = 11/18 (61%), Positives = 14/18 (77%)
 Frame = -3

Query: 198  LIYQILYQYRCKRKAHKY 145
            +IY ILY Y C+ KA+KY
Sbjct: 1111 IIYNILYSYNCRTKANKY 1128




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 262,568,158
Number of Sequences: 919285
Number of extensions: 5117458
Number of successful extensions: 19504
Number of sequences better than 10.0: 44
Number of HSP's better than 10.0 without gapping: 18743
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 19491
length of database: 288,558,979
effective HSP length: 112
effective length of database: 185,599,059
effective search space used: 12249537894
frameshift window, decay const: 50,  0.1
```

```
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020953866-015044-14056
```

7.1.9 Query= hy-9_13b236005

```
         (306 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 82 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|19263001|dbj|BAB85849.1|  (AB079614) pf16 protein [Ciona ...   130   2e-30
gi|6912678|ref|NP_036575.1|  (NM_012443) sperm associated an...   130   4e-30
gi|1101777|gb|AAC49169.1|  (U40057) PF16 [Chlamydomonas rein...   129   5e-30
gi|7657607|ref|NP_056588.1|  (NM_015773) sperm associated an...   128   1e-29
gi|9366742|emb|CAB95504.1|  (AL359782) probable axoneme cent...   121   2e-27
gi|10046851|emb|CAC07966.1|  (AJ293290) putative mitogen-act...    42   0.001
gi|8745041|emb|CAB95259.1|  (AL359777) probable MAP kinase k...    40   0.004
gi|11359260|pir||T40577  probable phosphatidylinositol 3-kin...    39   0.015
gi|4160337|emb|CAA22805.1|  (AL035216) putative phosphatidyl...    39   0.015
gi|20161458|dbj|BAB90382.1|  (AP003561) hypothetical protein...    37   0.033
gi|7295403|gb|AAF50720.1|  (AE003564) CG10478 gene product [...    37   0.033
gi|18550453|ref|XP_086974.1|  (XM_086974) similar to unnamed...    36   0.096
gi|15227784|ref|NP_179895.1|  (NM_127878) hypothetical prote...    35   0.13
gi|15239713|ref|NP_197434.1|  (NM_121938) putative protein [...    35   0.13
gi|15233106|ref|NP_191045.1|  (NM_115342) putative protein [...    34   0.28
gi|17529090|gb|AAL38755.1|  (AY065279) unknown protein [Arab...    34   0.28
gi|17540124|ref|NP_501227.1|  (NM_068826) serine-rich RNA po...    34   0.37
gi|2959322|emb|CAA74966.1|  (Y14616) Importin alpha-like pro...    33   0.63
gi|15242560|ref|NP_195908.1|  (NM_120366) putative protein [...    32   1.1
gi|5802657|gb|AAD51751.1|  (AF176651) pendulin [Oreochromis ...    32   1.1
gi|6320824|ref|NP_010903.1|  (NC_001137) An armadillo repeat...    32   1.1
gi|15242618|ref|NP_195927.1|  (NM_120385) importin alpha - l...    32   1.1
gi|7630132|emb|CAB71185.4|  (AL138518) importin alpha-like p...    32   1.4
gi|16758242|ref|NP_445935.1|  (NM_053483) karyopherin (impor...    32   1.4
```

```
gi|1363205|pir||S57873  pendulin - mouse >gi|555823|gb|AAA85...   32   1.4
gi|6754474|ref|NP_034785.1|  (NM_010655) karyopherin (import...   32   1.4
gi|17942565|pdb|1IQ1|C  Chain C, Crystal Structure Of The Im...   32   1.4
gi|15638984|gb|AAG42103.2|  (AY010065) karyopherin alpha 2 [...   32   1.4
gi|5542272|pdb|1IAL|A  Chain A, Importin Alpha, Mouse             32   1.4
gi|4505841|ref|NP_000290.1|  (NM_000299) plakophilin 1; Plak...   32   1.8
gi|17465256|ref|XP_069697.1|  (XM_069697) similar to IMPORTI...   32   1.8
gi|20138951|sp|Q13835|PKP1_HUMAN  Plakophilin 1 (Band-6-prot...   32   1.8
gi|17552816|ref|NP_499306.1|  (NM_066905) D2045.2.p [Caenorh...   32   1.8
gi|2134803|pir||S60712  band-6-protein - human >gi|535015|em...   32   1.8
gi|4504897|ref|NP_002257.1|  (NM_002266) karyopherin alpha 2...   32   1.8
gi|15238931|ref|NP_199049.1|  (NM_123599) arm repeat contain...   30   4.1
gi|9790161|ref|NP_062619.1|  (NM_019645) plakophilin 1 [Mus ...   30   4.1
gi|7288998|gb|AAF45332.1|  (AE003091) CG17514 gene product [...   30   5.3
gi|6822222|emb|CAB70928.1|  (AL137778) putative acyltransfer...   30   5.3
gi|4680191|gb|AAD27556.1|AF111710_2  (AF111710) putative myb...   30   5.3
gi|18017587|ref|NP_542121.1|  (NM_080554) proteasome (prosom...   30   5.3
gi|4325344|gb|AAD17343.1|  (AF128393) similar to beta-transd...   30   6.9
gi|15223673|ref|NP_175502.1|  (NM_103969) hypothetical prote...   30   6.9
gi|3299917|gb|AAC25988.1|  (AF047839) putative glucosyl hydr...   30   6.9
gi|1401236|gb|AAB03389.1|  (U59448) receptor tyrosine kinase...   30   6.9
gi|15219352|ref|NP_175078.1|  (NM_103538) hypothetical prote...   29   9.0
gi|15242247|ref|NP_200013.1|  (NM_124579) importin alpha sub...   29   9.0
gi|4504903|ref|NP_002260.1|  (NM_002269) karyopherin alpha 5...   29   9.0
gi|11418581|ref|XP_004308.1|  (XM_004308) karyopherin alpha ...   29   9.0
```

Alignments

```
>gi|19263001|dbj|BAB85849.1|  (AB079614) pf16 protein [Ciona intestinalis]
          Length = 476

 Score =  130 bits (328), Expect = 2e-30
 Identities = 69/92 (75%), Positives = 78/92 (84%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           LALGRLA +++DLAEAVV  +ILPQLV SL EQNRFYKKAAAFVLR VAKHS  LA +VV
Sbjct: 63  LALGRLANYNDDLAEAVVKGDILPQLVYSLAEQNRFYKKAAAFVLRAVAKHSPQLAQSVV 122

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           + GAL+ALV CLEEFDP VKE+AA AL YIA+
Sbjct: 123 DCGALDALVICLEEFDPGVKESAAWALGYIAR 154




 Score = 65.9 bits (159), Expect = 9e-11
 Identities = 34/96 (35%), Positives = 57/96 (58%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL +++KHS DL+E VV  EI P +++ L + + + +K  A ++R +AKH+  LA  +VN
Sbjct: 232 ALSQISKHSVDLSEMVVEAEIFPAVLTCLRDPDEYVRKNVATLIREIAKHTPELAQLIVN 291


Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNI 15
           +G + A+V  + E   +V+      L YIA    N+
Sbjct: 292 AGGVAAIVDYVGESKGNVRLPGVMTLGYIAAHSDNL 327
```

```
 Score = 63.2 bits (152), Expect = 6e-10
 Identities = 37/90 (41%), Positives = 58/90 (64%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  +AKHS  LA++VV    L  LV  L E +   K++AA+ L  +A+H++ L+ AVV++
Sbjct: 107 LRAVAKHSPQLAQSVVDCGALDALVICLEEFDPGVKESAAWALGYIARHNAELSQAVVDA 166

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
           GA+  LV C++E + S+K  +ASAL  + K
Sbjct: 167 GAVPLLVLCIQEPELSLKRISASALSDVCK 196
```

```
 Score = 58.5 bits (140), Expect = 1e-08
 Identities = 35/91 (38%), Positives = 51/91 (55%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           ALG +A+H+ +L++AVV    +P LV  + E     K+ +A  L  V KHS  LA  VV+
Sbjct: 148 ALGYIARHNAELSQAVVDAGAVPLLVLCIQEPELSLKRISASALSDVCKHSPELAQTVVD 207

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +GA+  L Q +   D  +K    SAL  I+K
Sbjct: 208 AGAIAHLAQMILNPDAKLKRQVFSALSQISK 238
```

```
 Score = 46.2 bits (108), Expect = 7e-05
 Identities = 28/91 (30%), Positives = 46/91 (49%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  + KHS +LA+ VV    +  L   +   +   K+     L  ++KHS  L+  VV
Sbjct: 190 ALSDVCKHSPELAQTVVDAGAIAHLAQMILNPDAKLKRQVFSALSQISKHSVDLSEMVVE 249

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +    A++ CL + D  V++  A+ +R IAK
Sbjct: 250 AEIFPAVLTCLRDPDEYVRKNVATLIREIAK 280
```

```
 Score = 42.0 bits (97), Expect = 0.001
 Identities = 26/88 (29%), Positives = 44/88 (49%), Gaps = 1/88 (1%)
 Frame = -2


Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAV-VNSGA 114
           +AKH+ +LA+ +V+   +  +V  +GE     +      L  +A HS  LAMAV V+ G
Sbjct: 278 IAKHTPELAQLIVNAGGVAAIVDYVGESKGNVRLPGVMTLGYIAAHSDNLAMAVIVSKGV 337
```

```
Query: 113 LEALVQCLEEFDPSVKEAAASALRYIAK 30
            +  +   EE +  +K A   +L  I +
Sbjct: 338 NQLAITLAEESEDHIKAAVGWSLGQIGR 365




 Score = 38.1 bits (87), Expect = 0.019
 Identities = 24/71 (33%), Positives = 42/71 (58%), Gaps = 1/71 (1%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAA-AFVLRCVAKHSSTLAMAV 129
           + LG +A HS++LA AV+ ++ + QL  +L E++  + KAA  + L  + +H+   A AV
Sbjct: 315 MTLGYIAAHSDNLAMAVIVSKGVNQLAITLAEESEDHIKAAVGWSLGQIGRHTPEHAKAV 374

Query: 128 VNSGALEALVQ 96
             +  L  L+Q
Sbjct: 375 AVANVLPRLLQ 385




 Score = 33.9 bits (76), Expect = 0.37
 Identities = 22/55 (40%), Positives = 30/55 (54%)
 Frame = -2

Query: 194 KKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           ++ AA  L  +A ++  LA AVV    L  LV  L E +   K+AAA  LR +AK
Sbjct: 58  QQTAALALGRLANYNDDLAEAVVKGDILPQLVYSLAEQNRFYKKAAAFVLRAVAK 112



>gi|6912678|ref|NP_036575.1|  (NM_012443) sperm associated antigen 6 [Homo sapiens]
 gi|7512612|pir||T12521 hypothetical protein DKFZp434I153.1 - human
 gi|3449383|gb|AAC32590.1|  (AF079363) sperm flagellar protein Repro-SA-1 [Homo sapiens]
 gi|5262578|emb|CAB45730.1|  (AL080136) hypothetical protein [Homo sapiens]
          Length = 509

 Score =  130 bits (326), Expect = 4e-30
 Identities = 68/92 (73%), Positives = 77/92 (82%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           LALGRLA +++DLAEAVV  +ILPQLV SL EQNRFYKKAAAFVLR V KHS  LA A+V
Sbjct: 63  LALGRLANYNDDLAEAVVKCDILPQLVYSLAEQNRFYKKAAAFVLRAVGKHSPQLAQAIV 122

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           + GAL+ LV CLE+FDP VKEAAA ALRYIA+
Sbjct: 123 DCGALDTLVICLEDFDPGVKEAAAWALRYIAR 154




 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 33/96 (34%), Positives = 59/96 (61%)
 Frame = -2
```

```
Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL +++KHS DLAE VV  EI P +++ L +++ + KK A+ ++R +AKH+  L+  VVN
Sbjct: 232 ALSQVSKHSVDLAEMVVEAEIFPVVLTCLKDKDEYVKKNASTLIREIAKHTPELSQLVVN 291

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNI 15
           +G + A++ C+     + +      L Y+A   +N+
Sbjct: 292 AGGVAAVIDCIGSCKGNTRLPGIMMLGYVAAHSENL 327




 Score = 67.0 bits (162), Expect = 4e-11
 Identities = 39/90 (43%), Positives = 59/90 (65%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  + KHS  LA+A+V    L  LV  L + +   K+AAA+ LR +A+H++ L+ AVV++
Sbjct: 107 LRAVGKHSPQLAQAIVDCGALDTLVICLEDFDPGVKEAAAWALRYIARHNAELSQAVVDA 166

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
           GA+  LV C++E + ++K  AASAL  IAK
Sbjct: 167 GAVPLLVLCIQEPEIALKRIAASALSDIAK 196




 Score = 56.6 bits (135), Expect = 5e-08
 Identities = 34/91 (37%), Positives = 51/91 (55%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  +A+H+ +L++AVV    +P LV  + E     K+ AA  L  +AKHS  LA  VV+
Sbjct: 148 ALRYIARHNAELSQAVVDAGAVPLLVLCIQEPEIALKRIAASALSDIAKHSPELAQTVVD 207

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +GA+  L Q +   D  +K    SAL  ++K
Sbjct: 208 AGAVAHLAQMILNPDAKLKHQILSALSQVSK 238




 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 31/91 (34%), Positives = 47/91 (51%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  +AKHS +LA+ VV    +  L   +   +   K      L  V+KHS  LA  VV
Sbjct: 190 ALSDIAKHSPELAQTVVDAGAVAHLAQMILNPDAKLKHQILSALSQVSKHSVDLAEMVVE 249

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +     ++ CL++ D  VK+ A++ +R IAK
Sbjct: 250 AEIFPVVLTCLKDKDEYVKKNASTLIREIAK 280
```

```
 Score = 40.8 bits (94), Expect = 0.003
 Identities = 26/70 (37%), Positives = 41/70 (58%), Gaps = 1/70 (1%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFY-KKAAAFVLRCVAKHSSTLAMAV 129
           + LG +A HSE+LA AV+ ++ +PQL   L E+   + K AAA+ L  + +H+   A AV
Sbjct: 315 MMLGYVAAHSENLAMAVIISKGVPQLSVCLSEEPEDHIKAAAAWALGQIGRHTPEHARAV 374

Query: 128 VNSGALEALV 99
             +  L  L+
Sbjct: 375 AVTNTLPVLL 384




 Score = 32.0 bits (71), Expect = 1.4
 Identities = 21/55 (38%), Positives = 29/55 (52%)
 Frame = -2

Query: 194 KKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           ++ AA  L  +A ++   LA AVV    L  LV  L E +   K+AAA  LR + K
Sbjct: 58  QQTAALALGRLANYNDDLAEAVVKCDILPQLVYSLAEQNRFYKKAAAFVLRAVGK 112



>gi|1101777|gb|AAC49169.1|  (U40057) PF16 [Chlamydomonas reinhardtii]
 gi|1587371|prf||2206442A  PF16 gene [Chlamydomonas reinhardtii]
          Length = 566

 Score =  129 bits (325), Expect = 5e-30
 Identities = 67/91 (73%), Positives = 77/91 (83%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           LALGRLA +S+DLAEAVV NEILPQLV SL EQNRFYK+AAAF LR VA+HS  LA +V+
Sbjct: 63  LALGRLANYSDDLAEAVVQNEILPQLVYSLSEQNRFYKQAAAFCLRAVARHSPELAQSVI 122

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIA 33
           +SGAL++LV CLEEFDP VKEA+A  L YIA
Sbjct: 123 DSGALDSLVTCLEEFDPGVKEASAWTLGYIA 153



 Score = 65.9 bits (159), Expect = 9e-11
 Identities = 39/90 (43%), Positives = 55/90 (60%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL ++AKHS DLAE VV  EI P++++ L   + F KK +A V+R VAKH+  LA  VV
Sbjct: 232 ALSQIAKHSVDLAEVVVGAEIFPKILTCLKFPDEFVKKHSATVVREVAKHTPELAHVVVG 291

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIA 33
           +G + ALV  + +   + +     AL YIA
Sbjct: 292 NGGVGALVDYISDSAGNNRLPGIMALGYIA 321
```

```
 Score = 63.5 bits (153), Expect = 4e-10
 Identities = 36/90 (40%), Positives = 60/90 (66%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  +A+HS +LA++V+ +  L  LV+ L E +   K+A+A+ L  +A H++ +A  VV++
Sbjct: 107 LRAVARHSPELAQSVIDSGALDSLVTCLEEFDPGVKEASAWTLGYIAGHNADVAQQVVDA 166

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
           GA+  LV C++E + S+K  AASAL  I+K
Sbjct: 167 GAVPLLVLCVQEPELSLKRIAASALSDISK 196




 Score = 53.5 bits (127), Expect = 4e-07
 Identities = 33/90 (36%), Positives = 48/90 (52%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           LG +A H+ D+A+ VV    +P LV  + E     K+ AA  L  ++KH+  LA AVV++
Sbjct: 149 LGYIAGHNADVAQQVVDAGAVPLLVLCVQEPELSLKRIAASALSDISKHTPELAQAVVDA 208

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
           GA+  L   +   D  +K     AL  IAK
Sbjct: 209 GAVAYLAPLVINQDAKLKRQVCCALSQIAK 238




 Score = 50.4 bits (119), Expect = 4e-06
 Identities = 35/104 (33%), Positives = 57/104 (54%), Gaps = 3/104 (2%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFY-KKAAAFVLRCVAKHSSTLAMAV 129
           +ALG +A  SE LA +V++ + LP LVS+L E+   + K A A+ L  + +H+   A AV
Sbjct: 315 MALGYIAAFSETLALSVIAEKALPPLVSALNEEPEDHLKSATAWTLGQIGRHTPDHAKAV 374

Query: 128 VNSGALEALVQCLEEFDPS--VKEAAASALRYIAKQLQNIYCLD 3
            ++G L  LV    +   S  +K     AL+ +  +L ++  LD
Sbjct: 375 ADTGCLATLVSLESDGASSDDLKTKCRRALKSVIAKLTHLPALD 418




 Score = 49.3 bits (116), Expect = 8e-06
 Identities = 30/91 (32%), Positives = 49/91 (52%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  ++KH+ +LA+AVV    +  L   +  Q+   K+     L  +AKHS  LA  VV
Sbjct: 190 ALSDISKHTPELAQAVVDAGAVAYLAPLVINQDAKLKRQVCCALSQIAKHSVDLAEVVVG 249

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +     ++ CL+  D  VK+ +A+ +R +AK
Sbjct: 250 AEIFPKILTCLKFPDEFVKKHSATVVREVAK 280
```

```
 Score = 36.2 bits (82), Expect = 0.074
 Identities = 29/91 (31%), Positives = 48/91 (51%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           A+  +AK+ +++ EA+     +  L   L +     +++AA  L  +A +S  LA AVV
Sbjct: 23  AVAEMAKNPQNI-EALQQAGAMALLRPLLLDNVPSIQQSAALALGRLANYSDDLAEAVVQ 81

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +  L  LV  L E +   K+AAA  LR +A+
Sbjct: 82  NEILPQLVYSLSEQNRFYKQAAAFCLRAVAR 112



>gi|7657607|ref|NP_056588.1| ☐ (NM_015773) sperm associated antigen 6 [Mus musculus]
 gi|7025325|gb|AAF35831.1|AF173866_1 ☐ (AF173866) axoneme central apparatus protein [Mus
musculus]
 gi|12838461|dbj|BAB24211.1| ☐ (AK005732) data source:MGD, source key:MGI:1354388,
evidence:ISS~putative~sperm associated antigen 6 [Mus musculus]
          Length = 507

 Score =  128 bits (322), Expect = 1e-29
 Identities = 67/92 (72%), Positives = 76/92 (81%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           LALGRLA +++DLAEAVV  +ILPQLV SL EQNRFYKKAAAFVLR V KHS  LA A+V
Sbjct: 63  LALGRLANYNDDLAEAVVKGDILPQLVYSLAEQNRFYKKAAAFVLRAVGKHSPQLAQAIV 122

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           + GAL+ LV CLE+FDP VKEAAA AL YIA+
Sbjct: 123 DCGALDTLVICLEDFDPGVKEAAAWALGYIAR 154



 Score = 69.7 bits (169), Expect = 6e-12
 Identities = 33/96 (34%), Positives = 59/96 (61%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL ++AKHS DLAE VV  EI P +++ L +++ + KK A  ++R +AKH+  L+  +VN
Sbjct: 232 ALSQIAKHSVDLAEMVVEAEIFPVVLTCLKDKDEYVKKNACTLIREIAKHTPELSQLIVN 291

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNI 15
           +G + A++ C+     +++      L Y+A   +N+
Sbjct: 292 AGGVAAVIDCIGSCKGNIRLPGIMMLGYVAAHSENL 327



 Score = 63.5 bits (153), Expect = 4e-10
 Identities = 37/90 (41%), Positives = 58/90 (64%)
 Frame = -2
```

```
Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  + KHS  LA+A+V    L  LV  L + +   K+AAA+ L  +A+H++ L+ AVV++
Sbjct: 107 LRAVGKHSPQLAQAIVDCGALDTLVICLEDFDPGVKEAAAWALGYIARHNTELSQAVVDA 166

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
           GA+  LV C++E + ++K  AASAL  I+K
Sbjct: 167 GAIPLLVLCIQEPEIALKRIAASALSDISK 196




 Score = 61.2 bits (147), Expect = 2e-09
 Identities = 36/91 (39%), Positives = 52/91 (56%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           ALG +A+H+ +L++AVV    +P LV  + E     K+ AA  L  ++KHS  LA  VV+
Sbjct: 148 ALGYIARHNTELSQAVVDAGAIPLLVLCIQEPEIALKRIAASALSDISKHSPELAQTVVD 207

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +GA+  L Q +   D  +K    SAL  IAK
Sbjct: 208 AGAIAHLAQMILNPDAKLKRQVLSALSQIAK 238




 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 30/91 (32%), Positives = 47/91 (50%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  ++KHS +LA+ VV    +  L    +   +   K+     L  +AKHS  LA  VV
Sbjct: 190 ALSDISKHSPELAQTVVDAGAIAHLAQMILNPDAKLKRQVLSALSQIAKHSVDLAEMVVE 249

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +     ++ CL++ D  VK+ A + +R IAK
Sbjct: 250 AEIFPVVLTCLKDKDEYVKKNACTLIREIAK 280




 Score = 40.0 bits (92), Expect = 0.005
 Identities = 26/70 (37%), Positives = 41/70 (58%), Gaps = 1/70 (1%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFY-KKAAAFVLRCVAKHSSTLAMAV 129
           + LG +A HSE+LA AV+ ++ +PQL   L E+   + K AAA+ L  + +H+   A AV
Sbjct: 315 MMLGYVAAHSENLAMAVIISKGVPQLSICLSEEPEDHIKAAAAWALGQLGRHTPEHARAV 374

Query: 128 VNSGALEALV 99
             +  L  L+
Sbjct: 375 AVTNTLPVLL 384
```

```
 Score = 32.3 bits (72), Expect = 1.1
 Identities = 21/55 (38%), Positives = 29/55 (52%)
 Frame = -2

Query: 194 KKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           ++ AA  L  +A ++  LA AVV    L  LV  L E +   K+AAA  LR + K
Sbjct: 58  QQTAALALGRLANYNDDLAEAVVKGDILPQLVYSLAEQNRFYKKAAAFVLRAVGK 112



>gi|9366742|emb|CAB95504.1|  (AL359782) probable axoneme central apparatus protein.
[Trypanosoma brucei]
          Length = 513

 Score =  121 bits (303), Expect = 2e-27
 Identities = 62/92 (67%), Positives = 79/92 (85%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           LALGRLA ++E+LAE+VVS +IL QLV SLG+Q+RFYKK+AAFVLR VA+H++ LA AVV
Sbjct: 63  LALGRLANYNEELAESVVSGDILAQLVYSLGDQSRFYKKSAAFVLRSVARHNAQLAQAVV 122

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIAK 30
            S A+EALV CLE+FDP+VKE+A  AL Y+A+
Sbjct: 123 GSQAVEALVGCLEDFDPTVKESAVWALGYVAR 154




 Score = 65.1 bits (157), Expect = 1e-10
 Identities = 37/95 (38%), Positives = 59/95 (61%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  +A+H+  LA+AVV ++ +  LV  L + +   K++A + L  VA+H++ LA  VV+
Sbjct: 107 LRSVARHNAQLAQAVVGSQAVEALVGCLEDFDPTVKESAVWALGYVARHNAHLAQEVVDK 166

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAKQLQNI 15
           GA+  LV C++E + S+K  AAS L  IAK L  +
Sbjct: 167 GAIPPLVLCVQEPELSLKRTAASTLADIAKHLPEL 201



 Score = 62.0 bits (149), Expect = 1e-09
 Identities = 35/89 (39%), Positives = 51/89 (56%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L ++AKHS +LAE VV  EI P++ S L + +   +K A+  +R VAKH+  LA  +VN+
Sbjct: 233 LAQIAKHSVELAELVVEGEIFPRIFSLLKDSDETVRKNASTCIREVAKHTPELAQLIVNA 292

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIA 33
           G + ALV    E   S +      L YI+
Sbjct: 293 GGVGALVDYTNESTGSARLPGIMTLGYIS 321
```

```
 Score = 52.8 bits (125), Expect = 8e-07
 Identities = 33/91 (36%), Positives = 45/91 (49%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           ALG +A+H+  LA+ VV    +P LV  + E     K+ AA  L  +AKH   LA AVV+
Sbjct: 148 ALGYVARHNAHLAQEVVDKGAIPPLVLCVQEPELSLKRTAASTLADIAKHLPELAQAVVD 207

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
             A+  L   +   D  +K      L  IAK
Sbjct: 208 QDAVTHLAPLIGSNDGKLKRQVCQCLAQIAK 238
```

```
 Score = 49.7 bits (117), Expect = 6e-06
 Identities = 28/90 (31%), Positives = 47/90 (52%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           L  +AKH  +LA+AVV  + +  L   +G  +   K+      L  +AKHS  LA  VV
Sbjct: 191 LADIAKHLPELAQAVVDQDAVTHLAPLIGSNDGKLKRQVCQCLAQIAKHSVELAELVVEG 250

Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIAK 30
                +   L++ D +V++ A++ +R +AK
Sbjct: 251 EIFPRIFSLLKDSDETVRKNASTCIREVAK 280
```

```
 Score = 32.7 bits (73), Expect = 0.82
 Identities = 24/70 (34%), Positives = 38/70 (54%), Gaps = 1/70 (1%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLG-EQNRFYKKAAAFVLRCVAKHSSTLAMAV 129
           + LG ++  SE LA AV+ +  +  L ++L  EQ    K AAA+ L  + +HS+  A AV
Sbjct: 315 MTLGYISAFSETLALAVIVSCGIEPLSNALEKEQEDHIKSAAAWSLGQIGRHSADHAKAV 374

Query: 128 VNSGALEALV 99
            +   L  L+
Sbjct: 375 ADCNVLPKLL 384
```

```
 Score = 29.6 bits (65), Expect = 6.9
 Identities = 25/101 (24%), Positives = 46/101 (44%), Gaps = 1/101 (0%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHS-STLAMAVV 126
           +LG++ +HS D A+AV    +LP+L+      + +   +++  LR  +K +   +    V
Sbjct: 359 SLGQIGRHSADHAKAVADCNVLPKLL------DLYLHPSSSDDLRMKSKRALKNIIQRCV 412

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNIYCLD 3
              ALE L+           EA   + L+Y+   Q   +   D
Sbjct: 413 QLPALEPLLH---------PEAPKNVLKYVCGQFAKVLPTD 444
```

```
>gi|10046851|emb|CAC07966.1|  (AJ293290) putative mitogen-activated protein kinase kinase 2
[Leishmania mexicana]
          Length = 1090

 Score = 42.4 bits (98), Expect = 0.001
 Identities = 29/85 (34%), Positives = 43/85 (50%)
 Frame = -2

Query: 299  LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
            +G L KHS+   EA+  N I+ +LV    + +   +K AAF +   A HS  L   +  S
Sbjct: 916  IGNLCKHSDFFFEALEKNHIVERLVKCCSDSDAQTQKLAAFAIGNAAFHSDYLYNLL--S 973

Query: 119  GALEALVQCLEEFDPSVKEAAASAL 45
             A+ +LV  L   D   ++ AA AL
Sbjct: 974  PAIPSLVGLLASGDAKTRQNAAGAL 998



 Score = 30.0 bits (66), Expect = 5.3
 Identities = 20/100 (20%), Positives = 44/100 (44%)
 Frame = -2

Query: 305  LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
            + L ++A+ S +  E +  + +   + + L    +  +     ++  + KHS     A+
Sbjct: 872  VVLSQMARMSAEFYEPIHRSNLYDCVAALLQHSEKDLRGKTCTLIGNLCKHSDFFFEALE 931

Query: 125  NSGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNIYCL 6
             +  +E LV+C  + D   ++ AA A+   A     +Y L
Sbjct: 932  KNHIVERLVKCCSDSDAQTQKLAAFAIGNAAFHSDYLYNL 971



>gi|8745041|emb|CAB95259.1|  (AL359777) probable MAP kinase kinase [Leishmania major]
 gi|13751123|emb|CAC37137.1|  (AL359683) probable mitogen-activated protein kinase kinase
[Leishmania major]
          Length = 1090

 Score = 40.4 bits (93), Expect = 0.004
 Identities = 28/85 (32%), Positives = 42/85 (48%)
 Frame = -2

Query: 299  LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
            +G L KHS+   E +  N I+ +LV    + +   +K AAF +   A HS  L   +  S
Sbjct: 916  IGNLCKHSDFFFEPLEKNHIVERLVKCCSDSDAQTQKLAAFAIGNAAFHSDYLYNLL--S 973

Query: 119  GALEALVQCLEEFDPSVKEAAASAL 45
             A+ +LV  L   D   ++ AA AL
Sbjct: 974  PAIPSLVGLLARGDAKTRQNAAGAL 998



>gi|11359260|pir||T40577  probable phosphatidylinositol 3-kinase - fission yeast
(Schizosaccharomyces pombe)
          Length = 2337
```

```
 Score = 38.5 bits (88), Expect = 0.015
 Identities = 31/91 (34%), Positives = 47/91 (51%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           ALGRLA     L    V+ E+   L    GE+N   + AA  +L+ +AK++STL  A ++
Sbjct: 107 ALGRLAVPGGALTSEFVNFEVKRALEWLQGERNENRRYAAVLILKELAKNTSTLIYAHID 166

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           S   E L   L +   +++ A+A AL    K
Sbjct: 167 S-IFELLWHGLRDPKVTIRIASADALSEFLK 196




>gi|4160337|emb|CAA22805.1|  (AL035216) putative phosphatidylinositol kinase related kinase
[Schizosaccharomyces pombe]
          Length = 398

 Score = 38.5 bits (88), Expect = 0.015
 Identities = 31/91 (34%), Positives = 47/91 (51%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           ALGRLA     L    V+ E+   L    GE+N   + AA  +L+ +AK++STL  A ++
Sbjct: 107 ALGRLAVPGGALTSEFVNFEVKRALEWLQGERNENRRYAAVLILKELAKNTSTLIYAHID 166

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           S   E L   L +   +++ A+A AL    K
Sbjct: 167 S-IFELLWHGLRDPKVTIRIASADALSEFLK 196




>gi|20161458|dbj|BAB90382.1|  (AP003561) hypothetical protein~similar to Arabidopsis
thaliana chromosome 2, T20D16.23 [Oryza sativa (japonica cultivar-group)]
          Length = 789

 Score = 37.4 bits (85), Expect = 0.033
 Identities = 20/62 (32%), Positives = 33/62 (52%)
 Frame = -2

Query: 230 LVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAAS 51
           L+  L   +   +++A   LR +++HS    +A+ N GA+  LV  L   DPS +E A +
Sbjct: 498 LIEELKSDSAEVQRSATGELRILSRHSLENRIAIANCGAIPFLVSLLHSTDPSTQENAVT 557

Query: 50  AL 45
            L
Sbjct: 558 IL 559




>gi|7295403|gb|AAF50720.1| ☐ (AE003564) CG10478 gene product [Drosophila melanogaster]
 gi|7295404|gb|AAF50721.1| ☐ (AE003564) CG11040 gene product [Drosophila melanogaster]
          Length = 442

 Score = 37.4 bits (85), Expect = 0.033
 Identities = 25/99 (25%), Positives = 52/99 (52%), Gaps = 1/99 (1%)
 Frame = -2
```

```
Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  +A  SE+  + +++N +LP L + +   +   +     +L  +A  +     A++N
Sbjct: 254 ALINIATGSEEQIQELLNNNLLPHLSALMSNSDPDIRCQVLKLLLNIADGNIFQRHAIMN 313

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIA-KQLQNIYC 9
           +G L  +++CL+    S+K AAA  +  +A  + +N+ C
Sbjct: 314 AGLLHKILECLKADAISLKSAAALTITTLAIDKDKNLLC 352



>gi|18550453|ref|XP_086974.1|  (XM_086974) similar to unnamed protein product [Homo
sapiens]
 gi|16554196|dbj|BAB71684.1|  (AK058144) unnamed protein product [Homo sapiens]
          Length = 237

 Score = 35.8 bits (81), Expect = 0.096
 Identities = 20/92 (21%), Positives = 44/92 (47%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           A+  + + ++++ +A+     +P LV+    +    +   A  +  +A H++ +  A +
Sbjct: 45  AIAEVGRDNKEIQDAIAMEGAIPPLVALFKGKQISVQMKGAMAVESLASHNALIQKAFLE 104

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAKQ 27
               + L++ L+ F   VKE  A AL  +A Q
Sbjct: 105 KSLTKYLLKLLKAFQIDVKEQGAVALWALAGQ 136



>gi|15227784|ref|NP_179895.1|  (NM_127878) hypothetical protein [Arabidopsis thaliana]
 gi|7487248|pir||T00518  hypothetical protein T20D16.23 - Arabidopsis thaliana
 gi|2642448|gb|AAB87116.1|  (AC002391) hypothetical protein [Arabidopsis thaliana]
 gi|20197130|gb|AAM14930.1|  (AC004401) hypothetical protein [Arabidopsis thaliana]
          Length = 924

 Score = 35.4 bits (80), Expect = 0.13
 Identities = 21/63 (33%), Positives = 34/63 (53%)
 Frame = -2

Query: 233 QLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAA 54
           +LV  L   +   ++ A   LR +AKH+    + + NSGA+  LV+ L   D + +E A
Sbjct: 624 KLVEELKSSSLDTQRQATAELRLLAKHNMDNRIVIGNSGAIVLLVELLYSTDSATQENAV 683

Query: 53  SAL 45
           +AL
Sbjct: 684 TAL 686



>gi|15239713|ref|NP_197434.1|  (NM_121938) putative protein [Arabidopsis thaliana]
          Length = 636

 Score = 35.4 bits (80), Expect = 0.13
 Identities = 33/98 (33%), Positives = 47/98 (47%), Gaps = 8/98 (8%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSL-----GEQNRFYK---KAAAFVLRCVAKHSS 147
           ALG LA   E   + +V    LP LV+ L     G  +R      + AA  +  +A  +S
Sbjct: 131 ALGLLAIKPE-YQKLIVDKGALPHLVNLLKRNKDGSSSRAVNSVIRRAADAITNLAHENS 189
```

```
Query: 146 TLAMAVVNSGALEALVQCLEEFDPSVKEAAASALRYIA 33
           ++   V   G +  LV+ LE  D  V+ AAA ALR +A
Sbjct: 190 SIKTRVRVEGGIPPLVELLEFSDSKVQRAAAGALRTLA 227




 Score = 31.2 bits (69), Expect = 2.4
 Identities = 30/108 (27%), Positives = 46/108 (41%), Gaps = 13/108 (12%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSL------------GEQNRFYKKAAAFVLRCVAK 156
           L  LAK++EDL   +V    +P L++ L                   +K +AF L  +A
Sbjct: 78  LAELAKNAEDLVNVIVDGGAVPALMTHLQAPPYNDGDLAEKPYEHEVEKGSAFALGLLAI 137

Query: 155 HSSTLAMAVVNSGALEALVQCLE-EFDPSVKEAAASALRYIAKQLQNI 15
                 + +V+ GAL  LV  L+   D S   A  S +R  A  + N+
Sbjct: 138 KPEYQKL-IVDKGALPHLVNLLKRNKDGSSSRAVNSVIRRAADAITNL 184




>gi|15233106|ref|NP_191045.1|  (NM_115342) putative protein [Arabidopsis thaliana]
 gi|7486336|pir||T06735  hypothetical protein F28P10.170 - Arabidopsis thaliana
 gi|4678308|emb|CAB41099.1|  (AL049655) putative protein [Arabidopsis thaliana]
          Length = 639

 Score = 34.3 bits (77), Expect = 0.28
 Identities = 22/66 (33%), Positives = 35/66 (52%)
 Frame = -2

Query: 242 ILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKE 63
           +L +L +   EQ R    AAA  LR +AK +    + +  +GA+  LV+ L   DP  +E
Sbjct: 357 LLEKLANGTTEQQR----AAAGELRLLAKRNVDNRVCIAEAGAIPLLVELLSSPDPRTQE 412

Query: 62  AAASAL 45
            + +AL
Sbjct: 413 HSVTAL 418




>gi|17529090|gb|AAL38755.1|  (AY065279) unknown protein [Arabidopsis thaliana]
 gi|20465441|gb|AAM20180.1|  (AY096530) unknown protein [Arabidopsis thaliana]
          Length = 632

 Score = 34.3 bits (77), Expect = 0.28
 Identities = 22/66 (33%), Positives = 35/66 (52%)
 Frame = -2

Query: 242 ILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKE 63
           +L +L +   EQ R    AAA  LR +AK +    + +  +GA+  LV+ L   DP  +E
Sbjct: 350 LLEKLANGTTEQQR----AAAGELRLLAKRNVDNRVCIAEAGAIPLLVELLSSPDPRTQE 405

Query: 62  AAASAL 45
            + +AL
Sbjct: 406 HSVTAL 411
```

```
>gi|17540124|ref|NP_501227.1|   (NM_068826) serine-rich RNA polymerase I suppressor
protein (SRP1) [Caenorhabditis elegans]
 gi|10720392|sp|Q19969|IMA_CAEEL  Putative importin alpha subunit (Karyopherin alpha
subunit)
 gi|7511453|pir||T42402  importin alpha 1 - Caenorhabditis elegans
 gi|2795929|gb|AAB97171.1|  (AF040995) importin alpha 3 [Caenorhabditis elegans]
 gi|4883510|gb|AAA83354.2|  (U41992) Hypothetical protein F32E10.4 [Caenorhabditis elegans]
          Length = 514

 Score = 33.9 bits (76), Expect = 0.37
 Identities = 27/87 (31%), Positives = 43/87 (49%), Gaps = 2/87 (2%)
 Frame = -2

Query: 257 VVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFD 78
           ++ + ILP LV  L   +   +  AA+ L  +A  +S    AVVN+GA+   +Q L   +
Sbjct: 104 LIGSGILPVLVQCLSSTDPNLQFEAAWALTNIASGTSEQTQAVVNAGAVPLFLQLLSCGN 163

Query: 77  PSVKEAAASALRYIAKQLQNI--YCLD 3
            +V E +  AL  I     +   YCL+
Sbjct: 164 LNVCEQSVWALGNIIGDGPHFRDYCLE 190
```

```
>gi|2959322|emb|CAA74966.1|  (Y14616) Importin alpha-like protein [Arabidopsis thaliana]
          Length = 528

 Score = 33.1 bits (74), Expect = 0.63
 Identities = 18/85 (21%), Positives = 39/85 (45%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           +G +    +   + ++ + +LP L +      +  KK A + +  +   +     AVV +
Sbjct: 306 VGNIVTGDDSQTQFIIESGVLPHLYNCDANHKKSIKKEACWTISNITAGNKLQIEAVVGA 365

Query: 119 GALEALVQCLEEFDPSVKEAAASAL 45
           G +  LV  L+  +  +K+ AA A+
Sbjct: 366 GIILPLVHLLQNAEFDIKKEAAWAI 390
```

```
>gi|15242560|ref|NP_195908.1|  (NM_120366) putative protein [Arabidopsis thaliana]
 gi|11357909|pir||T48309  hypothetical protein F9G14.190 - Arabidopsis thaliana
 gi|7413563|emb|CAB86042.1|  (AL162973) putative protein [Arabidopsis thaliana]
          Length = 1502

 Score = 32.3 bits (72), Expect = 1.1
 Identities = 23/82 (28%), Positives = 41/82 (49%)
 Frame = -2


Query: 263 EAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEE 84
           E +V ++ +P L   L      Y   A   L+ + K S    +A +N+GA+ A++  ++
Sbjct: 190 EFLVRHDTIPALCQRLLTIE--YLDVAEQCLQALEKISRDEPVACLNAGAIMAVLSFIDF 247

Query: 83  FDPSVKEAAASALRYIAKQLQN 18
           F  S++  A S +  I KQL +
Sbjct: 248 FSTSIQRVAISTVVNICKQLSS 269
```

```
>gi|5802657|gb|AAD51751.1|  (AF176651) pendulin [Oreochromis niloticus]
          Length = 523

 Score = 32.3 bits (72), Expect = 1.1
 Identities = 23/96 (23%), Positives = 43/96 (43%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           AL  +A  + +   AVV+   +P  +S +   ++   + A + L  +A   STL   V+
Sbjct: 138 ALTNIASGTSEQTAAVVTGGAIPAFISLVTSPHQHISEQAIWALGNIAGDGSTLRDKVIK 197

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNI 15
            G +  L+  L    P +   +A  L+ +   L N+
Sbjct: 198 HGTVAPLLSLLAV--PELSAFSAGYLKNVTWTLSNL 231



>gi|6320824|ref|NP_010903.1|  (NC_001137) An armadillo repeat-containing protein localized
on the vacuolar membrane; Vac8p [Saccharomyces cerevisiae]
 gi|731400|sp|P39968|VAC8_YEAST  Vacuolar protein 8
 gi|1077594|pir||S50446  VAC8 protein - yeast (Saccharomyces cerevisiae)
 gi|602380|gb|AAB64490.1|  (U18530) Yel013wp [Saccharomyces cerevisiae]
          Length = 578

 Score = 32.3 bits (72), Expect = 1.1
 Identities = 23/86 (26%), Positives = 37/86 (42%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGAL 111
           LA   ++  +   S  ++P    +  +  R  + A   +L     HS      +VN+GA+
Sbjct: 153 LATRDDNKHKIATSGALIPLTKLAKSKHIRVQRNATGALLNMT--HSEENRKELVNAGAV 210

Query: 110 EALVQCLEEFDPSVKEAAASALRYIA 33
             LV  L   DP V+    +AL  IA
Sbjct: 211 PVLVSLLSSTDPDVQYYCTTALSNIA 236



>gi|15242618|ref|NP_195927.1|  (NM_120385) importin alpha - like protein [Arabidopsis
thaliana]
 gi|11279266|pir||T48328  importin alpha-like protein - Arabidopsis thaliana
 gi|7413584|emb|CAB86074.1|  (AL163002) importin alpha-like protein [Arabidopsis thaliana]
          Length = 483

 Score = 32.3 bits (72), Expect = 1.1
 Identities = 24/92 (26%), Positives = 41/92 (44%), Gaps = 2/92 (2%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAK--HSSTLAMAV 129
           A+G +A   EDL   ++S   LP L   +        + AA+ L  + K   S   A  V
Sbjct: 191 AIGNVAGEGEDLRNVLLSQGALPPLARMIFPDKGSTVRTAAWALSNLIKGPESKAAAQLV 250

Query: 128 VNSGALEALVQCLEEFDPSVKEAAASALRYIA 33
              G L+A+++ L++ D       A  + Y++
Sbjct: 251 KIDGILDAILRHLKKTDEETATEIAWIIVYLS 282
```

```
>gi|7630132|emb|CAB71185.4|  (AL138518) importin alpha-like protein [Leishmania major]
          Length = 539

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 16/81 (19%), Positives = 38/81 (46%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNS 120
           +G +A  +++  + ++++  LP +   L    R  +K   + +  +A        A+VNS
Sbjct: 321 IGNIAAGTDEQTQVIINSGALPAMAELLRHPKRALRKETCWTISNIAAGQPYQIEALVNS 380

Query: 119 GALEALVQCLEEFDPSVKEAA 57
                +++CL   +  V++ A
Sbjct: 381 NVCIPILECLSAPELDVRKEA 401



>gi|16758242|ref|NP_445935.1|   (NM_053483) karyopherin (importin) alpha 2 [Rattus
norvegicus]
 gi|4456770|emb|CAB37408.1|   (AJ130946) importin alpha [Rattus norvegicus]
          Length = 529

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 43/87 (48%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG+ +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 104 LSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 163

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 164 IPAFISLLASPHAHISEQAVWALGNIA 190



>gi|1363205|pir||S57873  pendulin - mouse
 gi|555823|gb|AAA85281.1|   (U12270) pendulin [Mus musculus]
          Length = 529

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 43/87 (48%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG+ +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 104 LSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 163

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 164 IPAFISLLASPHAHISEQAVWALGNIA 190
```

```
>gi|6754474|ref|NP_034785.1| ☐ (NM_010655) karyopherin (importin) alpha 2; m-importin-
alpha-P1 [Mus musculus]
 gi|2829434|sp|P52293|IMA2_MOUSE  Importin alpha-2 subunit (Karyopherin alpha-2 subunit)
(SRP1-alpha) (RAG cohort protein 1) (Pendulin) (Pore targeting complex 58 kDa subunit)
(PTAC58) (Importin alpha P1)
 gi|1363192|pir||S57345  m-importin (nuclear pore-targeting complex component 58K) - mouse
 gi|893393|dbj|BAA09536.1| ☐ (D55720) nuclear pore-targeting complex component of 58 kDa
[Mus musculus]
 gi|998346|gb|AAC52451.1| ☐ (U34229) pendulin [Mus musculus]
 gi|13096958|gb|AAH03274.1|AAH03274 ☐ (BC003274) karyopherin (importin) alpha 2 [Mus
musculus]
 gi|13879484|gb|AAH06720.1|AAH06720 ☐ (BC006720) karyopherin (importin) alpha 2 [Mus
musculus]
 gi|1589534|prf||2211316B  pendulin [Mus musculus]
          Length = 529

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 43/87 (48%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG+ +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 104 LSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 163

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 164 IPAFISLLASPHAHISEQAVWALGNIA 190



>gi|17942565|pdb|1IQ1|C  Chain C, Crystal Structure Of The Importin-Alpha(44-54)-Importin-
Alpha(70-529) Complex
 gi|7766973|pdb|1EJL|I  Chain I, Mouse Importin Alpha-Sv40 Large T Antigen Nls Peptide
Complex
 gi|7766975|pdb|1EJY|I  Chain I, Mouse Importin Alpha-Nucleoplasmin Nls Peptide Complex
          Length = 460

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 43/87 (48%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG+ +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 35  LSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 94

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 95  IPAFISLLASPHAHISEQAVWALGNIA 121



>gi|15638984|gb|AAG42103.2|  (AY010065) karyopherin alpha 2 [Sus scrofa]
          Length = 240

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 42/87 (47%), Gaps = 1/87 (1%)
 Frame = -2
```

```
Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG  +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 96  LSREKQPPIDNIIRAGLIPKFVSFLGRMDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 155

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 156 IPAFISLLASPHAHISEQAVWALGNIA 182



>gi|5542272|pdb|1IAL|A  Chain A, Importin Alpha, Mouse
          Length = 453

 Score = 32.0 bits (71), Expect = 1.4
 Identities = 23/87 (26%), Positives = 43/87 (48%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG+ +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 61  LSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 120

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 121 IPAFISLLASPHAHISEQAVWALGNIA 147



>gi|4505841|ref|NP_000290.1|   (NM_000299) plakophilin 1; Plakophilin-1 [Homo sapiens]
 gi|550115|emb|CAA84426.1|   (Z34974) plakophilin [Homo sapiens]
 gi|1770488|emb|CAA98022.1|   (Z73678) plakophilin 1 [Homo sapiens]
          Length = 726

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 22/79 (27%), Positives = 36/79 (44%)
 Frame = -2

Query: 278 SEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALV 99
           SED+     S   +P+ V  L  Q+  Y+   A+ ++       +    V   G +  LV
Sbjct: 237 SEDIE---CSGLTIPKAVQYLSSQDEKYQAIGAYYIQHTCFQDESAKQQVYQLGGICKLV 293

Query: 98  QCLEEFDPSVKEAAASALR 42
             L   + +V++AAA ALR
Sbjct: 294 DLLRSPNQNVQQAAAGALR 312



>gi|17465256|ref|XP_069697.1|   (XM_069697) similar to IMPORTIN ALPHA-2 SUBUNIT
(KARYOPHERIN ALPHA-2 SUBUNIT) [Homo sapiens]
          Length = 553

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 25/89 (28%), Positives = 45/89 (50%), Gaps = 3/89 (3%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRFY---KKAAAFVLRCVAKHSSTLAMAVVNS 120
           L++      + V+   ++P++V  L  ++  Y   +  AA+ L  +A  +S    AVV
Sbjct: 98  LSQEKNPPLKLVIEAGLIPRMVEFL--KSSLYPCLQFEAAWALTNIASGTSEQTRAVVEG 155
```

```
Query: 119 GALEALVQCLEEFDPSVKEAAASALRYIA 33
           GA++ L++ L   + +V E A  AL  IA
Sbjct: 156 GAIQPLIELLSSSNVAVCEQAVWALGNIA 184



>gi|20138951|sp|Q13835|PKP1_HUMAN  Plakophilin 1 (Band-6-protein) (B6P)
          Length = 747

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 22/79 (27%), Positives = 36/79 (44%)
 Frame = -2

Query: 278 SEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALV 99
           SED+     S   +P+ V  L  Q+  Y+   A+ ++       +    V   G +  LV
Sbjct: 237 SEDIE---CSGLTIPKAVQYLSSQDEKYQAIGAYYIQHTCFQDESAKQQVYQLGGICKLV 293

Query: 98  QCLEEFDPSVKEAAASALR 42
             L   + +V++AAA ALR
Sbjct: 294 DLLRSPNQNVQQAAAGALR 312



>gi|17552816|ref|NP_499306.1|  (NM_066905) D2045.2.p [Caenorhabditis elegans]
 gi|7498209|pir||T20363  hypothetical protein D2045.2 - Caenorhabditis elegans
 gi|3875359|emb|CAA84693.1|  (Z35639) cDNA EST EMBL:T00176 comes from this gene~cDNA EST
yk31g3.5 comes from this gene~cDNA EST yk31g3.3 comes from this gene~cDNA EST yk115a11.3
comes from this gene~cDNA EST yk115a11.5 comes from this gene~cDNA EST yk54b12.3 comes from
this gene~c>
          Length = 1792

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 20/74 (27%), Positives = 40/74 (54%)
 Frame = -2

Query: 251  SNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPS 72
             +NEI  +L+ +L ++    +++A   L  + +   T+ M  +    LEA+++  ++   S
Sbjct: 1116 ANEIAKELLPALTDREYRVRESACLALSDLLRGHDTVEMHKMIPEYLEAILRVRDDVKES 1175

Query: 71   VKEAAASALRYIAK 30
             V+EAA  A   I+K
Sbjct: 1176 VREAADRAADSISK 1189



>gi|2134803|pir||S60712  band-6-protein - human
 gi|535015|emb|CAA55881.1|  (X79293) band-6-protein [Homo sapiens]
          Length = 725

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 22/79 (27%), Positives = 36/79 (44%)
 Frame = -2

Query: 278 SEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALV 99
           SED+     S   +P+ V  L  Q+  Y+   A+ ++       +    V   G +  LV
Sbjct: 236 SEDIE---CSGLTIPKAVQYLSSQDEKYQAIGAYYIQHTCFQDESAKQQVYQLGGICKLV 292
```

```
Query: 98  QCLEEFDPSVKEAAASALR 42
              L   + +V++AAA ALR
Sbjct: 293 DLLRSPNQNVQQAAAGALR 311




>gi|4504897|ref|NP_002257.1| ☐ (NM_002266) karyopherin alpha 2; RAG cohort 1; importin
alpha 1 [Homo sapiens]
 gi|1708480|sp|P52292|IMA2_HUMAN  Importin alpha-2 subunit (Karyopherin alpha-2 subunit)
(SRP1-alpha) (RAG cohort protein 1)
 gi|2135835|pir||A56516  nuclear localization sequence receptor SRP1 alpha - human
 gi|791185|gb|AAA65700.1| ☐ (U09559) Rch1 [Homo sapiens]
 gi|899539|gb|AAA69957.1| ☐ (U28386) hSRP1alpha [Homo sapiens]
 gi|15865453|emb|CAC83080.1|  (AJ303086) karyopherin alpha 2 [Homo sapiens]
          Length = 529

 Score = 31.6 bits (70), Expect = 1.8
 Identities = 23/87 (26%), Positives = 42/87 (47%), Gaps = 1/87 (1%)
 Frame = -2

Query: 290 LAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVVNSGA 114
           L++  +   + ++   ++P+ VS LG  +    +  +A+ L  +A  +S    AVV+ GA
Sbjct: 104 LSREKQPPIDNIIRAGLIPKFVSFLGRTDCSPIQFESAWALTNIASGTSEQTKAVVDGGA 163

Query: 113 LEALVQCLEEFDPSVKEAAASALRYIA 33
           + A +  L      + E A  AL  IA
Sbjct: 164 IPAFISLLASPHAHISEQAVWALGNIA 190




>gi|15238931|ref|NP_199049.1|  (NM_123599) arm repeat containing protein [Arabidopsis
thaliana]
 gi|9759470|dbj|BAB10475.1|  (AB016888) arm repeat containing protein [Arabidopsis thaliana]
          Length = 656

 Score = 30.4 bits (67), Expect = 4.1
 Identities = 19/60 (31%), Positives = 31/60 (51%)
 Frame = -2

Query: 224 SSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKEAAASAL 45
           S L EQ R  K+     +R +A+      + + N+GA+  LVQ L   D  ++E A + L
Sbjct: 387 SQLEEQRRSVKQ-----MRLLARENPENRVLIANAGAIPLLVQLLSYPDSGIQENAVTTL 441




>gi|9790161|ref|NP_062619.1| ☐ (NM_019645) plakophilin 1 [Mus musculus]
 gi|20138946|sp|P97350|PKP1_MOUSE  Plakophilin 1
 gi|1707594|emb|CAA69240.1| ☐ (Y07941) plakophilin 1 [Mus musculus]
          Length = 728

 Score = 30.4 bits (67), Expect = 4.1
 Identities = 22/79 (27%), Positives = 36/79 (44%)
 Frame = -2

Query: 278 SEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALV 99
           SED+     S   +P+ V  L  Q+  Y+   A+ ++       +    V   G +  LV
Sbjct: 238 SEDIE---CSGLTIPKAVQYLCSQDEKYQAIGAYYIQHTCFQDESAKQQVYQLGGICKLV 294
```

```
Query: 98  QCLEEFDPSVKEAAASALR 42
             L   + +V++AAA ALR
Sbjct: 295 DLLRSPNQNVQQAAAGALR 313



>gi|7288998|gb|AAF45332.1|  (AE003091) CG17514 gene product [Drosophila melanogaster]
          Length = 2225

 Score = 30.0 bits (66), Expect = 5.3
 Identities = 32/113 (28%), Positives = 49/113 (43%), Gaps = 20/113 (17%)
 Frame = -2

Query: 287  AKHSEDLAEAVVS---NEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSS---------- 147
             AK  E L   V S   +EILP ++  L + + F  +     LR V    S
Sbjct: 1579 AKTFESLHSTVGSRALDEILPFMLQGLSDADPFVAENTLDGLRQVMSIKSKVVLPYLVPQ 1638

Query: 146  -------TLAMAVVNSGALEALVQCLEEFDPSVKEAAASALRYIAKQLQNIYC 9
                    T A++++ S A EAL++ L +   S+ EA + A  Y  +  +N YC
Sbjct: 1639 LTSPPVNTKALSILVSVAGEALIKYLPKILSSLLEALSDAYGYPNEPQENEYC 1691



>gi|6822222|emb|CAB70928.1|  (AL137778) putative acyltransferase [Streptomyces coelicolor
A3(2)]
          Length = 311

 Score = 30.0 bits (66), Expect = 5.3
 Identities = 11/21 (52%), Positives = 15/21 (71%)
 Frame = -3

Query: 301 HWEDWLSIQKIWLKQLFPMKS 239
           H EDW  +Q++WLK L P K+
Sbjct: 280 HPEDWHMLQRLWLKDLDPAKA 300



>gi|4680191|gb|AAD27556.1|AF111710_2  (AF111710) putative myb-related protein [Oryza sativa
subsp. indica] [Oryza sativa (indica cultivar-group)]
          Length = 1297

 Score = 30.0 bits (66), Expect = 5.3
 Identities = 16/44 (36%), Positives = 24/44 (54%)
 Frame = +1

Query: 43  RRALAAASLTLGSNSSKHQTKASKAPEFTTAIAKVEECLATHLS 174
           R  +  A + L S +S H+ +AS+  E TTA  K +  + TH S
Sbjct: 304 RHDICIAGVLLFSCNSSHRRRASRTGEKTTATQKRKRSIQTHAS 347



>gi|18017587|ref|NP_542121.1|  (NM_080554) proteasome (prosome, macropain) 26S subunit,
non-ATPase, 5 [Mus musculus]
 gi|17512272|gb|AAH19112.1|AAH19112  (BC019112) Similar to proteasome (prosome,
macropain) 26S subunit, non-ATPase, 5 [Mus musculus]
```

```
          Length = 504

 Score = 30.0 bits (66), Expect = 5.3
 Identities = 22/76 (28%), Positives = 42/76 (54%), Gaps = 2/76 (2%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLA--MAVV 126
           +GR+ ++SE + E + + E+L Q+V  +G +N    KAA   ++ +++ S T A   A+
Sbjct: 107 IGRIVENSEAVTEILNNAELLKQIVYCIGGENLSVAKAA---IKSLSRISLTQAGLEALF 163

Query: 125 NSGALEALVQCLEEFD 78
            S  L+ L   ++  D
Sbjct: 164 ESNLLDDLKNVMKTND 179




>gi|4325344|gb|AAD17343.1|  (AF128393) similar to beta-transducins (Pfam: PF00400,
Score=71.7, E=1.5e-17, N=6) [Arabidopsis thaliana]
 gi|7267253|emb|CAB81036.1|  (AL161502) putative WD-repeat membrane protein [Arabidopsis
thaliana]
          Length = 931

 Score = 29.6 bits (65), Expect = 6.9
 Identities = 21/58 (36%), Positives = 28/58 (48%)
 Frame = -1

Query: 291 IGQAFRRFG*SSCFQ*NLAIISEFVGRTKSFLQKGSSIRTQMRCQTFFNFGNGSCEFR 118
           +G+AF+ +   +C + NL IIS+F    K  L K        R  TF  FGN    FR
Sbjct: 34  VGKAFQIY---NCAKLNLVIISKFFLLLKPKLPKKIRALASYRDYTFVAFGNEIAVFR 88



>gi|15223673|ref|NP_175502.1|  (NM_103969) hypothetical protein [Arabidopsis thaliana]
 gi|12321785|gb|AAG50927.1|AC079284_2  (AC079284) hypothetical protein [Arabidopsis
thaliana]
          Length = 788

 Score = 29.6 bits (65), Expect = 6.9
 Identities = 24/66 (36%), Positives = 34/66 (51%)
 Frame = -2

Query: 242 ILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVVNSGALEALVQCLEEFDPSVKE 63
           +L  +   LG  N   +KAAA VL  +A HSS+L +A      L AL  C  +    V+E
Sbjct: 217 LLHSIHECLGCTNWVTRKAAADVLISLAVHSSSL-VADKTDSTLTALEACRFDKIKPVRE 275

Query: 62  AAASAL 45
           + + AL
Sbjct: 276 SLSEAL 281




>gi|3299917|gb|AAC25988.1|  (AF047839) putative glucosyl hydrolase precursor
[Pseudoalteromonas sp. S9]
          Length = 809

 Score = 29.6 bits (65), Expect = 6.9
 Identities = 15/44 (34%), Positives = 26/44 (59%)
 Frame = +1
```

```
Query: 25  NYLAMQRRALAAASLTLGSNSSKHQTKASKAPEFTTAIAKVEEC 156
           +YLAM++R LA  +   LG+N     + +A     E+ T   A+++ C
Sbjct: 597 SYLAMEKRQLAQMAKRLGNNQDAQKYQA----EYQTLKARIDAC 636




>gi|1401236|gb|AAB03389.1|  (U59448) receptor tyrosine kinase 90 [Hydra vulgaris]
          Length = 848

 Score = 29.6 bits (65), Expect = 6.9
 Identities = 19/50 (38%), Positives = 25/50 (50%), Gaps = 5/50 (10%)
 Frame = +3

Query: 6   QTVYVLQLLGNVTESTCC-----SFLDTWIEFLQTLDQSFQGS*IHNCHC 140
           Q  YV++LLG V ES  C      S L    E+LQT    +    + +H C C
Sbjct: 647 QNSYVVRLLGLVRESPFCIITEYSELGDLKEYLQTNKNTSSSARLHMCTC 696




>gi|15219352|ref|NP_175078.1|  (NM_103538) hypothetical protein [Arabidopsis thaliana]
 gi|12320824|gb|AAG50555.1|AC074228_10  (AC074228) hypothetical protein [Arabidopsis
thaliana]
          Length = 2114

 Score = 29.3 bits (64), Expect = 9.0
 Identities = 25/91 (27%), Positives = 44/91 (47%), Gaps = 1/91 (1%)
 Frame = -2

Query: 299 LGRLAKHSEDLAEAVVSNEILPQLVSSLGEQ-NRFYKKAAAFVLRCVAKHSSTLAMAVVN 123
           L   L  + E+    AV S    +P L+   L     ++   K  A   V+   +   HS   + + V
Sbjct: 471 LAILTDNVEESRWAVTSAGGIPPLLQILETGVSQKAKDDAVRVILNLCCHSEEIRLCVEK 530

Query: 122 SGALEALVQCLEEFDPSVKEAAASALRYIAK 30
           +GA+ AL+   L+    P  +E++A+ L  + K
Sbjct: 531 AGAIPALLGLLKNGGPKSQESSANTLLKLIK 561




>gi|15242247|ref|NP_200013.1|  (NM_124579) importin alpha subunit [Arabidopsis thaliana]
 gi|10177735|dbj|BAB11048.1|  (AB015478) importin alpha subunit [Arabidopsis thaliana]
          Length = 441

 Score = 29.3 bits (64), Expect = 9.0
 Identities = 24/102 (23%), Positives = 46/102 (44%), Gaps = 5/102 (4%)
 Frame = -2

Query: 305 LALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRFYKKAAAFVLRCVAKHSSTLAMAVV 126
           +AL   L++ SED   ++V++     +P+LV   L    +        A    +   +    +        V+
Sbjct: 192 MALCHLSEGSEDGIQSVIEAGFVPKLVQILQLPSPVVLVPALLTIGAMTAGNHQQTQCVI 251

Query: 125 NSGALEALVQCL-EEFDPSVKEAAASALRYIA----KQLQNI 15
           NSGAL   +    L    +   +K+ A   +   I      +Q+Q++
Sbjct: 252 NSGALPIISNMLTRNHENKIKKCACWVISNITAGTKEQIQSV 293
```

```
>gi|4504903|ref|NP_002260.1| ☐ (NM_002269) karyopherin alpha 5 (importin alpha 6);
importin alpha 6 [Homo sapiens]
 gi|3122273|sp|O15131|IMA5_HUMAN  Importin alpha-6 subunit (Karyopherin alpha-5 subunit)
 gi|2343116|gb|AAC51868.1| ☐ (AF005361) importin alpha 6 [Homo sapiens]
          Length = 536

 Score = 29.3 bits (64), Expect = 9.0
 Identities = 25/91 (27%), Positives = 40/91 (43%), Gaps = 1/91 (1%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVV 126
           ALG +A  + +  + V++ EILP L+  L   NR    + A + L  + +  +
Sbjct: 190 ALGNIAGDNAECRDFVLNCEILPPLLELLTNSNRLTTTRNAVWALSNLCRGKNPPPNFSK 249

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIA 33
            S  L  L + L   DP V      AL Y++
Sbjct: 250 VSPCLNVLSRLLFSSDPDVLADVCWALSYLS 280




>gi|11418581|ref|XP_004308.1| ☐ (XM_004308) karyopherin alpha 5 (importin alpha 6) [Homo
sapiens]
 gi|12314096|emb|CAC00772.1|  (AL132795) dJ412I7.2 (karyopherin alpha 5 (importin alpha 6))
[Homo sapiens]
          Length = 536

 Score = 29.3 bits (64), Expect = 9.0
 Identities = 25/91 (27%), Positives = 40/91 (43%), Gaps = 1/91 (1%)
 Frame = -2

Query: 302 ALGRLAKHSEDLAEAVVSNEILPQLVSSLGEQNRF-YKKAAAFVLRCVAKHSSTLAMAVV 126
           ALG +A  + +  + V++ EILP L+  L   NR    + A + L  + +  +
Sbjct: 190 ALGNIAGDNAECRDFVLNCEILPPLLELLTNSNRLTTTRNAVWALSNLCRGKNPPPNFSK 249

Query: 125 NSGALEALVQCLEEFDPSVKEAAASALRYIA 33
            S  L  L + L   DP V      AL Y++
Sbjct: 250 VSPCLNVLSRLLFSSDPDVLADVCWALSYLS 280




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.312    0.120    0.326

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 162,396,508
```

```
Number of Sequences: 919285
Number of extensions: 2853173
Number of successful extensions: 8030
Number of sequences better than 10.0: 98
Number of HSP's better than 10.0 without gapping: 7804
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 8025
length of database: 288,558,979
effective HSP length: 77
effective length of database: 217,774,034
effective search space used: 5226576816
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.2 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (22.0 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020954194-021295-4666
```

7.1.10 Query= **hy-10_15b145025**

```
         (411 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 30 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|15837445|ref|NP_298133.1|  (NC_002488) conserved hypothet...    94   3e-19
gi|16799833|ref|NP_470101.1|  (NC_003212) lin0759 [Listeria ...    84   3e-16
gi|16802807|ref|NP_464292.1|  (NC_003210) lmo0765 [Listeria ...    84   3e-16
gi|15673481|ref|NP_267655.1|  (NC_002662) conserved hypothet...    82   1e-15
gi|7480237|pir||T37125  hypothetical protein SCJ4.42c - Stre...    80   3e-15
gi|15901957|ref|NP_346561.1|  (NC_003028) conserved hypothet...    78   1e-14
gi|15903993|ref|NP_359543.1|  (NC_003098) Conserved hypothet...    78   1e-14
gi|19115560|ref|NP_594648.1|  (NC_003424) hypothetical prote...    77   4e-14
gi|18309408|ref|NP_561342.1|  (NC_003366) conserved hypothet...    76   5e-14
gi|16801187|ref|NP_471455.1|  (NC_003212) similar to unknown...    75   1e-13
gi|16804052|ref|NP_465537.1|  (NC_003210) similar to unknown...    74   3e-13
gi|19746530|ref|NP_607666.1|  (NC_003485) conserved hypothet...    71   2e-12
gi|15675485|ref|NP_269659.1|  (NC_002737) conserved hypothet...    71   2e-12
gi|15613353|ref|NP_241656.1|  (NC_002570) BH0790~unknown con...    70   5e-12
gi|15897551|ref|NP_342156.1|  (NC_002754) Argininosuccinate ...    32   1.2
gi|6320562|ref|NP_010643.1|  (NC_001136) may be involved in ...    31   2.6
gi|16804984|ref|NP_473013.1|  (NC_000910) protein of the YMR...    31   2.6
gi|6322974|ref|NP_013046.1|  (NC_001144) Hypothetical ORF; Y...    31   2.6
gi|1083855|pir||S41309  cyclosporin synthetase - cyclosporin...    31   2.6
gi|14719841|gb|AAK73196.1|  (AF320327) putative ruminococcin...    30   3.5
gi|1085199|pir||S46260  creatine transporter - marbled elect...    30   3.5
gi|15237573|ref|NP_198932.1|  (NM_123481) purine permease-li...    30   4.5
gi|11132043|sp|O14653|GS27_HUMAN  27 kDa Golgi SNARE protein...    30   5.9
```

```
gi|16905522|ref|NP_004278.2|  (NM_004287) golgi SNAP recepto...   30   5.9
gi|19114410|ref|NP_593498.1|  (NC_003424) putative glycosylp...   30   5.9
gi|17564302|ref|NP_504743.1|  (NM_072342) T15B7.1.p [Caenorh...   30   5.9
gi|16905520|ref|NP_473363.1|  (NM_054022) golgi SNAP recepto...   30   5.9
gi|15896721|ref|NP_350070.1|  (NC_003030) Short-chain alcoho...   29   7.7
gi|10954561|ref|NP_044683.1|  (NC_001811) unknown [Methanoco...   29   7.7
gi|15895879|ref|NP_349228.1|  (NC_003030) Cell wall hydrolas...   29   7.7
```

Alignments

```
>gi|15837445|ref|NP_298133.1|  (NC_002488) conserved hypothetical protein [Xylella
fastidiosa 9a5c]
 gi|11292265|pir||H82756  conserved hypothetical protein XF0843 [imported] - Xylella
fastidiosa (strain 9a5c)
 gi|9105747|gb|AAF83653.1|AE003923_11  (AE003923) conserved hypothetical protein [Xylella
fastidiosa 9a5c]
          Length = 516

 Score = 93.6 bits (231), Expect = 3e-19
 Identities = 40/78 (51%), Positives = 57/78 (72%)
 Frame = +3

Query: 9   IWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGLF 188
           IWPM+++   L+++D ++I+ CL  L TT  G  FMHE+F +D+PS+FTRSWFAWAN LF
Sbjct: 431 IWPMSLIHYALSSHDDTQIQHCLHWLKTTHAGTGFMHEAFHQDDPSTFTRSWFAWANSLF 490

Query: 189 GELVLKIDKEYPELLEEN 242
           GEL++ + +  P LL  N
Sbjct: 491 GELIVDLAQRKPHLLGVN 508
```

```
>gi|16799833|ref|NP_470101.1|  (NC_003212) lin0759 [Listeria innocua]
 gi|16413210|emb|CAC95991.1|  (AL596166) lin0759 [Listeria innocua]
          Length = 432

 Score = 84.0 bits (206), Expect = 3e-16
 Identities = 37/68 (54%), Positives = 46/68 (67%)
 Frame = +3

Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           YIWP+ +  QGLT ND SE  E L+ML+ T  G   MHESF  D+P  FTR WF+WAN +
Sbjct: 354 YIWPIGLAIQGLTANDTSEKLEILQMLLRTDAGTGLMHESFHPDHPEDFTREWFSWANMM 413

Query: 186 FGELVLKI 209
           F EL+L +
Sbjct: 414 FCELILDV 421
```

```
>gi|16802807|ref|NP_464292.1|  (NC_003210) lmo0765 [Listeria monocytogenes EGD-e]
 gi|16410154|emb|CAC98843.1|  (AL591976) lmo0765 [Listeria monocytogenes]
          Length = 432

 Score = 83.6 bits (205), Expect = 3e-16
 Identities = 38/68 (55%), Positives = 46/68 (66%)
 Frame = +3
```

```
Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           YIWP+ +  QGLT +DKSE  E L+ML+ T  G   MHESF  D P  FTR WF+WAN +
Sbjct: 354 YIWPIGLAIQGLTASDKSEKLEILQMLLRTDAGTGLMHESFHPDYPEDFTREWFSWANMM 413

Query: 186 FGELVLKI 209
           F ELVL +
Sbjct: 414 FCELVLDV 421




>gi|15673481|ref|NP_267655.1|  (NC_002662) conserved hypothetical protein [Lactococcus
lactis subsp. lactis]
 gi|12724495|gb|AAK05597.1|AE006380_5  (AE006380) conserved hypothetical protein
[Lactococcus lactis subsp. lactis]
          Length = 431

 Score = 82.0 bits (201), Expect = 1e-15
 Identities = 35/67 (52%), Positives = 46/67 (68%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           NY+WP+A+  +G+TTNDKSE    L  LV T  G   MHE F  +NP ++TR WF+WAN
Sbjct: 356 NYVWPIALAMEGMTTNDKSEKERILNHLVETDAGTHLMHEGFDVNNPQNYTREWFSWANM 415

Query: 183 LFGELVL 203
           +F ELV+
Sbjct: 416 MFCELVM 422




>gi|7480237|pir||T37125  hypothetical protein SCJ4.42c - Streptomyces coelicolor
 gi|5738802|emb|CAB52976.1|  (AL109950) hypothetical protein [Streptomyces coelicolor A3(2)]
          Length = 434

 Score = 80.5 bits (197), Expect = 3e-15
 Identities = 36/66 (54%), Positives = 45/66 (67%)
 Frame = +3

Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           +IWP+AI  QGLT+ND +E R  L  L+ T  G   MHESF KD+P  FTR WF+WAN +
Sbjct: 354 HIWPIAIAVQGLTSNDTTERRWALTTLLATDAGTGQMHESFHKDDPHRFTRPWFSWANAM 413

Query: 186 FGELVL 203
           + EL L
Sbjct: 414 YAELAL 419




>gi|15901957|ref|NP_346561.1|  (NC_003028) conserved hypothetical protein [Streptococcus
pneumoniae TIGR4]
 gi|14973657|gb|AAK76201.1|  (AE007503) conserved hypothetical protein [Streptococcus
pneumoniae TIGR4]
          Length = 426

 Score = 78.2 bits (191), Expect = 1e-14
 Identities = 36/66 (54%), Positives = 46/66 (69%)
 Frame = +3
```

```
Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           YIWP+A+  QGLTT DK+E +  L  LV   GG   MHESF  D+P+ ++R WF+WAN +
Sbjct: 355 YIWPIALSIQGLTTRDKAEKKFLLDQLVACDGGTGVMHESFHVDDPTLYSREWFSWANMM 414

Query: 186 FGELVL 203
           F ELVL
Sbjct: 415 FCELVL 420



>gi|15903993|ref|NP_359543.1|  (NC_003098) Conserved hypothetical protein [Streptococcus
pneumoniae R6]
 gi|15459651|gb|AAL00754.1|  (AE008559) Conserved hypothetical protein [Streptococcus
pneumoniae R6]
          Length = 434

 Score = 78.2 bits (191), Expect = 1e-14
 Identities = 36/66 (54%), Positives = 46/66 (69%)
 Frame = +3

Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           YIWP+A+  QGLTT DK+E +  L  LV   GG   MHESF  D+P+ ++R WF+WAN +
Sbjct: 363 YIWPIALSIQGLTTRDKAEKKFLLDQLVACDGGTGVMHESFHVDDPTLYSREWFSWANMM 422

Query: 186 FGELVL 203
           F ELVL
Sbjct: 423 FCELVL 428



>gi|19115560|ref|NP_594648.1|  (NC_003424) hypothetical protein [Schizosaccharomyces pombe]
 gi|1723558|sp|Q10449|YDEG_SCHPO  Hypothetical protein C12B10.16c in chromosome I
 gi|7490785|pir||T37583  hypothetical protein SPAC12B10.16c - fission yeast
(Schizosaccharomyces pombe)
 gi|1262429|emb|CAA94706.1|  (Z70721) hypothetical protein [Schizosaccharomyces pombe]
          Length = 509

 Score = 76.6 bits (187), Expect = 4e-14
 Identities = 35/78 (44%), Positives = 50/78 (63%)
 Frame = +3

Query: 9   IWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGLF 188
           +WPM+++ Q LT++D  EI   L +L  +T G   MHES    +  SFTR WF+WAN LF
Sbjct: 430 VWPMSLIVQALTSDDDDEIMSLLDVLKHSTAGLGLMHESVDVSSFKSFTRPWFSWANSLF 489

Query: 189 GELVLKIDKEYPELLEEN 242
            EL+L + +  P LL++N
Sbjct: 490 AELILDLLERKPHLLKKN 507



>gi|18309408|ref|NP_561342.1|  (NC_003366) conserved hypothetical protein [Clostridium
perfringens]
 gi|18144084|dbj|BAB80132.1|  (AP003186) conserved hypothetical protein [Clostridium
perfringens]
          Length = 427

 Score = 76.3 bits (186), Expect = 5e-14
```

```
 Identities = 31/65 (47%), Positives = 44/65 (67%)
 Frame = +3

Query: 6   YIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANGL 185
           YIW +A+  QGLTTN++ EI + +++L  T  G  +MHE F  D+P+ FTR WFAW+N L
Sbjct: 355 YIWHIALSMQGLTTNNQEEIDQLIKLLKETDAGTGYMHEGFHVDDPTKFTRDWFAWSNSL 414

Query: 186 FGELV 200
           F   +
Sbjct: 415 FSHFI 419



>gi|16801187|ref|NP_471455.1|  (NC_003212) similar to unknown proteins [Listeria innocua]
 gi|16414635|emb|CAC97351.1|  (AL596171) similar to unknown proteins [Listeria innocua]
          Length = 434

 Score = 75.1 bits (183), Expect = 1e-13
 Identities = 31/69 (44%), Positives = 44/69 (62%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           +Y+WP+A+  +GLT   +SE +  L ML+   GG  +MHE     NP+ FTR WFAW+N
Sbjct: 354 HYVWPIALSIEGLTAETESEKQAILEMLIAGDGGTDYMHEGVNASNPAEFTRDWFAWSNA 413

Query: 183 LFGELVLKI 209
           +F E VL +
Sbjct: 414 MFSEFVLSL 422



>gi|16804052|ref|NP_465537.1|  (NC_003210) similar to unknown proteins [Listeria
monocytogenes EGD-e]
 gi|16411466|emb|CAD00091.1|  (AL591981) similar to unknown proteins [Listeria
monocytogenes]
          Length = 434

 Score = 73.9 bits (180), Expect = 3e-13
 Identities = 30/69 (43%), Positives = 44/69 (63%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           +Y+WP+A+  +GLT   ++E +  L ML+   GG  +MHE     NP+ FTR WFAW+N
Sbjct: 354 HYVWPIALSIEGLTAETEAEKQAILEMLIAGDGGTDYMHEGVNASNPAEFTRDWFAWSNA 413

Query: 183 LFGELVLKI 209
           +F E VL +
Sbjct: 414 MFSEFVLSL 422



>gi|19746530|ref|NP_607666.1|  (NC_003485) conserved hypothetical protein [Streptococcus
pyogenes MGAS8232]
 gi|19748739|gb|AAL98165.1|  (AE010073) conserved hypothetical protein [Streptococcus
pyogenes MGAS8232]
          Length = 437

 Score = 70.9 bits (172), Expect = 2e-12
```

```
 Identities = 32/67 (47%), Positives = 43/67 (63%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           NYIW +A+  QGLT  D+   +E L +LV T  G   MHE F  ++P  +TR WF+WAN
Sbjct: 362 NYIWHIALALQGLTALDQDSKKEMLDLLVATDAGTHLMHEGFDVNDPYQYTREWFSWANM 421

Query: 183 LFGELVL 203
           +F EL+L
Sbjct: 422 MFCELLL 428



>gi|15675485|ref|NP_269659.1|  (NC_002737) conserved hypothetical protein [Streptococcus
pyogenes] [Streptococcus pyogenes M1 GAS]
 gi|13622679|gb|AAK34380.1|  (AE006591) conserved hypothetical protein [Streptococcus
pyogenes M1 GAS]
          Length = 437

 Score = 70.9 bits (172), Expect = 2e-12
 Identities = 32/67 (47%), Positives = 43/67 (63%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           NYIW +A+  QGLT  D+   +E L +LV T  G   MHE F  ++P  +TR WF+WAN
Sbjct: 362 NYIWHIALALQGLTALDQDSKKEMLDLLVATDAGTHLMHEGFDVNDPYQYTREWFSWANM 421

Query: 183 LFGELVL 203
           +F EL+L
Sbjct: 422 MFCELLL 428



>gi|15613353|ref|NP_241656.1|  (NC_002570) BH0790~unknown conserved protein in others
[Bacillus halodurans]
 gi|10173404|dbj|BAB04509.1|  (AP001509) BH0790~unknown conserved protein in others
[Bacillus halodurans]
          Length = 454

 Score = 69.7 bits (169), Expect = 5e-12
 Identities = 28/67 (41%), Positives = 41/67 (60%)
 Frame = +3

Query: 3   NYIWPMAIVTQGLTTNDKSEIRECLRMLVTTTGGKWFMHESFFKDNPSSFTRSWFAWANG 182
           +Y+W +++   QG+T  D  E ++ + M   T    +FMHE F  D P  +TRSWFAWAN
Sbjct: 352 HYVWHISLAIQGMTAIDSKEKKQIVAMFKQTHADTYFMHEGFDVDRPEQYTRSWFAWANS 411

Query: 183 LFGELVL 203
           +F E +L
Sbjct: 412 MFSEFLL 418



>gi|15897551|ref|NP_342156.1|  (NC_002754) Argininosuccinate synthetase (argG)
[Sulfolobussolfataricus]
 gi|6015836|emb|CAB57663.1|  (Y18930) argininosuccinate synthase [Sulfolobus solfataricus]
 gi|13813806|gb|AAK40946.1|  (AE006692) Argininosuccinate synthetase (argG) [Sulfolobus
solfataricus]
```

```
          Length = 406

 Score = 32.0 bits (71), Expect = 1.2
 Identities = 18/51 (35%), Positives = 22/51 (42%)
 Frame = +2

Query: 89  DNHWRQMVYARKLFQGQSFQFHKILVCMGKWIVWRISVEDRQGIPRIIRRE 241
           D  W  +VY    F+      HK+   M KWI     VE   G  RI+ RE
Sbjct: 312 DQLWSDLVYQGLWFEPLRETLHKVADEMNKWISGEAKVEVSNGSFRIVGRE 362




>gi|6320562|ref|NP_010643.1|  (NC_001136) may be involved in connecting nuclear microtubules
to the spindle pole body; Spc110p [Saccharomyces cerevisiae]
 gi|417392|sp|P32380|NUF1_YEAST  NUF1 protein (Spindle poly body spacer protein SPC110)
 gi|283221|pir||S26710  spindle pole body protein NUF1 - yeast (Saccharomyces cerevisiae)
 gi|4070|emb|CAA77668.1|  (Z11582) nuf1 [Saccharomyces cerevisiae]
 gi|312175|emb|CAA51733.1|  (X73297) SPC110/NUF1 [Saccharomyces cerevisiae]
 gi|849173|gb|AAB64791.1|  (U28372) Nuf1p: Probable essential component of the
nucleoskeleton (Swiss Prot. accession number P32380) [Saccharomyces cerevisiae]
          Length = 944

 Score = 30.8 bits (68), Expect = 2.6
 Identities = 18/59 (30%), Positives = 32/59 (53%)
 Frame = -1

Query: 336 NQAREDGCQSVYEIKSIFTQIALKYYYIIKHNSLLIILGIPYLSSTLILQTIHLPMQTK 160
           N+ R+D   S  ++  I +   LKY+  ++HN+ L ++   YL+  L L T   L + T+
Sbjct: 792 NERRKDN-DSTMQLNDIISYYKLKYHSEVRHNNDLKVIN-DYLNKVLALGTRRLRLDTR 848




>gi|16804984|ref|NP_473013.1|  (NC_000910) protein of the YMR7 family [Plasmodium
falciparum]
 gi|7494452|pir||C71615  YMR7 homolog PFB0425c - malaria parasite (Plasmodium falciparum)
 gi|3845181|gb|AAC71874.1|  (AE001394) protein of the YMR7 family [Plasmodium falciparum
3D7]
          Length = 273

 Score = 30.8 bits (68), Expect = 2.6
 Identities = 18/50 (36%), Positives = 26/50 (52%)
 Frame = +3

Query: 207 IDKEYPELLEENYA*QYNNISVLFKQICSQSHKHFDIHLLWLGCQVVNHF 356
           IDK Y     EN     N+I+ ++K +  Q HK FDI L  +G + +  F
Sbjct: 178 IDKLYNSFYMENIL--CNDINDIYKDMTVQKHKRFDIVLDNMGVEFITDF 225




>gi|6322974|ref|NP_013046.1|  (NC_001144) Hypothetical ORF; Yll054cp [Saccharomyces
cerevisiae]
 gi|1077325|pir||S50966  probable membrane protein YLL054c - yeast (Saccharomyces
cerevisiae)
 gi|642330|emb|CAA88003.1|  (Z47973) ORF L0584 [Saccharomyces cerevisiae]
 gi|1360258|emb|CAA97507.1|  (Z73159) ORF YLL054c [Saccharomyces cerevisiae]
          Length = 769

 Score = 30.8 bits (68), Expect = 2.6
```

```
 Identities = 19/58 (32%), Positives = 25/58 (42%)
 Frame = +3

Query: 201 LKIDKEYPELLEENYA*QYNNISVLFKQICSQSHKHFDIHLLWLGCQVVNHFQMGPQY 374
           +K D+  PEL E NYA   N    L     +  H H  + +L   C  +N    GP Y
Sbjct: 444 MKTDESLPELFERNYASTNNLDLSLMATAENIYHLHLSLQVLLTSCFPIN----GPSY 497




>gi|1083855|pir||S41309  cyclosporin synthetase - cyclosporin fungus
 gi|440169|emb|CAA82227.1|  (Z28383) cyclosporine synthetase [Tolypocladium inflatum]
          Length = 15281

 Score = 30.8 bits (68), Expect = 2.6
 Identities = 17/59 (28%), Positives = 31/59 (51%)
 Frame = +2

Query: 59   RNQRMLEDVSDNHWRQMVYARKLFQGQSFQFHKILVCMGKWIVWRISVEDRQGIPRIIR 235
            +N ++  ++S   +  +++ R   Q QS   HKI      WI +  S  DRQG+ R+++
Sbjct: 2275 KNMKVNNELSSYRYGAVLHIRNHNQNQSRSIHKINA--ESWIDFASSQMDRQGLARLLK 2331




>gi|14719841|gb|AAK73196.1|  (AF320327) putative ruminococcin B modifying enzyme
[Ruminococcus gnavus]
          Length = 248

 Score = 30.4 bits (67), Expect = 3.5
 Identities = 14/44 (31%), Positives = 25/44 (56%)
 Frame = +3

Query: 108 WFMHESFFKDNPSSFTRSWFAWANGLFGELVLKIDKEYPELLEE 239
           +F   +++KD PS+ T+    W N L  +++  IDK Y    +E+
Sbjct: 11  YFYERNYYKDVPSTDTKMLEYWNNILGKDVIESIDKAYGVSIEQ 54




>gi|1085199|pir||S46260  creatine transporter - marbled electric ray
 gi|541754|emb|CAA54512.1|  (X77306) creatine transporter [Torpedo marmorata]
          Length = 611

 Score = 30.4 bits (67), Expect = 3.5
 Identities = 12/23 (52%), Positives = 17/23 (73%)
 Frame = +1

Query: 271 CYLSKYALNLINTLTSIFSGLVV 339
           CY   + L+L+N+ TS F+GLVV
Sbjct: 314 CYKDAFILSLVNSATSFFAGLVV 336




>gi|15237573|ref|NP_198932.1|  (NM_123481) purine permease-like protein [Arabidopsis
thaliana]
 gi|9759162|dbj|BAB09718.1|  (AB010072) purine permease-like protein [Arabidopsis thaliana]
          Length = 358

 Score = 30.0 bits (66), Expect = 4.5
```

```
 Identities = 15/42 (35%), Positives = 25/42 (58%)
 Frame = -3

Query: 127 KLSCINHLPPVVVTNILKHSLISDLSLVVNP*VTIAIGHIQL 2
           +L C+  +  + + + L  +LIS LSL+V P   IA+ H +L
Sbjct: 279 QLGCVGAVSLIFLVSSLFSNLISTLSLIVTPLAAIAVFHDKL 320



>gi|11132043|sp|O14653|GS27_HUMAN  27 kDa Golgi SNARE protein (Golgi SNAP receptor complex
member 2) (Membrin)
 gi|2316088|gb|AAB82651.1|   (AF007548) golgi SNARE [Homo sapiens]
          Length = 212

 Score = 29.6 bits (65), Expect = 5.9
 Identities = 20/60 (33%), Positives = 30/60 (49%)
 Frame = +2

Query: 125 LFQGQSFQFHKILVCMGKWIVWRISVEDRQGIPRIIRRELCLIIQQYFSAIQVNMLLISQ 304
           LFQ    Q H+I  CMG     R+   D+Q + I+  E+   I Q FS ++   +L S+
Sbjct: 4   LFQQTHKQVHEIQSCMG-----RLETADKQSV-HIVENEIQASIDQIFSRLERLEILSSK 57



>gi|16905522|ref|NP_004278.2|   (NM_004287) golgi SNAP receptor complex member 2 isoform
A; 27 kDa golgi SNARE protein; membrin  [Homo sapiens]
 gi|16307241|gb|AAH09710.1|AAH09710   (BC009710) Unknown (protein for MGC:9419) [Homo
sapiens]
          Length = 212

 Score = 29.6 bits (65), Expect = 5.9
 Identities = 20/60 (33%), Positives = 30/60 (49%)
 Frame = +2

Query: 125 LFQGQSFQFHKILVCMGKWIVWRISVEDRQGIPRIIRRELCLIIQQYFSAIQVNMLLISQ 304
           LFQ    Q H+I  CMG     R+   D+Q + I+  E+   I Q FS ++   +L S+
Sbjct: 4   LFQQTHKQVHEIQSCMG-----RLETADKQSV-HIVENEIQASIDQIFSRLERLEILSSK 57



>gi|19114410|ref|NP_593498.1|  (NC_003424) putative glycosylphosphatidylinositol anchor
attachment protein [Schizosaccharomyces pombe]
 gi|6689268|emb|CAB65611.1|  (AL136078) putative glycosylphosphatidylinositol anchor
attachment protein [Schizosaccharomyces pombe]
          Length = 581

 Score = 29.6 bits (65), Expect = 5.9
 Identities = 17/49 (34%), Positives = 27/49 (54%), Gaps = 3/49 (6%)
 Frame = +2

Query: 248 LIIQQYFSAIQVN--MLLISQTL*HPSSLAWLSSCQSFLNGPLI-FWLW 385
           L++   F + + N  + L+   L + S+  +L S    LNGPL+ FWLW
Sbjct: 486 LLVPLQFISFRFNRRLSLLFAVLTYFSTFIFLCSLSKILNGPLVPFWLW 534
```

```
>gi|17564302|ref|NP_504743.1| ☐ (NM_072342) T15B7.1.p [Caenorhabditis elegans]
 gi|7507850|pir||T32255  hypothetical protein T15B7.1 - Caenorhabditis elegans
 gi|2384954|gb|AAB69973.1|  (AF022985) Hypothetical protein T15B7.1 [Caenorhabditis elegans]
          Length = 146

 Score = 29.6 bits (65), Expect = 5.9
 Identities = 13/35 (37%), Positives = 21/35 (59%)
 Frame = +3

Query: 93  TTGGKWFMHESFFKDNPSSFTRSWFAWANGLFGEL 197
           T GG W + ++ F D+ S + R W  + NG FG++
Sbjct: 3   TYGGGWVLFQNRFDDSESYWDRKWDEYKNG-FGDV 36



>gi|16905520|ref|NP_473363.1| ☐ (NM_054022) golgi SNAP receptor complex member 2 isoform
B; 27 kDa golgi SNARE protein; membrin [Homo sapiens]
 gi|12711467|gb|AAK01855.1|AF229796_1 ☐ (AF229796) golgi SNARE isoform B [Homo sapiens]
          Length = 213

 Score = 29.6 bits (65), Expect = 5.9
 Identities = 20/60 (33%), Positives = 30/60 (49%)
 Frame = +2

Query: 125 LFQGQSFQFHKILVCMGKWIVWRISVEDRQGIPRIIRRELCLIIQQYFSAIQVNMLLISQ 304
           LFQ    Q H+I  CMG     R+   D+Q +  I+  E+   I Q FS ++   +L S+
Sbjct: 4   LFQQTHKQVHEIQSCMG-----RLETADKQSV-HIVENEIQASIDQIFSRLERLEILSSK 57



>gi|15896721|ref|NP_350070.1|  (NC_003030) Short-chain alcohol dehydrogenase family protein
[Clostridium acetobutylicum]
 gi|15026575|gb|AAK81410.1|AE007846_1  (AE007846) Short-chain alcohol dehydrogenase family
protein [Clostridium acetobutylicum]
          Length = 291

 Score = 29.3 bits (64), Expect = 7.7
 Identities = 16/43 (37%), Positives = 26/43 (60%)
 Frame = -3

Query: 187 NNPFAHANQDLVKLEGLSLKKLSCINHLPPVVVTNILKHSLIS 59
           NN   HA +  V  +GL+  ++  +N+L P ++TN L  SLI+
Sbjct: 89  NNAGLHAFEQRVTSDGLA--EMMAVNYLAPWLLTNTLLQSLIA 129



>gi|10954561|ref|NP_044683.1|  (NC_001811) unknown [Methanococcus maripaludis]
 gi|2072768|gb|AAC45253.1|  (U47023) unknown [Methanococcus maripaludis]
          Length = 99

 Score = 29.3 bits (64), Expect = 7.7
 Identities = 14/36 (38%), Positives = 24/36 (65%)
 Frame = +1

Query: 253 NIIIFQCYLSKYALNLINTLTSIFSGLVVKLSIISK 360
           NII+F C ++   L+LI+  ++ F  + + +SIISK
Sbjct: 22  NIILFSCSVNTNKLDLISIFSTAFLIMSLTISIISK 57
```

```
>gi|15895879|ref|NP_349228.1|  (NC_003030) Cell wall hydrolase (autolysin), family 25 of
glycosyl hydrolases; peptodoglycan-binding domain [Clostridium acetobutylicum]
 gi|15025646|gb|AAK80568.1|AE007759_6  (AE007759) Cell wall hydrolase (autolysin), family 25
of glycosyl hydrolases; peptodoglycan-binding domain [Clostridium acetobutylicum]
          Length = 325

 Score = 29.3 bits (64), Expect = 7.7
 Identities = 17/58 (29%), Positives = 28/58 (47%), Gaps = 5/58 (8%)
 Frame = -3

Query: 292  EHIYLNSTEILLYY-----QAQFSSNNSGYSLSIFNTNSPNNPFAHANQDLVKLEGLS 134
           E +Y+ +TE + Y      Q Q   + N+G   +   ++    NNP   A+  L   + GLS
Sbjct: 24   EIVYIKATEGVTYVDKTIRQFQSDAKNAGLKVGFYHFLRANNPINEADNFLSAISGLS 81




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 227,113,096
Number of Sequences: 919285
Number of extensions: 4517927
Number of successful extensions: 14228
Number of sequences better than 10.0: 60
Number of HSP's better than 10.0 without gapping: 13864
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 14221
length of database: 288,558,979
effective HSP length: 112
effective length of database: 185,599,059
effective search space used: 4454377416
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020954469-026213-20966
```

7.1.11 Query= **hy-11_16b150013**

```
         (382 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 300 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. prot...    75   7e-21
gi|16805189|ref|NP_473217.1|  (NC_000521) predicted using he...    72   5e-20
gi|15383902|emb|CAB11118.3|  (Z98547) Calcium-dependent prot...    72   5e-20
gi|7527460|gb|AAF63154.1|AF106064_1  (AF106064) calcium-depe...    72   5e-20
gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falci...    70   2e-19
gi|6318615|gb|AAF06969.1|AF162661_1  (AF162661) phosphoenolp...    75   7e-19
gi|6625808|gb|AAF19401.1|AF203479_1  (AF203479) phosphoenolp...    73   7e-19
gi|6983879|dbj|BAA90814.1|  (AP001168) ESTs AU030197(E50746)...    67   8e-19
gi|1076202|pir||S54788  calcium-stimulated protein kinase - ...    65   1e-18
gi|15222023|ref|NP_172719.1|  (NM_101129) calcium-dependent ...    67   4e-18
gi|12830367|emb|CAC29064.1|  (AJ303380) serine/threonine kin...    69   3e-17
gi|6644464|gb|AAF21062.1|  (AF216527) calcium-dependent prot...    59   6e-17
gi|20340969|ref|XP_110303.1|  (XM_110303) similar to calcium...    60   7e-17
gi|2077934|dbj|BAA19880.1|  (D86557) Protein Kinase [Rattus ...    60   7e-17
gi|6175636|gb|AAF05112.1|AF158091_1  (AF158091) phosphoenolp...    69   7e-17
gi|7428007|pir||T02784  calcium-dependent protein kinase (EC...    60   9e-17
gi|4678722|emb|CAB41259.1|  (AL049688) hypothetical protein ...    59   1e-16
gi|16755792|gb|AAL28100.1|AF428261_1  (AF428261) calcium/cal...    59   1e-16
gi|14196445|ref|NP_065172.1|  (NM_020439) calcium/calmodulin...    59   1e-16
gi|4007153|emb|CAA19296.1|  (AL023754) dJ272L16.1 (Rat Ca2+/...    59   1e-16
gi|1362152|pir||S56639  ribosomal protein S6 kinase homolog ...    61   1e-16
```

```
gi|15224978|ref|NP_181425.1|  (NM_129449) putative calcium-d...   60   2e-16
gi|15289758|dbj|BAB63463.1|  (AB051808) calcium dependent pr...   60   2e-16
gi|12484155|gb|AAG53994.1|AF333959_1  (AF333959) calmodulin-...   60   2e-16
gi|4325074|gb|AAD17247.1|  (AF118101) protein kinase 6 [Toxo...   60   2e-16
gi|15223870|ref|NP_172341.1|  (NM_100738) putative calcium-d...   67   2e-16
gi|7434322|pir||T00718  calcium-dependent protein kinase hom...   67   2e-16
gi|15238528|ref|NP_197831.1|  (NM_122351) calcium dependent ...   53   2e-16
gi|17975557|ref|NP_524622.1|  (NM_079883) Calcium/calmodulin...   56   2e-16
gi|6625812|gb|AAF19403.1|AF203481_1  (AF203481) phosphoenolp...   65   2e-16
gi|6625810|gb|AAF19402.1|AF203480_1  (AF203480) phosphoenolp...   65   2e-16
gi|16215471|emb|CAC82999.1|  (AJ344155) calcium-dependent pr...   59   3e-16
gi|9294489|dbj|BAB02708.1|  (AB019230) IRE homolog; protein ...   65   3e-16
gi|15229535|ref|NP_188412.1|  (NM_112666) protein kinase, pu...   65   3e-16
gi|6729348|dbj|BAA89784.1|  (AB037134) IRE homolog 1 [Arabid...   65   3e-16
gi|19111951|ref|NP_595159.1|  (NC_003423) camp-dependent pro...   70   4e-16
gi|3282250|gb|AAC24961.1|  (AF009337) CDPK-related protein k...   54   4e-16
gi|7446432|pir||T02033  calcium/calmodulin-dependent protein...   54   5e-16
gi|7446433|pir||T02994  CDPK-related protein kinase - maize ...   54   5e-16
gi|12039322|gb|AAG46110.1|AC073166_8  (AC073166) calcium-dep...   59   5e-16
gi|1705733|sp|P53682|CDP1_ORYSA  Calcium-dependent protein k...   59   5e-16
gi|14484895|gb|AAK62812.1|  (AF276999) calcium-dependent pro...   55   5e-16
gi|7446430|pir||T01694  calcium-dependent protein kinase-rel...   54   5e-16
gi|16902296|dbj|BAB71853.1|  (AB065100) phosphoenolpyruvate ...   64   5e-16
gi|15236560|ref|NP_194096.1|  (NM_118496) calcium-dependent ...   57   6e-16
gi|15450437|gb|AAK96512.1|  (AY052319) AT4g23650/F9D16_120 [...   57   6e-16
gi|1730055|sp|P25323|KMLC_DICDI  Myosin light chain kinase (...   62   6e-16
gi|17542684|ref|NP_501900.1|  (NM_069499) calcium/calmodulin...   57   8e-16
gi|7505788|pir||T23616  hypothetical protein K11E8.1c - Caen...   57   8e-16
gi|15229129|ref|NP_190506.1|  (NM_114797) calcium dependent ...   54   8e-16
gi|15229002|ref|NP_191235.1|  (NM_115535) calcium-dependent ...   53   8e-16
gi|17542680|ref|NP_501902.1|  (NM_069501) Ca2+/calmodulin-de...   57   8e-16
gi|17542678|ref|NP_501896.1|  (NM_069495) Ca2+/calmodulin-de...   57   8e-16
gi|7505786|pir||T23614  hypothetical protein K11E8.1a - Caen...   57   8e-16
gi|17542674|ref|NP_501901.1|  (NM_069500) Ca2+/calmodulin-de...   57   8e-16
gi|17542682|ref|NP_501898.1|  (NM_069497) Ca2+/calmodulin-de...   57   8e-16
gi|17542676|ref|NP_501899.1|  (NM_069498) Ca2+/calmodulin-de...   57   8e-16
gi|17542672|ref|NP_501897.1|  (NM_069496) Ca2+/calmodulin-de...   57   8e-16
gi|17542688|ref|NP_501904.1|  (NM_069503) calcium/calmodulin...   57   8e-16
gi|17542690|ref|NP_501905.1|  (NM_069504) Ca2+/calmodulin-de...   57   8e-16
gi|17542686|ref|NP_501903.1|  (NM_069502) Ca2+/calmodulin-de...   57   8e-16
gi|102256|pir||A40811  myosin-light-chain kinase (EC 2.7.1.1...   62   8e-16
gi|16215467|emb|CAC82998.1|  (AJ344154) calcium-dependent pr...   59   1e-15
gi|18416872|ref|NP_568281.1|  (NM_121286) calcium-dependent ...   55   1e-15
gi|20343902|ref|XP_110633.1|  (XM_110633) RIKEN cDNA 4921505...   64   1e-15
gi|2271459|gb|AAC13354.1|  (AF009560) calcium-dependent prot...   67   1e-15
gi|14148952|emb|CAC39171.1|  (AJ307671) Serine/Threonine kin...   64   1e-15
gi|17064926|gb|AAL32617.1|  (AY062539) calcium-dependent pro...   55   1e-15
gi|15231060|ref|NP_190753.1|  (NM_115044) calcium-dependent ...   55   1e-15
gi|17560450|ref|NP_505770.1|  (NM_073369) calcium/calmodulin...   59   1e-15
gi|16904226|gb|AAL30820.1|AF435452_1  (AF435452) calcium/cal...   52   2e-15
gi|3298603|gb|AAC25955.1|  (AF054621) serine/threonine kinas...   60   2e-15
gi|4521233|dbj|BAA76292.1|  (AB017332) Aurora/Ipl1-related k...   60   2e-15
gi|4507273|ref|NP_003151.1|  (NM_003160) serine/threonine ki...   60   2e-15
gi|3289986|gb|AAC25618.1|  (AC005261) STK13 [Homo sapiens]        60   2e-15
gi|15144311|gb|AAK84452.1|AC087192_13  (AC087192) putative k...   52   2e-15
gi|15225092|ref|NP_180708.1|  (NM_128707) putative calcium-d...   54   2e-15
```

```
gi|3859672|emb|CAA22010.1|  (AL033502) serine-threonine prot...    61   2e-15
gi|15226841|ref|NP_181647.1|  (NM_129679) CPDK-related prote...    52   3e-15
gi|6063536|dbj|BAA85396.1|  (AP000615) ESTs C22369(C12239),C...    56   3e-15
gi|1705735|sp|P53684|CDP3_ORYSA  Calcium-dependent protein k...    56   3e-15
gi|2944385|gb|AAC05270.1|  (AF048691) calcium dependent prot...    56   3e-15
gi|14594815|emb|CAC43293.1|  (AJ309171) putative phosphoenol...    63   3e-15
gi|4758666|ref|NP_004681.1|  (NM_004690) LATS homolog 1 [Hom...    58   4e-15
gi|20127418|ref|NP_001211.2|  (NM_001220) calcium/calmodulin...    58   4e-15
gi|15228350|ref|NP_187677.1|  (NM_111902) calmodulin-domain ...    56   4e-15
gi|7434373|pir||S68470  Ca2+/calmodulin-dependent protein ki...    58   4e-15
gi|6671660|ref|NP_031621.1|  (NM_007595) calcium/calmodulin-...    58   4e-15
gi|4139270|gb|AAD03744.1|  (AF112472) calcium/calmodulin-dep...    58   4e-15
gi|5326757|gb|AAD42035.1|AF078803_1  (AF078803) calcium/calm...    58   4e-15
gi|11120682|ref|NP_068507.1|  (NM_021739) Ca++/calmodulin-de...    58   4e-15
gi|466360|gb|AAA81938.1|  (U06636) calmodulin dependent prot...    58   4e-15
gi|3668373|gb|AAC79460.1|  (AF085249) calcium/calmodulin-dep...    58   4e-15
gi|603213|gb|AAA57338.1|  (U18196) calcium/calmodulin-depend...    58   4e-15
gi|6688228|emb|CAB65122.1|  (AJ252238) calcium/calmodulin de...    58   4e-15
gi|5326764|gb|AAD42038.1|AF083419_1  (AF083419) calcium/calm...    58   4e-15
gi|4139268|gb|AAD03743.1|  (AF112471) calcium/calmodulin-dep...    58   4e-15
gi|6688224|emb|CAB65120.1|  (AJ252236) calcium/calmodulin de...    58   4e-15
gi|5326759|gb|AAD42036.1|AF081572_1  (AF081572) calcium/calm...    58   4e-15
gi|5326762|gb|AAD42037.1|AF081924_1  (AF081924) calcium/calm...    58   4e-15
```

Alignments

```
>gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. protein kinase (C-term. EF
hand) [Plasmodium falciparum]
 gi|422320|pir||A45472  protein kinase (EC 2.7.1.37) - malaria parasite (Plasmodium
falciparum)
 gi|9878|emb|CAA47704.1|  (X67288) protein kinase [Plasmodium falciparum]
 gi|3845284|gb|AAC71952.1|  (AE001419) calcium-dept. protein kinase (C-term. EF
hand) [Plasmodium falciparum 3D7]
          Length = 524

 Score = 74.7 bits (182), Expect(2) = 7e-21
 Identities = 31/65 (47%), Positives = 47/65 (71%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           +++KL+  +GT  Y+APEV+   YNEKCDVWS GV+L+ +L GYPPF G+N Q++ + ++
Sbjct: 221 KDNKLRDRLGTAYYIAPEVLRKKYNEKCDVWSCGVILYILLCGYPPFGGQNDQDIIKKVE 280

Query: 46  YQNVY 32
               Y
Sbjct: 281 KGKYY 285




 Score = 45.1 bits (105), Expect(2) = 7e-21
 Identities = 18/36 (50%), Positives = 26/36 (72%)
 Frame = -2
```

```
Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+H+  I+HRDIKPENI+  ++      K++DFGLS
Sbjct: 181 YLHKHNIVHRDIKPENILLENKHSLLNIKIVDFGLS 216



>gi|16805189|ref|NP_473217.1|  (NC_000521) predicted using hexExon; MAL3P3.17
(PFC0420w), Calcium-dependent protein kinase, len: 591 aa; Similarity to calcium-
dependent protein kinases. P.falciparum calcium-dependent protein kinase CDPK2
(TR:O15865) BLAST Score: 1069, sum P(1) = 2.3>
 gi|7494209|pir||T18445  hypothetical protein C0420w - malaria parasite (Plasmodium
falciparum)
          Length = 608

 Score = 72.0 bits (175), Expect(2) = 5e-20
 Identities = 28/61 (45%), Positives = 43/61 (69%)
 Frame = -1

Query: 229 YQNSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + N ++K   GTP Y+AP+V+ G+YN KCD+WS GVL + +L GYPPF+G +  E+   +
Sbjct: 268 HNNYEMKTKAGTPYYVAPQVLTGSYNYKCDMWSSGVLFYILLCGYPPFFGESDHEILSMV 327

Query: 49  Q 47
           +
Sbjct: 328 K 328



 Score = 45.1 bits (105), Expect(2) = 5e-20
 Identities = 21/45 (46%), Positives = 29/45 (63%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKYDAFSYK 244
           Y+H + I HRDIKPEN +F D    S  K+IDFGL+  +   +Y+
Sbjct: 228 YLHIRNICHRDIKPENFLFYDMTPESLIKIIDFGLASYFTHNNYE 272



>gi|15383902|emb|CAB11118.3|  (Z98547) Calcium-dependent protein kinase [Plasmodium
falciparum 3D7]
          Length = 580

 Score = 72.0 bits (175), Expect(2) = 5e-20
 Identities = 28/61 (45%), Positives = 43/61 (69%)
 Frame = -1

Query: 229 YQNSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + N ++K   GTP Y+AP+V+ G+YN KCD+WS GVL + +L GYPPF+G +  E+   +
Sbjct: 268 HNNYEMKTKAGTPYYVAPQVLTGSYNYKCDMWSSGVLFYILLCGYPPFFGESDHEILSMV 327

Query: 49  Q 47
           +
Sbjct: 328 K 328
```

```
 Score = 45.1 bits (105), Expect(2) = 5e-20
 Identities = 21/45 (46%), Positives = 29/45 (63%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKYDAFSYK 244
           Y+H + I HRDIKPEN +F D    S  K+IDFGL+  +   +Y+
Sbjct: 228 YLHIRNICHRDIKPENFLFYDMTPESLIKIIDFGLASYFTHNNYE 272
```

```
>gi|7527460|gb|AAF63154.1|AF106064_1  (AF106064) calcium-dependent protein kinase-3
[Plasmodium falciparum] [Plasmodium falciparum 3D7]
          Length = 562

 Score = 72.0 bits (175), Expect(2) = 5e-20
 Identities = 28/61 (45%), Positives = 43/61 (69%)
 Frame = -1

Query: 229 YQNSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + N ++K   GTP Y+AP+V+ G+YN KCD+WS GVL + +L GYPPF+G +  E+   +
Sbjct: 268 HNNYEMKTKAGTPYYVAPQVLTGSYNYKCDMWSSGVLFYILLCGYPPFFGESDHEILSMV 327

Query: 49  Q 47
           +
Sbjct: 328 K 328
```

```
 Score = 45.1 bits (105), Expect(2) = 5e-20
 Identities = 21/45 (46%), Positives = 29/45 (63%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKYDAFSYK 244
           Y+H + I HRDIKPEN +F D    S  K+IDFGL+  +   +Y+
Sbjct: 228 YLHIRNICHRDIKPENFLFYDMTPESLIKIIDFGLASYFTHNNYE 272
```

```
>gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falciparum]
          Length = 513

 Score = 70.1 bits (170), Expect(2) = 2e-19
 Identities = 25/51 (49%), Positives = 40/51 (78%)
 Frame = -1

Query: 199 GTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           GTP Y+AP+V++G Y++KCD+WS GV+++ +L GYPPFYG    E+ + ++
Sbjct: 231 GTPYYVAPQVLDGKYDKKCDIWSSGVIMYTLLCGYPPFYGDTDNEVLKKVK 281
```

```
 Score = 44.7 bits (104), Expect(2) = 2e-19
 Identities = 19/36 (52%), Positives = 25/36 (68%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+H   I+HRD+KPEN +F    +  S  K+IDFGLS
Sbjct: 182 YLHSLNIVHRDLKPENFLFQSENKDSLLKIIDFGLS 217



>gi|6318615|gb|AAF06969.1|AF162661_1  (AF162661) phosphoenolpyruvate carboxylase
kinase [Kalanchoe fedtschenkoi]
 gi|6318617|gb|AAF06970.1|AF162662_1  (AF162662) phosphoenolpyruvate carboxylase
kinase [Kalanchoe fedtschenkoi]
          Length = 274

 Score = 74.7 bits (182), Expect(2) = 7e-19
 Identities = 32/57 (56%), Positives = 43/57 (75%), Gaps = 1/57 (1%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV++G  YNEK DVWS GV+L+ ML G+PPFYG   Q+++E +   N+
Sbjct: 168 VGTPYYVAPEVLQGREYNEKVDVWSAGVILYTMLAGFPPFYGETAQDIFEAVMRGNL 224




 Score = 38.5 bits (88), Expect(2) = 7e-19
 Identities = 17/35 (48%), Positives = 23/35 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           A+ H +G+ HRD+KPEN++F   D     KL DFG
Sbjct: 122 AHCHRRGVCHRDVKPENVLF---DSVGRLKLADFG 153



>gi|6625808|gb|AAF19401.1|AF203479_1  (AF203479) phosphoenolpyruvate carboxylase
kinase [Glycine max]
          Length = 274

 Score = 72.8 bits (177), Expect(2) = 7e-19
 Identities = 33/65 (50%), Positives = 46/65 (70%), Gaps = 1/65 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
            + S +   VGTP Y+APEVI G  Y+EK DVWS GV+L+ ML G+PPFYG +  E++E++
Sbjct: 154 EGSSMSGVVGTPYYVAPEVIMGREYDEKVDVWSSGVILYAMLAGFPPFYGESAPEIFESV 213
```

```
Query: 49  QYQNV 35
              N+
Sbjct: 214 LRANL 218




 Score = 40.4 bits (93), Expect(2) = 7e-19
 Identities = 20/35 (57%), Positives = 25/35 (71%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           A+ H QG+ HRDIKPENI+F   D  ++ KL DFG
Sbjct: 116 AHCHAQGLAHRDIKPENILF---DEGNKLKLSDFG 147



>gi|6983879|dbj|BAA90814.1|  (AP001168) ESTs AU030197(E50746),AU030196(E50746)
correspond to a region of the predicted gene.~Similar to calcium-dependent
calmodulin-independent protein kinase CDPK (U90262) [Oryza sativa (japonica
cultivar-group)]
          Length = 587

 Score = 67.4 bits (163), Expect(2) = 8e-19
 Identities = 28/56 (50%), Positives = 41/56 (73%)
 Frame = -1

Query: 214 LKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           L   VG+P Y+APEV+ G YNE  DVWS GV+L+ +L G PPF+G+ + +++E I+
Sbjct: 276 LSGMVGSPFYIAPEVLAGGYNEAADVWSAGVILYILLSGIPPFWGKTKSKIFECIR 331



 Score = 45.4 bits (106), Expect(2) = 8e-19
 Identities = 19/37 (51%), Positives = 26/37 (69%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H +GI+HRD+KPENI+   +   S  KL DFGL+
Sbjct: 231 AHCHSKGIVHRDLKPENILLVSKSPSSPIKLADFGLA 267




>gi|1076202|pir||S54788  calcium-stimulated protein kinase - Chlamydomonas
eugametos
 gi|806542|emb|CAA89202.1|  (Z49233) calcium-stimulated protein kinase
[Chlamydomonas eugametos]
          Length = 591
```

```
 Score = 65.5 bits (158), Expect(2) = 1e-18
 Identities = 25/56 (44%), Positives = 42/56 (74%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VG+P Y+APEV+  +Y ++ D+WS GV+L+ +L GYPPF+G N   +++E I  +++
Sbjct: 314 VGSPFYVAPEVVRRSYGKEADIWSCGVILYILLCGYPPFHGDNETKIFEAIMNKSL 369




 Score = 46.6 bits (109), Expect(2) = 1e-18
 Identities = 20/40 (50%), Positives = 26/40 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKY 262
           A+ H+ G+IHRD+KPEN +  D+      K IDFGLS  Y
Sbjct: 265 AHCHDMGVIHRDLKPENFLLLDKSANPVLKSIDFGLSSFY 304



>gi|15222023|ref|NP_172719.1|  (NM_101129) calcium-dependent protein kinase,
putative [Arabidopsis thaliana]
 gi|8778638|gb|AAF79646.1|AC025416_20  (AC025416) F5O11.32 [Arabidopsis thaliana]
 gi|9502372|gb|AAF88079.1|AC025417_7  (AC025417) T12C24.12 [Arabidopsis thaliana]
 gi|16323176|gb|AAL15322.1|  (AY057692) At1g12580/T12C24_10 [Arabidopsis thaliana]
          Length = 522

 Score = 67.0 bits (162), Expect(2) = 4e-18
 Identities = 26/57 (45%), Positives = 42/57 (73%)
 Frame = -1

Query: 217 KLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           KL   VG+P Y+APEV+ G YN+  DVWS GV+L+ +L G PPF+G+ + ++++ ++
Sbjct: 201 KLSGTVGSPFYIAPEVLAGGYNQAADVWSAGVILYILLSGAPPFWGKTKSKIFDAVR 257




 Score = 43.5 bits (101), Expect(2) = 4e-18
 Identities = 18/36 (50%), Positives = 24/36 (66%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           + H+ GI+HRD+KPENI+       S  KL DFGL+
Sbjct: 158 FCHDSGIVHRDLKPENILMATMSSSSPIKLADFGLA 193



>gi|12830367|emb|CAC29064.1|   (AJ303380) serine/threonine kinase 33 [Homo
sapiens]
```

```
          Length = 514

 Score = 68.6 bits (166), Expect(2) = 3e-17
 Identities = 31/83 (37%), Positives = 54/83 (64%), Gaps = 1/83 (1%)
 Frame = -1

Query: 250 LQVRKSQYQNSKLKAGVGTPVYLAPEVIEG-TYNEKCDVWSLGVLLFNMLVGYPPFYGRN 74
           L V+K      + L+A  GTP+Y+APEVI    Y+++CD+WS+GV+++ +L G PPF   +
Sbjct: 268 LAVKKQSRSEAMLQATCGTPIYMAPEVISAHDYSQQCDIWSIGVVMYMLLRGEPPFLASS 327

Query: 73  RQELYENIQYQNVYNLNYIYQLI 5
            ++L+E I+    ++  N ++   I
Sbjct: 328 EEKLFELIRKGELHFENAVWNSI 350




 Score = 39.3 bits (90), Expect(2) = 3e-17
 Identities = 20/45 (44%), Positives = 27/45 (59%), Gaps = 6/45 (13%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMF------TDRDIRSEPKLIDFGLSVK 265
           AY+H   I+HRD+K ENIM        + +I    K+ DFGL+VK
Sbjct: 227 AYLHNNDIVHRDLKLENIMVKSSLIDDNNEINLNIKVTDFGLAVK 271



>gi|6644464|gb|AAF21062.1|  (AF216527) calcium-dependent protein kinase [Dunaliella
tertiolecta]
          Length = 595

 Score = 58.9 bits (141), Expect(2) = 6e-17
 Identities = 22/56 (39%), Positives = 42/56 (74%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VG+  Y+APEV+  +Y ++ D+WS G++L+ +L G+PPF+G + ++++E I  ++V
Sbjct: 317 VGSAYYVAPEVLRRSYGKEADIWSCGIILYILLCGFPPFHGDSEKKIFEAIISKSV 372




 Score = 47.8 bits (112), Expect(2) = 6e-17
 Identities = 19/37 (51%), Positives = 27/37 (72%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H  G+IHRD+KPEN + +D+  R++ K  DFGLS
Sbjct: 268 AHCHNMGVIHRDLKPENFLLSDKSPRAQLKATDFGLS 304
```

```
>gi|20340969|ref|XP_110303.1|  (XM_110303) similar to calcium/calmodulin-dependent
protein kinase IG [Mus musculus]
 gi|14422219|emb|CAC41379.1|  (AL365314) bM106N23.1 (novel protein kinase (ortholog
of human dJ272L16.1 and rat Ca2+/calmodulin dependent protein kinase)) [Mus
musculus]
 gi|16755794|gb|AAL28101.1|AF428262_1  (AF428262) calcium/calmodulin-dependent
protein kinase I gamma [Mus musculus]
 gi|18256867|gb|AAH21840.1|AAH21840  (BC021840) Unknown (protein for MGC:30513)
[Mus musculus]
          Length = 477

 Score = 59.7 bits (143), Expect(2) = 7e-17
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +    GTP Y+APEV+ +  Y++  D WS+GV+ + +L GYPPFY     +L+E I
Sbjct: 172 QNGVMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 231

Query: 49  Q 47
           +
Sbjct: 232 K 232



 Score = 46.6 bits (109), Expect(2) = 7e-17
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2


Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++   +  S+  + DFGLS
Sbjct: 133 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 168



>gi|2077934|dbj|BAA19880.1|  (D86557) Protein Kinase [Rattus norvegicus]
          Length = 309

 Score = 59.7 bits (143), Expect(2) = 7e-17
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +    GTP Y+APEV+ +  Y++  D WS+GV+ + +L GYPPFY     +L+E I
Sbjct: 172 QNGVMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 231

Query: 49  Q 47
           +
Sbjct: 232 K 232
```

```
 Score = 46.6 bits (109), Expect(2) = 7e-17
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++   +  S+  + DFGLS
Sbjct: 133 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 168



>gi|6175636|gb|AAF05112.1|AF158091_1  (AF158091) phosphoenolpyruvate carboxylase-
kinase [Mesembryanthemum crystallinum]
          Length = 279

 Score = 68.6 bits (166), Expect(2) = 7e-17
 Identities = 32/62 (51%), Positives = 43/62 (68%), Gaps = 1/62 (1%)
 Frame = -1

Query: 217 KLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQ 41
           +++  VGTP Y+APEV+ G  YNEK DVWS GV+L+ ML G PPFYG   +E +E +
Sbjct: 169 EMRGVVGTPYYVAPEVLSGKDYNEKADVWSAGVILYIMLGGVPPFYGETVEETFEAVLRG 228

Query: 40  NV 35
           N+
Sbjct: 229 NL 230



 Score = 37.7 bits (86), Expect(2) = 7e-17
 Identities = 18/35 (51%), Positives = 23/35 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           A+ H    + HRDIKP+NI+F   D R+  KL DFG
Sbjct: 127 AHCHRNYVAHRDIKPDNILF---DSRNRLKLCDFG 158



>gi|7428007|pir||T02784  calcium-dependent protein kinase (EC 2.7.1.-) - maize
(strain W64A)
 gi|3320104|emb|CAA07481.1|  (AJ007366) calcium-dependent protein kinase [Zea mays]
          Length = 639

 Score = 60.1 bits (144), Expect(2) = 9e-17
 Identities = 23/51 (45%), Positives = 36/51 (70%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV++  Y  + DVWS GV+++ +L G PPF+  N Q ++E +
Sbjct: 315 VGSPYYVAPEVLKKRYGPEADVWSAGVIIYILLCGVPPFWAENEQGIFEEV 365
```

```
 Score = 45.8 bits (107), Expect(2) = 9e-17
 Identities = 18/35 (51%), Positives = 24/35 (68%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +F D    +  K IDFGLS+
Sbjct: 269 HSMGVMHRDLKPENFLFADHSEEAALKTIDFGLSI 303



>gi|4678722|emb|CAB41259.1| ☐ (AL049688) hypothetical protein [Homo sapiens]
          Length = 481

 Score = 59.3 bits (142), Expect(2) = 1e-16
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +    GTP Y+APEV+ +  Y++  D WS+GV+ + +L GYPPFY     +L+E I
Sbjct: 177 QNGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 236

Query: 49  Q 47
           +
Sbjct: 237 K 237



 Score = 46.6 bits (109), Expect(2) = 1e-16
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++   +  S+  + DFGLS
Sbjct: 138 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 173



>gi|16755792|gb|AAL28100.1|AF428261_1 ☐ (AF428261) calcium/calmodulin-dependent
protein kinase I gamma [Homo sapiens]
          Length = 476

 Score = 59.3 bits (142), Expect(2) = 1e-16
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1
```

```
Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +     GTP Y+APEV+ +   Y++  D WS+GV+ + +L GYPPFY      +L+E I
Sbjct: 172 QNGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 231

Query: 49  Q 47
           +
Sbjct: 232 K 232




 Score = 46.6 bits (109), Expect(2) = 1e-16
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++    +   S+   + DFGLS
Sbjct: 133 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 168



>gi|14196445|ref|NP_065172.1| ☐ (NM_020439) calcium/calmodulin-dependent protein
kinase IG [Homo sapiens]
          Length = 476

 Score = 59.3 bits (142), Expect(2) = 1e-16
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +     GTP Y+APEV+ +   Y++  D WS+GV+ + +L GYPPFY      +L+E I
Sbjct: 172 QNGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 231

Query: 49  Q 47
           +
Sbjct: 232 K 232




 Score = 46.6 bits (109), Expect(2) = 1e-16
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++    +   S+   + DFGLS
Sbjct: 133 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 168
```

```
>gi|4007153|emb|CAA19296.1|  (AL023754) dJ272L16.1 (Rat Ca2+/Calmodulin
dependent Protein Kinase LIKE protein) [Homo sapiens]
          Length = 460

 Score = 59.3 bits (142), Expect(2) = 1e-16
 Identities = 26/61 (42%), Positives = 39/61 (63%), Gaps = 1/61 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           QN  +    GTP Y+APEV+ +  Y++  D WS+GV+ + +L GYPPFY     +L+E I
Sbjct: 172 QNGIMSTACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKI 231

Query: 49  Q 47
           +
Sbjct: 232 K 232




 Score = 46.6 bits (109), Expect(2) = 1e-16
 Identities = 18/36 (50%), Positives = 27/36 (75%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+HE GI+HRD+KPEN+++   +  S+  + DFGLS
Sbjct: 133 YLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLS 168




>gi|1362152|pir||S56639  ribosomal protein S6 kinase homolog (clone Aspk11) - oat
 gi|871986|emb|CAA56313.1|  (X79992) putative pp70 ribosomal protein S6 kinase
[Avena sativa]
          Length = 480

 Score = 60.8 bits (146), Expect(2) = 1e-16
 Identities = 27/71 (38%), Positives = 47/71 (66%), Gaps = 1/71 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVIEGTYNEKC-DVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           +N++  +  GT  Y+APE+++G  ++K  D WS+G+LLF ML G PPF+G NR ++ + I
Sbjct: 301 ENTRSNSMCGTVEYMAPEIVQGRGHDKAADWWSVGILLFEMLTGKPPFFGGNRDKIQQKI 360

Query: 49  QYQNVYNLNYI 17
             + +   +Y+
Sbjct: 361 VREKMKLPSYL 371




 Score = 44.7 bits (104), Expect(2) = 1e-16
 Identities = 20/41 (48%), Positives = 28/41 (67%)
 Frame = -2
```

```
Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKYD 259
           A++H  GI+HRD+KPENI+    D R    L DFGL+ ++D
Sbjct: 263 AHLHANGIMHRDLKPENILL---DARGHAMLTDFGLAKEFD 300



>gi|15224978|ref|NP_181425.1|  (NM_129449) putative calcium-dependent protein
kinase [Arabidopsis thaliana]
 gi|3928078|gb|AAC79604.1|  (AC005770) putative calcium-dependent protein kinase
[Arabidopsis thaliana]
          Length = 583

 Score = 60.5 bits (145), Expect(2) = 2e-16
 Identities = 23/51 (45%), Positives = 36/51 (70%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV+   Y+ +CDVWS GV+++ +L G PPF+    Q ++E +
Sbjct: 296 VGSPYYVAPEVLRKHYSHECDVWSAGVIIYILLSGVPPFWDETEQGIFEQV 346




 Score = 44.7 bits (104), Expect(2) = 2e-16
 Identities = 19/35 (54%), Positives = 24/35 (68%)
 Frame = -2


Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +F   D  +  K IDFGLSV
Sbjct: 250 HSLGVMHRDLKPENFLFVSGDEEAALKTIDFGLSV 284



>gi|15289758|dbj|BAB63463.1|  (AB051808) calcium dependent protein kinase [Solanum
tuberosum]
          Length = 578

 Score = 60.1 bits (144), Expect(2) = 2e-16
 Identities = 24/51 (47%), Positives = 35/51 (68%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV+   Y  + DVWS GV+L+ +L G PPF+  N Q ++E +
Sbjct: 276 VGSPYYVAPEVLRKRYGPEADVWSAGVILYILLSGVPPFWAENEQGIFEQV 326




 Score = 45.1 bits (105), Expect(2) = 2e-16
 Identities = 23/49 (46%), Positives = 31/49 (62%), Gaps = 3/49 (6%)
```

```
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKY---DAFSYKQGS 235
           H  G++HRD+KPEN +F D+   S  K IDFGLS+ +   D F+   GS
Sbjct: 230 HSLGVMHRDLKPENFLFVDQKEDSLLKAIDFGLSIFFKPGDRFTDVVGS 278



>gi|12484155|gb|AAG53994.1|AF333959_1  (AF333959) calmodulin-domain protein kinase
2 [Toxoplasma gondii]
          Length = 520

 Score = 60.5 bits (145), Expect(2) = 2e-16
 Identities = 22/43 (51%), Positives = 33/43 (76%)
 Frame = -1

Query: 214 LKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPF 86
           ++   GTP Y++P+V+EG Y  +CDVWS GV+++ +L GYPPF
Sbjct: 203 MRTRAGTPYYVSPQVLEGRYGPECDVWSAGVMMYILLCGYPPF 245



 Score = 44.7 bits (104), Expect(2) = 2e-16
 Identities = 19/40 (47%), Positives = 26/40 (64%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKY 262
           AY H   + HRD+KPEN +F   +  S  KLIDFGL+ ++
Sbjct: 158 AYCHAHRVAHRDLKPENFLFLHDNPESPIKLIDFGLAARF 197



>gi|4325074|gb|AAD17247.1|  (AF118101) protein kinase 6 [Toxoplasma gondii]
          Length = 482

 Score = 60.5 bits (145), Expect(2) = 2e-16
 Identities = 22/43 (51%), Positives = 33/43 (76%)
 Frame = -1

Query: 214 LKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPF 86
           ++   GTP Y++P+V+EG Y  +CDVWS GV+++ +L GYPPF
Sbjct: 165 MRTRAGTPYYVSPQVLEGRYGPECDVWSAGVMMYILLCGYPPF 207



 Score = 44.7 bits (104), Expect(2) = 2e-16
 Identities = 19/40 (47%), Positives = 26/40 (64%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKY 262
           AY H   + HRD+KPEN +F   +  S  KLIDFGL+ ++
Sbjct: 120 AYCHAHRVAHRDLKPENFLFLHDNPESPIKLIDFGLAARF 159
```

```
>gi|15223870|ref|NP_172341.1|  (NM_100738) putative calcium-dependent protein
kinase (U90439) [Arabidopsis thaliana]
 gi|6318613|gb|AAF06968.1|AF162660_1  (AF162660) phosphoenolpyruvate carboxylase
kinase [Arabidopsis thaliana]
 gi|17644151|gb|AAL38773.1|  (AY065297) putative calcium-dependent protein kinase
[Arabidopsis thaliana]
          Length = 284

 Score = 67.4 bits (163), Expect(2) = 2e-16
 Identities = 30/57 (52%), Positives = 41/57 (71%), Gaps = 1/57 (1%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEG-TYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV+ G +Y EK D+WS GV+L+ ML G PPFYG   +E++E +   N+
Sbjct: 176 VGTPYYVAPEVLMGYSYGEKVDLWSAGVVLYTMLAGTPPFYGETAEEIFEAVLRGNL 232




 Score = 37.7 bits (86), Expect(2) = 2e-16
 Identities = 17/39 (43%), Positives = 27/39 (68%), Gaps = 1/39 (2%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEP-KLIDFGLSV 268
           ++ H  G++HRDIKPENI+    D+R++  K+ DFG  +
Sbjct: 129 SHCHRYGVVHRDIKPENILV---DLRNDTVKICDFGSGI 164




>gi|7434322|pir||T00718  calcium-dependent protein kinase homolog F22013.13 -
Arabidopsis thaliana
 gi|9802556|gb|AAF99758.1|AC003981_8  (AC003981) F22013.13 [Arabidopsis thaliana]
          Length = 280

 Score = 67.4 bits (163), Expect(2) = 2e-16
 Identities = 30/57 (52%), Positives = 41/57 (71%), Gaps = 1/57 (1%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEG-TYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV+ G +Y EK D+WS GV+L+ ML G PPFYG   +E++E +   N+
Sbjct: 176 VGTPYYVAPEVLMGYSYGEKVDLWSAGVVLYTMLAGTPPFYGETAEEIFEAVLRGNL 232




 Score = 37.7 bits (86), Expect(2) = 2e-16
 Identities = 17/39 (43%), Positives = 27/39 (68%), Gaps = 1/39 (2%)
 Frame = -2
```

```
Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEP-KLIDFGLSV 268
           ++ H  G++HRDIKPENI+    D+R++  K+ DFG  +
Sbjct: 129 SHCHRYGVVHRDIKPENILV---DLRNDTVKICDFGSGI 164




>gi|15238528|ref|NP_197831.1|  (NM_122351) calcium dependent protein kinase-like
protein
           [Arabidopsis thaliana]
 gi|10177904|dbj|BAB11236.1|  (AB016884) calcium dependent protein kinase-like
protein [Arabidopsis thaliana]
          Length = 594

 Score = 52.8 bits (125), Expect(2) = 2e-16
 Identities = 22/62 (35%), Positives = 38/62 (60%)
 Frame = -1

Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQY 44
           + +L   VG+  Y+APEV+  +Y+ + D+WS+GV+ + +L G  PFYGR    ++  +
Sbjct: 302 DQRLNDVVGSAYYVAPEVLHRSYSTEADMWSIGVISYILLCGSRPFYGRTESAIFRCVLR 361

Query: 43  QN 38
            N
Sbjct: 362 AN 363




 Score = 52.0 bits (123), Expect(2) = 2e-16
 Identities = 23/43 (53%), Positives = 33/43 (76%), Gaps = 2/43 (4%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS--VKYD 259
           A+ H QG++HRD+KPEN +FT R+  +  K+IDFGLS  ++YD
Sbjct: 260 AFFHLQGVVHRDLKPENFLFTSRNEDAILKVIDFGLSDFIRYD 302




>gi|17975557|ref|NP_524622.1|  (NM_079883) Calcium/calmodulin-dependent protein
kinase I [Drosophila melanogaster]
 gi|3893099|emb|CAA76937.1|  (Y17917) calcium/calmodulin dependent protein
kinase I [Drosophila melanogaster]
 gi|7304311|gb|AAF59343.1|  (AE003844) CaMKI gene product [alt 2] [Drosophila
melanogaster]
          Length = 405

 Score = 55.8 bits (133), Expect(2) = 2e-16
 Identities = 28/64 (43%), Positives = 41/64 (63%), Gaps = 2/64 (3%)
 Frame = -1

Query: 235 SQYQNSKLKA-GVGTPVYLAPEVI-EGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQEL 62
           S+ ++S + A   GTP Y+APEV+ +  Y +  DVWS+GV+ + +L GYPPFY  N   L
Sbjct: 193 SKMEDSGIMATACGTPGYVAPEVLAQKPYGKAVDVWSIGVISYILLCGYPPFYDENDANL 252
```

```
Query: 61  YENI 50
           +  I
Sbjct: 253 FAQI 256
```

```
 Score = 48.9 bits (115), Expect(2) = 2e-16
 Identities = 20/41 (48%), Positives = 30/41 (72%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKYDA 256
           Y+HEQG++HRD+KPEN+++    D  S+   + DFGLS   D+
Sbjct: 158 YMHEQGVVHRDLKPENLLYYSPDDDSKIMISDFGLSKMEDS 198
```

```
>gi|6625812|gb|AAF19403.1|AF203481_1  (AF203481) phosphoenolpyruvate carboxylase
kinase [Lycopersicon esculentum]
          Length = 279

 Score = 65.1 bits (157), Expect(2) = 2e-16
 Identities = 31/57 (54%), Positives = 39/57 (68%), Gaps = 1/57 (1%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV+ G  YNEK DVWS GV+L+ ML G PPFYG    E ++ +   N+
Sbjct: 169 VGTPYYVAPEVLMGKEYNEKVDVWSAGVILYIMLSGVPPFYGETPTETFQAVLRGNL 225
```

```
 Score = 39.7 bits (91), Expect(2) = 2e-16
 Identities = 17/35 (48%), Positives = 23/35 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           +Y H  G+ HRDIKP+N++F   D   +  KL DFG
Sbjct: 121 SYCHHMGVAHRDIKPDNVLF---DSENRLKLADFG 152
```

```
>gi|6625810|gb|AAF19402.1|AF203480_1  (AF203480) phosphoenolpyruvate carboxylase
kinase [Lycopersicon esculentum]
          Length = 276

 Score = 65.1 bits (157), Expect(2) = 2e-16
 Identities = 31/57 (54%), Positives = 39/57 (68%), Gaps = 1/57 (1%)
 Frame = -1
```

```
Query: 202 VGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV+ G  YNEK DVWS GV+L+ ML G PPFYG    E ++ +   N+
Sbjct: 166 VGTPYYVAPEVLMGKEYNEKVDVWSAGVILYIMLSGVPPFYGETPTETFQAVLRGNL 222




 Score = 39.7 bits (91), Expect(2) = 2e-16
 Identities = 17/35 (48%), Positives = 23/35 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           +Y H  G+ HRDIKP+N++F   D  +  KL DFG
Sbjct: 118 SYCHHMGVAHRDIKPDNVLF---DSENRLKLADFG 149



>gi|16215471|emb|CAC82999.1|  (AJ344155) calcium-dependent protein kinase 3
[Nicotiana tabacum]
          Length = 578

 Score = 59.3 bits (142), Expect(2) = 3e-16
 Identities = 23/51 (45%), Positives = 35/51 (68%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV+   Y  + DVWS GV+++ +L G PPF+   N Q ++E +
Sbjct: 276 VGSPYYVAPEVLRKRYGPEADVWSAGVIIYILLSGVPPFWAENEQGIFEQV 326




 Score = 45.1 bits (105), Expect(2) = 3e-16
 Identities = 23/49 (46%), Positives = 31/49 (62%), Gaps = 3/49 (6%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVKY---DAFSYKQGS 235
           H  G++HRD+KPEN +F D+   S  K IDFGLS+ +   D F+   GS
Sbjct: 230 HSLGVMHRDLKPENFLFVDQKEDSLLKAIDFGLSIFFKPGDRFTDVVGS 278



>gi|9294489|dbj|BAB02708.1|  (AB019230) IRE homolog; protein kinase-like protein
[Arabidopsis thaliana]
          Length = 1398

 Score = 64.7 bits (156), Expect(2) = 3e-16
 Identities = 28/69 (40%), Positives = 46/69 (66%), Gaps = 1/69 (1%)
 Frame = -1

Query: 253  QLQVRKSQYQNSKLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGR 77
            +L   + Q +  K ++ VGTP YLAPE++ GT +    D WS+G++LF ++VG PPF
Sbjct: 1158 RLAASEEQLERRKKRSAVGTPDYLAPEILLGTGHGATADWWSVGIILFELIVGIPPFNAE 1217
```

```
Query: 76   NRQELYENI 50
            + Q++++NI
Sbjct: 1218 HPQQIFDNI 1226
```

```
 Score = 39.3 bits (90), Expect(2) = 3e-16
 Identities = 16/36 (44%), Positives = 24/36 (66%)
 Frame = -2

Query: 378  YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
            Y+H +G++HRD+KP+N++          KL DFGLS
Sbjct: 1098 YLHSEGVVHRDLKPDNLLIAH---DGHIKLTDFGLS 1130
```

```
>gi|15229535|ref|NP_188412.1|  (NM_112666) protein kinase, putative [Arabidopsis
thaliana]
          Length = 1263

 Score = 64.7 bits (156), Expect(2) = 3e-16
 Identities = 28/69 (40%), Positives = 46/69 (66%), Gaps = 1/69 (1%)
 Frame = -1

Query: 253  QLQVRKSQYQNSKLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGR 77
            +L   + Q +  K ++ VGTP YLAPE++ GT +    D WS+G++LF ++VG PPF
Sbjct: 1063 RLAASEEQLERRKKRSAVGTPDYLAPEILLGTGHGATADWWSVGIILFELIVGIPPFNAE 1122

Query: 76   NRQELYENI 50
            + Q++++NI
Sbjct: 1123 HPQQIFDNI 1131
```

```
 Score = 39.3 bits (90), Expect(2) = 3e-16
 Identities = 16/36 (44%), Positives = 24/36 (66%)
 Frame = -2

Query: 378  YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
            Y+H +G++HRD+KP+N++          KL DFGLS
Sbjct: 1003 YLHSEGVVHRDLKPDNLLIAH---DGHIKLTDFGLS 1035
```

```
>gi|6729348|dbj|BAA89784.1|  (AB037134) IRE homolog 1 [Arabidopsis thaliana]
          Length = 1023

 Score = 64.7 bits (156), Expect(2) = 3e-16
 Identities = 28/69 (40%), Positives = 46/69 (66%), Gaps = 1/69 (1%)
 Frame = -1
```

```
Query: 253 QLQVRKSQYQNSKLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGR 77
           +L   + Q +  K ++ VGTP YLAPE++ GT +    D WS+G++LF ++VG PPF
Sbjct: 782 RLAASEEQLERRKKRSAVGTPDYLAPEILLGTGHGATADWWSVGIILFELIVGIPPFNAE 841

Query: 76  NRQELYENI 50
           + Q++++NI
Sbjct: 842 HPQQIFDNI 850




 Score = 39.3 bits (90), Expect(2) = 3e-16
 Identities = 16/36 (44%), Positives = 24/36 (66%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+H +G++HRD+KP+N++          KL DFGLS
Sbjct: 722 YLHSEGVVHRDLKPDNLLIAH---DGHIKLTDFGLS 754



>gi|19111951|ref|NP_595159.1|  (NC_003423) camp-dependent protein kinase catalytic
subunit [Schizosaccharomyces pombe]
 gi|729868|sp|P40376|KAPB_SCHPO  cAMP-dependent protein kinase catalytic subunit
 gi|626070|pir||A54400  protein kinase (EC 2.7.1.37), cAMP-dependent, catalytic
chain - fission yeast (Schizosaccharomyces pombe)
 gi|473894|gb|AAA70165.1|  (U08622) cAMP-dependent protein kinase
[Schizosaccharomyces pombe]
 gi|484305|dbj|BAA04891.1|  (D23667) catalytic subunit of the cAMP-dependent
protein kinase [Schizosaccharomyces pombe]
 gi|5817276|emb|CAB53726.1|  (AL110295) camp-dependent protein kinase catalytic
subunit [Schizosaccharomyces pombe]
 gi|1093611|prf||2104277A  cAMP-dependent protein kinase [Schizosaccharomyces
pombe]
          Length = 512

 Score = 70.1 bits (170), Expect(2) = 4e-16
 Identities = 32/51 (62%), Positives = 37/51 (71%), Gaps = 1/51 (1%)
 Frame = -1

Query: 199 GTPVYLAPEVIE-GTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           GTP YLAPE+I    YN+  D WSLG+L+F ML GYPPFY  N  +LYENI
Sbjct: 359 GTPDYLAPEIISLKPYNKAADWWSLGILIFEMLAGYPPFYSENPMKLYENI 409




 Score = 33.9 bits (76), Expect(2) = 4e-16
 Identities = 14/36 (38%), Positives = 23/36 (63%)
 Frame = -2
```

```
Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+H   I++RD+KPEN++    D     K++DFG +
Sbjct: 314 YLHHNQIVYRDLKPENLLL---DRFGHLKIVDFGFA 346



>gi|3282250|gb|AAC24961.1|  (AF009337) CDPK-related protein kinase [Tradescantia
virginiana]
          Length = 415

 Score = 54.3 bits (129), Expect(2) = 4e-16
 Identities = 23/37 (62%), Positives = 29/37 (78%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           AY H QG++HRD+KPEN +F +RD   +  KLIDFGLS
Sbjct: 78  AYCHLQGVVHRDLKPENFLFINRDENAPMKLIDFGLS 114




 Score = 49.7 bits (117), Expect(2) = 4e-16
 Identities = 19/58 (32%), Positives = 37/58 (63%)
 Frame = -1

Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + +L   VG+  Y+APEV+  +Y+ + D+WS+GV+ + +L G  PF+ R    ++ ++
Sbjct: 120 DERLNDIVGSAYYVAPEVLHRSYSMEADIWSIGVITYILLCGSRPFWARTESGIFRSV 177



>gi|7446432|pir||T02033  calcium/calmodulin-dependent protein kinase homolog -
maize
 gi|1839597|gb|AAB47181.1|  (S82324) calcium/calmodulin-dependent protein kinase
homolog|CaM kinase homolog|MCK1 [Zea mays]
          Length = 625

 Score = 53.9 bits (128), Expect(2) = 5e-16
 Identities = 23/37 (62%), Positives = 29/37 (78%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H QG++HRD+KPEN +FT RD   +  KLIDFGLS
Sbjct: 291 AFCHLQGVVHRDLKPENFLFTTRDESAPMKLIDFGLS 327




 Score = 49.7 bits (117), Expect(2) = 5e-16
 Identities = 19/58 (32%), Positives = 37/58 (63%)
 Frame = -1
```

```
Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + +L   VG+  Y+APEV+  +Y+ + D+WS+GV+ + +L G  PF+ R    ++ ++
Sbjct: 333 DERLNDIVGSAYYVAPEVLHRSYSMEADIWSIGVITYILLCGSRPFWARTESGIFRSV 390
```

```
>gi|7446433|pir||T02994  CDPK-related protein kinase - maize (fragment)
 gi|1313907|dbj|BAA12691.1|  (D84507) CDPK-related protein kinase [Zea mays]
          Length = 599

 Score = 53.9 bits (128), Expect(2) = 5e-16
 Identities = 23/37 (62%), Positives = 29/37 (78%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H QG++HRD+KPEN +FT RD  +  KLIDFGLS
Sbjct: 265 AFCHLQGVVHRDLKPENFLFTTRDESAPMKLIDFGLS 301



 Score = 49.7 bits (117), Expect(2) = 5e-16
 Identities = 19/58 (32%), Positives = 37/58 (63%)
 Frame = -1

Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + +L   VG+  Y+APEV+  +Y+ + D+WS+GV+ + +L G  PF+ R    ++ ++
Sbjct: 307 DERLNDIVGSAYYVAPEVLHRSYSMEADIWSIGVITYILLCGSRPFWARTESGIFRSV 364
```

```
>gi|12039322|gb|AAG46110.1|AC073166_8  (AC073166) calcium-dependent protein kinase
[Oryza sativa]
          Length = 534

 Score = 58.5 bits (140), Expect(2) = 5e-16
 Identities = 22/51 (43%), Positives = 35/51 (68%)
 Frame = -1


Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV+E  Y  + D+W+ GV+L+ +L G PPF+   +  +YE +
Sbjct: 235 VGSPYYIAPEVLEKRYGPEADIWTAGVILYVLLTGVPPFWADTQSGIYEKV 285



 Score = 45.1 bits (105), Expect(2) = 5e-16
 Identities = 19/35 (54%), Positives = 23/35 (65%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +  D D     K IDFGLSV
Sbjct: 189 HSHGVMHRDLKPENFLLLDADDEFSVKAIDFGLSV 223
```

```
>gi|1705733|sp|P53682|CDP1_ORYSA  Calcium-dependent protein kinase, isoform 1 (CDPK
1)
 gi|542156|pir||JC1515  calcium-dependent protein kinase (EC 2.7.1.-) - rice
 gi|435466|dbj|BAA02698.1|  (D13436) calcium-dependent protein kinase [Oryza sativa
(japonica cultivar-group)]
          Length = 534

 Score = 58.5 bits (140), Expect(2) = 5e-16
 Identities = 22/51 (43%), Positives = 35/51 (68%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV+E  Y  + D+W+ GV+L+ +L G PPF+   +  +YE +
Sbjct: 235 VGSPYYIAPEVLEKRYGPEADIWTAGVILYVLLTGVPPFWADTQSGIYEKV 285



 Score = 45.1 bits (105), Expect(2) = 5e-16
 Identities = 19/35 (54%), Positives = 23/35 (65%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +  D D     K IDFGLSV
Sbjct: 189 HSHGVMHRDLKPENFLLLDADDEFSVKAIDFGLSV 223



>gi|14484895|gb|AAK62812.1|  (AF276999) calcium-dependent protein kinase [Funaria
hygrometrica]
          Length = 518

 Score = 55.1 bits (131), Expect(2) = 5e-16
 Identities = 23/51 (45%), Positives = 34/51 (66%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           VG+P Y+APEV++  Y  + DVWS GV+L+ +L G PPF+    Q + + I
Sbjct: 202 VGSPYYMAPEVLKRNYGPEVDVWSAGVILYILLCGVPPFWAETEQGVAQAI 252



 Score = 48.5 bits (114), Expect(2) = 5e-16
 Identities = 20/35 (57%), Positives = 27/35 (77%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H QG++HRD+KPEN +F +++  S  K IDFGLSV
Sbjct: 156 HRQGVMHRDLKPENFLFANKNENSPLKAIDFGLSV 190
```

```
>gi|7446430|pir||T01694  calcium-dependent protein kinase-related kinase - maize
(fragment)
 gi|2443388|dbj|BAA22410.1|  (D38452) calcium-dependent protein kinase-related
kinase [Zea mays]
          Length = 452

 Score = 53.9 bits (128), Expect(2) = 5e-16
 Identities = 23/37 (62%), Positives = 29/37 (78%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H QG++HRD+KPEN +FT RD  +  KLIDFGLS
Sbjct: 118 AFCHLQGVVHRDLKPENFLFTTRDESAPMKLIDFGLS 154



 Score = 49.7 bits (117), Expect(2) = 5e-16
 Identities = 19/58 (32%), Positives = 37/58 (63%)
 Frame = -1

Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           + +L   VG+  Y+APEV+  +Y+ + D+WS+GV+ + +L G  PF+ R    ++ ++
Sbjct: 160 DERLNDIVGSAYYVAPEVLHRSYSMEADIWSIGVITYILLCGSRPFWARTESGIFRSV 217



>gi|16902296|dbj|BAB71853.1|  (AB065100) phosphoenolpyruvate carboxylase kinase
[Flaveria trinervia]
          Length = 281

 Score = 64.3 bits (155), Expect(2) = 5e-16
 Identities = 30/57 (52%), Positives = 40/57 (69%), Gaps = 1/57 (1%)
 Frame = -1

Query: 202 VGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQYQNV 35
           VGTP Y+APEV+ G  YNEK DVWS GV+L+ ML G PPF+G +   + +E +   N+
Sbjct: 171 VGTPYYVAPEVLSGMEYNEKVDVWSAGVILYIMLAGVPPFHGDSPADTFEAVLRGNL 227



 Score = 39.3 bits (90), Expect(2) = 5e-16
 Identities = 18/35 (51%), Positives = 23/35 (65%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFG 277
           +Y H  GI HRD+KP+N++F   D R   KL DFG
Sbjct: 123 SYCHRLGIAHRDLKPDNVLF---DSRGGLKLADFG 154
```

```
>gi|15236560|ref|NP_194096.1|  (NM_118496) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|2129550|pir||S71197  calcium-dependent protein kinase (EC 2.7.1.-) CDPK6 -
Arabidopsis thaliana
 gi|2129554|pir||S71901  calcium-dependent protein kinase 6 - Arabidopsis thaliana
 gi|7428013|pir||S71774  calcium-dependent protein kinase (EC 2.7.1.-) 6 -
Arabidopsis thaliana
 gi|836940|gb|AAA67654.1|  (U20623) calcium-dependent protein kinase [Arabidopsis
thaliana]
 gi|836944|gb|AAA67656.1|  (U20625) calcium-dependent protein kinase [Arabidopsis
thaliana]
 gi|4454034|emb|CAA23031.1|  (AL035394) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|7269213|emb|CAB79320.1|  (AL161559) calcium-dependent protein kinase (CDPK6)
[Arabidopsis thaliana]
 gi|14326514|gb|AAK60302.1|AF385710_1  (AF385710) AT4g23650/F9D16_120 [Arabidopsis
thaliana]
 gi|19548043|gb|AAL87385.1|  (AY081732) AT4g23650/F9D16_120 [Arabidopsis thaliana]
          Length = 529

 Score = 57.4 bits (137), Expect(2) = 6e-16
 Identities = 24/56 (42%), Positives = 37/56 (65%)
 Frame = -1


Query: 217 KLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           K K  VG+  Y+APEV++  Y  + D+WS GV+L+ +L G PPF+G N   +++ I
Sbjct: 235 KFKDLVGSAYYVAPEVLKRNYGPEADIWSAGVILYILLSGVPPFWGENETGIFDAI 290




 Score = 45.8 bits (107), Expect(2) = 6e-16
 Identities = 19/35 (54%), Positives = 24/35 (68%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +F  +D  S  K  DFGLSV
Sbjct: 194 HSMGVMHRDLKPENFLFLSKDENSPLKATDFGLSV 228




>gi|15450437|gb|AAK96512.1|  (AY052319) AT4g23650/F9D16_120 [Arabidopsis thaliana]
 gi|17386100|gb|AAL38596.1|AF446863_1  (AF446863) AT4g23650/F9D16_120 [Arabidopsis
thaliana]
          Length = 529

 Score = 57.4 bits (137), Expect(2) = 6e-16
 Identities = 24/56 (42%), Positives = 37/56 (65%)
 Frame = -1
```

```
Query: 217 KLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           K K  VG+  Y+APEV++  Y  + D+WS GV+L+ +L G PPF+G N    +++ I
Sbjct: 235 KFKDLVGSAYYVAPEVLKRNYGPEADIWSAGVILYILLSGVPPFWGENETGIFDAI 290




 Score = 45.8 bits (107), Expect(2) = 6e-16
 Identities = 19/35 (54%), Positives = 24/35 (68%)
 Frame = -2

Query: 372 HEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSV 268
           H  G++HRD+KPEN +F  +D  S  K  DFGLSV
Sbjct: 194 HSMGVMHRDLKPENFLFLSKDENSPLKATDFGLSV 228



>gi|1730055|sp|P25323|KMLC_DICDI  Myosin light chain kinase (MLCK)
 gi|1498250|gb|AAB06337.1|  (M64176) myosin light chain kinase [Dictyostelium
discoideum]
          Length = 295

 Score = 62.0 bits (149), Expect(2) = 6e-16
 Identities = 27/64 (42%), Positives = 42/64 (65%), Gaps = 1/64 (1%)
 Frame = -1

Query: 226 QNSKLKAGVGTPVYLAPEVIEGT-YNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENI 50
           Q   ++   GTP Y+APEV+  T Y+++ D+WS+GV+ + +L G+PPFYG    E++E I
Sbjct: 160 QTLVMQTACGTPSYVAPEVLNATGYDKEVDMWSIGVITYILLCGFPPFYGDTVPEIFEQI 219

Query: 49  QYQN 38
               N
Sbjct: 220 MEAN 223




 Score = 41.2 bits (95), Expect(2) = 6e-16
 Identities = 16/36 (44%), Positives = 24/36 (66%)
 Frame = -2

Query: 378 YIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           Y+H   I+HRD+KPEN++   ++   E  + DFGLS
Sbjct: 120 YLHGLNIVHRDLKPENLLLKSKENHLEVAIADFGLS 155



>gi|17542684|ref|NP_501900.1|  (NM_069499) calcium/calmodulin dependent protein
kinase [Caenorhabditis elegans]
 gi|14530497|emb|CAA94244.2|  (Z70279) contains similarity to Pfam domain: PF00069
(Eukaryotic protein kinase domain), Score=309.7, E-value=1.1e-89, N=1~cDNA EST
```

```
yk213a11.5 comes from this gene~cDNA EST yk238f10.5 comes from this gene~cDNA EST
yk330e12.5 comes from this gene~cDNA E>
 gi|14530613|emb|CAC42359.1|  (AL023841) contains similarity to Pfam domain:
PF00069 (Eukaryotic protein kinase domain), Score=309.7, E-value=1.1e-89, N=1~cDNA
EST yk213a11.5 comes from this gene~cDNA EST yk238f10.5 comes from this gene~cDNA
EST yk330e12.5 comes from this gene~cDNA>
          Length = 720

 Score = 57.0 bits (136), Expect(2) = 8e-16
 Identities = 24/52 (46%), Positives = 38/52 (72%), Gaps = 1/52 (1%)
 Frame = -1

Query: 199 GTPVYLAPEVIE-GTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           GTP YL+PEV++   Y++  D+W+ GV+L+ +LVGYPPF+  ++   LY  I+
Sbjct: 173 GTPGYLSPEVLKKDPYSKPVDIWACGVILYILLVGYPPFWDEDQHRLYAQIK 224




 Score = 45.8 bits (107), Expect(2) = 8e-16
 Identities = 18/39 (46%), Positives = 27/39 (69%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVK 265
           AY H  GI+HRD+KPEN++   +   +  KL DFGL+++
Sbjct: 123 AYCHSNGIVHRDLKPENLLLASKAKGAAVKLADFGLAIE 161



>gi|7505788|pir||T23616  hypothetical protein K11E8.1c - Caenorhabditis elegans
          Length = 708

 Score = 57.0 bits (136), Expect(2) = 8e-16
 Identities = 24/52 (46%), Positives = 38/52 (72%), Gaps = 1/52 (1%)
 Frame = -1


Query: 199 GTPVYLAPEVIE-GTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQ 47
           GTP YL+PEV++   Y++  D+W+ GV+L+ +LVGYPPF+  ++   LY  I+
Sbjct: 161 GTPGYLSPEVLKKDPYSKPVDIWACGVILYILLVGYPPFWDEDQHRLYAQIK 212




 Score = 45.8 bits (107), Expect(2) = 8e-16
 Identities = 18/39 (46%), Positives = 27/39 (69%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLSVK 265
           AY H  GI+HRD+KPEN++   +   +  KL DFGL+++
Sbjct: 111 AYCHSNGIVHRDLKPENLLLASKAKGAAVKLADFGLAIE 149
```

```
>gi|15229129|ref|NP_190506.1|  (NM_114797) calcium dependent protein kinase - like
[Arabidopsis thaliana]
 gi|11346394|pir||T45842  calcium dependent protein kinase-like - Arabidopsis
thaliana
 gi|6723407|emb|CAB66416.1|  (AL132956) calcium dependent protein kinase-like
[Arabidopsis thaliana]
 gi|12324436|gb|AAG52176.1|AC012329_3  (AC012329) putative calcium dependent
protein kinase; 28698-25746 [Arabidopsis thaliana]
          Length = 594

 Score = 53.5 bits (127), Expect(2) = 8e-16
 Identities = 22/62 (35%), Positives = 38/62 (60%)
 Frame = -1

Query: 223 NSKLKAGVGTPVYLAPEVIEGTYNEKCDVWSLGVLLFNMLVGYPPFYGRNRQELYENIQY 44
           + +L   VG+   Y+APEV+   +Y+ + D+WS+GV+ + +L G  PFYGR    ++   +
Sbjct: 301 DQRLNDVVGSAYYVAPEVLHRSYSTEADIWSIGVISYILLCGSRPFYGRTESAIFRCVLR 360

Query: 43  QN 38
            N
Sbjct: 361 AN 362




 Score = 49.3 bits (116), Expect(2) = 8e-16
 Identities = 20/37 (54%), Positives = 29/37 (78%)
 Frame = -2

Query: 381 AYIHEQGIIHRDIKPENIMFTDRDIRSEPKLIDFGLS 271
           A+ H QG++HRD+KPEN +FT ++   +  K+IDFGLS
Sbjct: 259 AFFHLQGVVHRDLKPENFLFTSKNEDAVLKVIDFGLS 295




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

```
Number of Hits to DB: 216,470,125
Number of Sequences: 919285
Number of extensions: 4660936
Number of successful extensions: 29785
Number of sequences better than 10.0: 7570
Number of HSP's better than 10.0 without gapping: 22076
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 27541
length of database: 288,558,979
effective HSP length: 102
effective length of database: 194,791,909
effective search space used: 4675005816
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020955089-08388-19794
```

7.1.12 Query= **hy-12_17a150017**

```
          (500 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 66 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|9790083|ref|NP_062659.1|  (NM_019685) RuvB-like protein 1...   244   2e-64
gi|4506753|ref|NP_003698.1|  (NM_003707) TATA binding protei...   244   3e-64
gi|19115695|ref|NP_594783.1|  (NC_003424) Putative 3' to 5' ...   241   2e-63
gi|12004636|gb|AAG44127.1|AF218072_1  (AF218072) pontin [Xen...   241   2e-63
gi|20160820|dbj|BAB89760.1|  (AP003271) hypothetical protein...   237   3e-62
gi|18376072|emb|CAD21100.1|  (AL669989) probable RUVB-like p...   236   6e-62
gi|6320396|ref|NP_010476.1|  (NC_001136) RUVB-like protein, ...   236   1e-61
gi|7208771|emb|CAB76908.1|  (AJ276264) putative Ruv DNA-heli...   233   5e-61
gi|15242217|ref|NP_197625.1|  (NM_122138) Ruv DNA-helicase-l...   230   4e-60
gi|7299320|gb|AAF54514.1|  (AE003686) pontin gene product [a...   230   5e-60
gi|7299321|gb|AAF54515.1|  (AE003686) pontin gene product [a...   230   5e-60
gi|7243680|gb|AAF43411.1|AF233278_1  (AF233278) pontin [Dros...   228   2e-59
gi|17558290|ref|NP_505567.1|  (NM_073166) Yeast hypothetical...   210   6e-54
gi|19173260|ref|NP_597063.1|  (NC_003238) DNA helicase domai...   177   3e-44
gi|12004634|gb|AAG44126.1|AF218071_1  (AF218071) reptin [Xen...   119   1e-26
gi|19112432|ref|NP_595640.1|  (NC_003423) Putative 3' to 5' ...   119   2e-26
gi|6755382|ref|NP_035434.1|  (NM_011304) RuvB-like protein 2...   118   2e-26
gi|5730023|ref|NP_006657.1|  (NM_006666) RuvB-like 2; erythr...   118   2e-26
gi|4929561|gb|AAD34041.1|AF151804_1  (AF151804) CGI-46 prote...   117   4e-26
gi|20093447|ref|NP_613294.1|  (NC_003551) DNA helicase TIP49...   115   1e-25
gi|6325021|ref|NP_015089.1|  (NC_001148) RUVB-like protein, ...   115   2e-25
gi|15240788|ref|NP_201564.1|  (NM_126163) RuvB DNA helicase-...   114   4e-25
gi|11499401|ref|NP_070640.1|  (NC_000917) TBP-interacting pr...   114   5e-25
gi|17737635|ref|NP_524156.1|  (NM_079432) reptin [Drosophila...   113   9e-25
gi|17542510|ref|NP_501067.1|  (NM_068666) T22D1.10.p [Caenor...   110   5e-24
```

```
gi|18978219|ref|NP_579576.1|  (NC_003413) tbp-interacting pr...   108   2e-23
gi|15920793|ref|NP_376462.1|  (NC_003106) 452aa long hypothe...   108   2e-23
gi|15229623|ref|NP_190552.1|  (NM_114843) RuvB DNA helicase ...   107   4e-23
gi|14520579|ref|NP_126054.1|  (NC_000868) TBP-interacting pr...   107   4e-23
gi|18312811|ref|NP_559478.1|  (NC_003364) TBP-interacting pr...   107   5e-23
gi|14591557|ref|NP_143639.1|  (NC_000961) hypothetical prote...   106   9e-23
gi|19074927|ref|NP_586433.1|  (NC_003237) similarity to HYPO...   103   6e-22
gi|14600633|ref|NP_147150.1|  (NC_000854) TATA-binding prote...   102   2e-21
gi|15899195|ref|NP_343800.1|  (NC_002754) TATA binding prote...    95   3e-19
gi|17977678|ref|NP_524220.1|  (NM_079496) abstrakt; abstract...    35   0.32
gi|453389|gb|AAA73419.1|  (U05814) cytochrome C oxidase subu...    34   0.55
gi|6682808|dbj|BAA88886.1|  (AB018404) similar to vaccinia A...    34   0.72
gi|18311394|ref|NP_563328.1|  (NC_003366) RNA polymerase bet...    33   1.6
gi|11466188|ref|NP_066511.1|  (NC_002573) NADH dehydrogenase...    32   2.1
gi|7021194|dbj|BAA91402.1|  (AK000868) unnamed protein produ...    32   2.7
gi|267421|sp|P14629|XPG_XENLA  DNA-repair protein complement...    32   2.7
gi|16552402|dbj|BAB71301.1|  (AK056881) unnamed protein prod...    32   2.7
gi|14042323|dbj|BAB55198.1|  (AK027561) unnamed protein prod...    32   2.7
gi|17864498|ref|NP_524849.1|  (NM_080110) Ecdysone-inducible...    32   2.7
gi|14737288|ref|XP_041928.1|  (XM_041928) similar to unnamed...    32   2.7
gi|7295103|gb|AAF50429.1|  (AE003555) ImpE1 gene product [Dr...    32   2.7
gi|3152919|gb|AAC17172.1|  (AF065141) ABC transporter ATP-bi...    32   3.6
gi|6321327|ref|NP_011404.1|  (NC_001139) Nop seven associate...    32   3.6
gi|11023511|gb|AAG26468.1|AF285774_8  (AF285774) putative fl...    31   4.7
gi|6689106|emb|CAB65393.1|  (AJ235845) NADH dehydrogenase, p...    31   4.7
gi|6325354|ref|NP_015422.1|  (NC_001148) Hypothetical ORF; Y...    31   4.7
gi|15964188|ref|NP_384541.1|  (NC_003047) CONSERVED HYPOTHET...    31   4.7
gi|15677717|ref|NP_274878.1|  (NC_003112) TonB-dependent rec...    31   6.1
gi|15232255|ref|NP_189412.1|  (NM_113691) unknown protein [A...    31   6.1
gi|17552634|ref|NP_498077.1|  (NM_065676) C45G9.11.p [Caenor...    31   6.1
gi|17510485|ref|NP_491069.1|  (NM_058668) coatomer, alpha ch...    31   6.1
gi|16799815|ref|NP_470083.1|  (NC_003212) probable cell surf...    31   6.1
gi|13325397|gb|AAH04503.1|AAH04503  (BC004503) Similar to Ru...    31   6.1
gi|1176784|sp|Q09282|YQIB_CAEEL  HYPOTHETICAL 39.1 KD PROTEI...    31   6.1
gi|15027651|ref|NP_149383.1|  (NC_003029) ymf65 [Tetrahymena...    30   8.0
gi|17557832|ref|NP_505637.1|  (NM_073236) PH (pleckstrin  ho...    30   8.0
gi|15925300|ref|NP_372834.1|  (NC_002758) hypothetical prote...    30   8.0
gi|18311723|ref|NP_558390.1|  (NC_003364) hypothetical prote...    30   8.0
gi|18310263|ref|NP_562197.1|  (NC_003366) hypothetical prote...    30   8.0
gi|13812027|ref|NP_113158.1|  (NC_002751) hypothetical prote...    30   8.0
gi|1480429|gb|AAC44464.1|  (U18943) putative transcriptional...    30   8.0
```

Alignments

```
>gi|9790083|ref|NP_062659.1|  (NM_019685) RuvB-like protein 1; Pontin 52 [Mus musculus]
 gi|7449960|pir||JC5521  TATA-binding protein-interacting protein 49 - rat
 gi|2225877|dbj|BAA20875.1|  (AB002406) TIP49 [Rattus norvegicus]
 gi|4106528|gb|AAD02877.1|  (AF100694) Pontin52 [Mus musculus]
 gi|4521276|dbj|BAA76313.1|  (AB001581) DNA helicase p50 [Rattus norvegicus]
 gi|13435708|gb|AAH04718.1|AAH04718  (BC004718) RuvB-like protein 1 [Mus musculus]
          Length = 456

 Score =  244 bits (624), Expect = 2e-64
 Identities = 118/151 (78%), Positives = 135/151 (89%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
            DPS  E+L +E+V  GDVIYIEANSG VKR GRCD Y +E+DLEAEEYVPLPKGDVHKKK
Sbjct: 173 DPSIFESLQKERVEAGDVIYIEANSGAVKRQGRCDTYATEFDLEAEEYVPLPKGDVHKKK 232
```

```
Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EI+QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLRGEINKVV+KYIDQGVA
Sbjct: 233 EIIQDVTLHDLDVANARPQGGQDILSMMGQLMKPKKTEITDKLRGEINKVVNKYIDQGVA 292

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+F+DE HMLDIE FT+L+RALES++
Sbjct: 293 ELVPGVLFVDEVHMLDIECFTYLHRALESSI 323



>gi|4506753|ref|NP_003698.1|  (NM_003707) TATA binding protein interacting protein 49
kDa; RuvB (E coli homolog)-like 1 [Homo sapiens]
 gi|7449961|pir||JE0334  nuclear matrix protein NMP 238 - human
 gi|3132308|dbj|BAA28169.1|  (AB012122) TIP49 [Homo sapiens]
 gi|3243035|gb|AAC77819.1|  (AF070735) RuvB-like protein RUVBL1 [Homo sapiens]
 gi|3892584|emb|CAA08986.1|  (AJ010058) Nuclear matrix protein NMP238 [Homo sapiens]
 gi|4151525|gb|AAD04427.1|  (AF099084) Pontin52 [Homo sapiens]
 gi|5327000|emb|CAB46271.1|  (Y18418) erythrocyte cytosolic protein of 54 kDa, ECP-54 [Homo
sapiens]
 gi|12653495|gb|AAH00519.1|AAH00519  (BC000519) RuvB (E coli homolog)-like 1 [Homo sapiens]
 gi|12804269|gb|AAH02993.1|  (BC002993) RuvB-like 1 (E. coli) [Homo sapiens]
 gi|15277588|gb|AAH12886.1|AAH12886  (BC012886) Similar to RuvB (E coli homolog)-like 1
[Homo sapiens]
          Length = 456

 Score =  244 bits (623), Expect = 3e-64
 Identities = 117/151 (77%), Positives = 135/151 (88%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS   E+L +E+V  GDVIYIEANSG VKR GRCD Y +E+DLEAEEYVPLPKGDVHKKK
Sbjct: 173 DPSIFESLQKERVEAGDVIYIEANSGAVKRQGRCDTYATEFDLEAEEYVPLPKGDVHKKK 232

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EI+QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLRGEINKVV+KYIDQG+A
Sbjct: 233 EIIQDVTLHDLDVANARPQGGQDILSMMGQLMKPKKTEITDKLRGEINKVVNKYIDQGIA 292

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+F+DE HMLDIE FT+L+RALES++
Sbjct: 293 ELVPGVLFVDEVHMLDIECFTYLHRALESSI 323



>gi|19115695|ref|NP_594783.1|  (NC_003424) Putative 3' to 5' DNA/RNA helicase involved in
ribosomal RNA processing; by similarity to S. cerevisiae RVB1 [Schizosaccharomyces pombe]
 gi|13810239|emb|CAC37428.1|  (AL590902) putative 3' to 5' DNA/RNA helicase; possibly
involved in ribosomal RNA processing; similar to S. cerevisiae RVB1 [Schizosaccharomyces
pombe]
          Length = 456

 Score =  241 bits (615), Expect = 2e-63
 Identities = 114/151 (75%), Positives = 137/151 (90%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS +E+L RE+V+ GDVIYIEAN+G VKRVGR D Y +E+DLEAEEYVP+PKG+VHK+K
Sbjct: 174 DPSIYESLQREQVSTGDVIYIEANTGAVKRVGRSDAYATEFDLEAEEYVPMPKGEVHKRK 233
```

```
Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDVTLHDLD+ANA+PQGG D +S+M QL K KKT+ITDKLRGEINKVV+KYI+QG+A
Sbjct: 234 EIVQDVTLHDLDIANARPQGGQDIMSMMGQLMKPKKTEITDKLRGEINKVVNKYIEQGIA 293

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           EL+PGV+FIDE HMLDIE FT+LN+ALEST+
Sbjct: 294 ELIPGVLFIDEVHMLDIECFTYLNQALESTI 324



>gi|12004636|gb|AAG44127.1|AF218072_1  (AF218072) pontin [Xenopus laevis]
          Length = 456

 Score =  241 bits (615), Expect = 2e-63
 Identities = 115/151 (76%), Positives = 136/151 (89%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS +E+L +E+V +GDVIYIEANSG VKR GR D Y +E+DLEAEEYVPLPKGDVH+KK
Sbjct: 173 DPSIYESLQKERVEVGDVIYIEANSGAVKRQGRSDTYATEFDLEAEEYVPLPKGDVHQKK 232

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           E++QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLRG+INKVV+KYIDQG+A
Sbjct: 233 EVIQDVTLHDLDVANARPQGGQDILSMMGQLMKPKKTEITDKLRGQINKVVNKYIDQGIA 292

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLDIE FT+L+RALES+L
Sbjct: 293 ELVPGVLFIDEVHMLDIECFTYLHRALESSL 323



>gi|20160820|dbj|BAB89760.1|  (AP003271) hypothetical protein~similar toRuv DNA-helicase
[Oryza sativa (japonica cultivar-group)]
          Length = 448

 Score =  237 bits (605), Expect = 3e-62
 Identities = 113/151 (74%), Positives = 134/151 (87%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DP+ ++ L +EKV +GDVIYIEANSG VKRVGRCD + +EYDLEAEEYVP+PKG+VHKKK
Sbjct: 204 DPTIYDALIKEKVAVGDVIYIEANSGAVKRVGRCDSFATEYDLEAEEYVPIPKGEVHKKK 263

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDVTLHDLD ANA+PQGG D +SLM Q+ K +KT+ITDKLR EINKVV++YID+G+A
Sbjct: 264 EIVQDVTLHDLDAANAQPQGGQDILSLMGQMMKPRKTEITDKLRQEINKVVNRYIDEGIA 323

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLDIE F++LNRALES L
Sbjct: 324 ELVPGVLFIDEVHMLDIECFSYLNRALESPL 354



>gi|18376072|emb|CAD21100.1|  (AL669989) probable RUVB-like protein [Neurospora crassa]
          Length = 458

 Score =  236 bits (603), Expect = 6e-62
```

```
 Identities = 113/151 (74%), Positives = 133/151 (87%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS +E + +E+VT+GDVIYIEAN+G  KRVGR D Y +E+DLEAEEYVP+PKG+VHKKK
Sbjct: 174 DPSIYEAIQKERVTVGDVIYIEANTGACKRVGRSDAYATEFDLEAEEYVPIPKGEVHKKK 233

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDV+LHDLDVANA+PQGG D +S+M QL K K T+ITDKLR EINKVVSKYIDQGVA
Sbjct: 234 EIVQDVSLHDLDVANARPQGGQDIMSMMGQLMKPKMTEITDKLRSEINKVVSKYIDQGVA 293

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD+E FT+LN+ALES +
Sbjct: 294 ELVPGVLFIDEAHMLDVECFTYLNKALESPI 324




>gi|6320396|ref|NP_010476.1|  (NC_001136) RUVB-like protein, TIP49a Homologue; Rvb1p
[Saccharomyces cerevisiae]
 gi|1077539|pir||S52698  hypothetical protein YDR190c - yeast (Saccharomyces cerevisiae)
 gi|755784|emb|CAA88704.1|  (Z48784) unknown [Saccharomyces cerevisiae]
          Length = 463

 Score =  236 bits (601), Expect = 1e-61
 Identities = 113/151 (74%), Positives = 135/151 (88%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DP+ +E++ REKV+IGDVIYIEAN+G VKRVGR D Y +E+DLE EEYVPLPKG+VHKKK
Sbjct: 182 DPTIYESIQREKVSIGDVIYIEANTGAVKRVGRSDAYATEFDLETEEYVPLPKGEVHKKK 241

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDVTLHDLDVANA+PQGG D +S+M QL K KKT+IT+KLR E+NKVV+KYIDQGVA
Sbjct: 242 EIVQDVTLHDLDVANARPQGGQDVISMMGQLLKPKKTEITEKLRQEVNKVVAKYIDQGVA 301

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           EL+PGV+FIDE +MLDIE FT+LN+ALES +
Sbjct: 302 ELIPGVLFIDEVNMLDIEIFTYLNKALESNI 332




>gi|7208771|emb|CAB76908.1|  (AJ276264) putative Ruv DNA-helicase [Cicer arietinum]
          Length = 458

 Score =  233 bits (595), Expect = 5e-61
 Identities = 112/151 (74%), Positives = 134/151 (88%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DP+ ++ L +EKV +GDVIYIEANSG VKRVGR D + +E+DLEAEEYVPLPKG+VHKKK
Sbjct: 176 DPTIYDALIKEKVAVGDVIYIEANSGAVKRVGRSDAFATEFDLEAEEYVPLPKGEVHKKK 235

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDVTLHDLD ANA+PQGG D +SLM Q+ K +KT+ITDKLR EINKVV++YID+GVA
Sbjct: 236 EIVQDVTLHDLDAANARPQGGQDILSLMGQMMKPRKTEITDKLRQEINKVVNRYIDEGVA 295

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD+E F++LNRALES+L
Sbjct: 296 ELVPGVLFIDEVHMLDMECFSYLNRALESSL 326
```

```
>gi|15242217|ref|NP_197625.1|  (NM_122138) Ruv DNA-helicase-like protein [Arabidopsis
thaliana]
 gi|9757813|dbj|BAB08331.1|  (AB007651) Ruv DNA-helicase-like protein [Arabidopsis thaliana]
          Length = 458

 Score =  230 bits (587), Expect = 4e-60
 Identities = 111/151 (73%), Positives = 133/151 (87%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DP+ ++ L +EKV +GDVIYIEANSG VKRVGR D + +E+DLEAEEYVPLPKG+VHKKK
Sbjct: 176 DPTIYDALIKEKVAVGDVIYIEANSGAVKRVGRSDAFATEFDLEAEEYVPLPKGEVHKKK 235

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           EIVQDVTL DLD ANA+PQGG D +SLM Q+ K +KT+ITDKLR EINKVV++YID+GVA
Sbjct: 236 EIVQDVTLQDLDAANARPQGGQDILSLMGQMMKPRKTEITDKLRQEINKVVNRYIDEGVA 295

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD+E F++LNRALES+L
Sbjct: 296 ELVPGVLFIDEVHMLDMECFSYLNRALESSL 326



>gi|7299320|gb|AAF54514.1| ☐ (AE003686) pontin gene product [alt 2] [Drosophila
melanogaster]
 gi|16768848|gb|AAL28643.1| ☐ (AY061095) LD08555p [Drosophila melanogaster]
          Length = 456

 Score =  230 bits (586), Expect = 5e-60
 Identities = 110/151 (72%), Positives = 131/151 (85%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS  + L +EKV +GDVIYIEANSG VKR GR D + +E+DLE EEYVPLPKGDVHKKK
Sbjct: 173 DPSIFDALQKEKVEVGDVIYIEANSGAVKRQGRSDTFATEFDLETEEYVPLPKGDVHKKK 232

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           E++QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLR EINKVV+KYIDQG+A
Sbjct: 233 EVIQDVTLHDLDVANARPQGGQDVLSMMGQLMKPKKTEITDKLRMEINKVVNKYIDQGIA 292

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD+E FT+L+++LES +
Sbjct: 293 ELVPGVLFIDEIHMLDLETFTYLHKSLESPI 323



>gi|7299321|gb|AAF54515.1| ☐ (AE003686) pontin gene product [alt 1] [Drosophila
melanogaster]
          Length = 396

 Score =  230 bits (586), Expect = 5e-60
 Identities = 110/151 (72%), Positives = 131/151 (85%)
 Frame = -2
```

```
Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS  + L +EKV +GDVIYIEANSG VKR GR D + +E+DLE EEYVPLPKGDVHKKK
Sbjct: 113 DPSIFDALQKEKVEVGDVIYIEANSGAVKRQGRSDTFATEFDLETEEYVPLPKGDVHKKK 172

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           E++QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLR EINKVV+KYIDQG+A
Sbjct: 173 EVIQDVTLHDLDVANARPQGGQDVLSMMGQLMKPKKTEITDKLRMEINKVVNKYIDQGIA 232

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD+E FT+L+++LES +
Sbjct: 233 ELVPGVLFIDEIHMLDLETFTYLHKSLESPI 263




>gi|7243680|gb|AAF43411.1|AF233278_1  (AF233278) pontin [Drosophila melanogaster]
          Length = 456

 Score =  228 bits (582), Expect = 2e-59
 Identities = 109/151 (72%), Positives = 131/151 (86%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS  + L +EKV +GDVIYIEANSG VKR GR D + +E+DLE EEYVPLPKGDVHKKK
Sbjct: 173 DPSIFDALQKEKVEVGDVIYIEANSGAVKRQGRSDTFATEFDLETEEYVPLPKGDVHKKK 232

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           E++QDVTLHDLDVANA+PQGG D +S+M QL K KKT+ITDKLR EINKVV+KYIDQG+A
Sbjct: 233 EVIQDVTLHDLDVANARPQGGQDVLSMMGQLMKPKKTEITDKLRMEINKVVNKYIDQGIA 292

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FIDE HMLD++ FT+L+++LES +
Sbjct: 293 ELVPGVLFIDEIHMLDLKTFTYLHKSLESPI 323




>gi|17558290|ref|NP_505567.1|  (NM_073166) Yeast hypothetical 50.5 KD protein like
[Caenorhabditis elegans]
 gi|7496628|pir||T19534  hypothetical protein C27H6.2 - Caenorhabditis elegans
 gi|3874559|emb|CAB02793.1|  (Z81042) predicted using Genefinder~Similarity to Yeast
hypothetical 50.5 KD protein (TR:G755784)~cDNA EST EMBL:T01921 comes from this gene~cDNA EST
yk61b3.3 comes from this gene~cDNA EST EMBL:T01580 comes from this gene~cDNA EST yk81a11.3
comes from t>
          Length = 458

 Score =  210 bits (534), Expect = 6e-54
 Identities = 98/152 (64%), Positives = 131/152 (85%), Gaps = 1/152 (0%)
 Frame = -2

Query: 454 DPSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKK 275
           DPS ++++ +++V +GDVIYIEANSG VKRVGRCD+Y SE+DLEA+E+VP+PKGDV K K
Sbjct: 191 DPSIYDSILKQRVEVGDVIYIEANSGIVKRVGRCDVYASEFDLEADEFVPMPKGDVRKSK 250

Query: 274 EIVQDVTLHDLDVANAKPQGGH-DFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGV 98
           +IVQ+V+LHDLD+ANA+PQG   D  ++++QL   KKT++TD+LR EINKVV++YI+ GV
Sbjct: 251 DIVQNVSLHDLDIANARPQGRQGDVSNIVSQLMTPKKTEVTDRLRSEINKVVNEYIESGV 310

Query: 97  AELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           AEL+PGV+FIDE HMLD+E FT+L RALES +
Sbjct: 311 AELMPGVLFIDEVHMLDVECFTYLYRALESPM 342
```

```
>gi|19173260|ref|NP_597063.1|  (NC_003238) DNA helicase domain [Encephalitozoon cuniculi]
 gi|19171386|emb|CAD27111.1|  (AL590451) DNA helicase domain [Encephalitozoon cuniculi]
          Length = 426

 Score =  177 bits (450), Expect = 3e-44
 Identities = 78/150 (52%), Positives = 122/150 (81%)
 Frame = -2

Query: 451 PSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKE 272
           PS +E + ++++  GDV+YIE NSG +K++GR + + +++DLEA+ YVP+PKG+V K+KE
Sbjct: 166 PSLYEQIDKQRIVNGDVVYIEVNSGVIKKLGRSEAHMNDFDLEADTYVPIPKGEVLKRKE 225

Query: 271 IVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAE 92
           ++Q VTLHDLD+ANA+P  G D +SL+ ++   +KT+IT++LRG++N++V+ Y++ G AE
Sbjct: 226 VMQSVTLHDLDMANARP-SGQDMLSLVFRILSPRKTEITERLRGDVNRMVNGYLENGNAE 284

Query: 91  LVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           +VPGV+FIDE HMLD+E FTFL++ +ES L
Sbjct: 285 IVPGVLFIDEVHMLDVECFTFLHKVIESPL 314



>gi|12004634|gb|AAG44126.1|AF218071_1  (AF218071) reptin [Xenopus laevis]
          Length = 462

 Score =  119 bits (298), Expect = 1e-26
 Identities = 63/148 (42%), Positives = 100/148 (67%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E+LT+EKV  GDVI I+  +G + ++GR      +YD    +  +V  P G++ K+KE+V
Sbjct: 179 ESLTKEKVQAGDVITIDKATGKITKLGRAFTRARDYDAMGSQTKFVQCPDGELQKRKEVV 238

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LH++DV N++ QG   F++L +    + K+++    R +IN  V+++ ++G AE++
Sbjct: 239 HTVSLHEIDVINSRTQG---FLALFSGDTGEIKSEV----REQINAKVAEWREEGKAEII 291

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE F+FLNRALES +
Sbjct: 292 PGVLFIDEVHMLDIECFSFLNRALESDM 319



>gi|19112432|ref|NP_595640.1|  (NC_003423) Putative 3' to 5' DNA/RNA helicase involved in
ribosomal RNA processing and for snoRNA accumulation; by similarity to yeast rvb1
[Schizosaccharomyces pombe]
 gi|7493023|pir||T40697  probable tata binding protein-interacting protein - fission yeast
(Schizosaccharomyces pombe)
 gi|4455781|emb|CAB36870.1|  (AL035536) Putative 3' to 5' DNA/RNA helicase involved in
ribosomal RNA processing and for snoRNA accumulation; by similarity to yeast rvb1
[Schizosaccharomyces pombe]
          Length = 465

 Score =  119 bits (297), Expect = 2e-26
 Identities = 64/148 (43%), Positives = 98/148 (65%), Gaps = 2/148 (1%)
 Frame = -2
```

```
Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           ++LT+EKV  GDVI I+ + G V ++GR      +YD    +  +V  P+G++ K+KE+V
Sbjct: 174 DSLTKEKVLAGDVISIDKSVGRVTKLGRSFSRARDYDAMGADTRFVQCPQGEIQKRKEVV 233

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LHD+DV N++ QG   F++L +        +I  ++R +IN  VS++ ++G AE+V
Sbjct: 234 HTVSLHDIDVINSRTQG---FLALFSG----DTGEIKPEVREQINTKVSEWREEGKAEIV 286

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+F+DE HMLDIE F+F NRALE  L
Sbjct: 287 PGVLFVDEVHMLDIECFSFFNRALEDDL 314



>gi|6755382|ref|NP_035434.1| ☐ (NM_011304) RuvB-like protein 2 [Mus musculus]
 gi|4521249|dbj|BAA76297.1| ☐ (AB013912) DNA helicase [Mus musculus]
          Length = 463

 Score =  118 bits (296), Expect = 2e-26
 Identities = 62/148 (41%), Positives = 101/148 (67%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E+LT++KV  GDVI I+  +G + ++GR      +YD    +  +V  P G++ K+KE+V
Sbjct: 180 ESLTKDKVQAGDVITIDKATGKISKLGRSFTRARDYDAMGSQTKFVQCPDGELQKRKEVV 239

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LH++DV N++ QG   F++L +    + K+++    R +IN  V+++ ++G AE++
Sbjct: 240 HTVSLHEIDVINSRTQG---FLALFSGDTGEIKSEV----REQINAKVAEWREEGKAEII 292

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE+F+FLNRALES +
Sbjct: 293 PGVLFIDEVHMLDIESFSFLNRALESDM 320



>gi|5730023|ref|NP_006657.1| ☐ (NM_006666) RuvB-like 2; erythrocyte cytosolic protein, 51-
KD; TBP-interacting protein, 48-KD; Reptin52 [Homo sapiens]
 gi|11280157|pir||T46313  hypothetical protein DKFZp434K1011.1 - human
 gi|4587311|dbj|BAA76708.1| ☐ (AB024301) RuvB-like DNA helicase TIP49b [Homo sapiens]
 gi|5020422|gb|AAD38073.1|AF155138_1 ☐ (AF155138) RUVBL2 protein [Homo sapiens]
 gi|5326998|emb|CAB46270.1| ☐ (Y18417) erythrocyte cytosolic protein of 51 kDa, ECP-51
[Homo sapiens]
 gi|6807657|emb|CAB66677.1| ☐ (AL136743) hypothetical protein [Homo sapiens]
 gi|9367027|gb|AAF87087.1|AF124607_1 ☐ (AF124607) Reptin52 [Homo sapiens]
 gi|12653319|gb|AAH00428.1|AAH00428 ☐ (BC000428) RuvB (E coli homolog)-like 2 [Homo
sapiens]
 gi|13528657|gb|AAH04531.1|AAH04531 ☐ (BC004531) RuvB (E coli homolog)-like 2 [Homo
sapiens]
          Length = 463

 Score =  118 bits (296), Expect = 2e-26
 Identities = 62/148 (41%), Positives = 101/148 (67%), Gaps = 2/148 (1%)
 Frame = -2
```

```
Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E+LT++KV  GDVI I+  +G + ++GR      +YD    +  +V  P G++ K+KE+V
Sbjct: 180 ESLTKDKVQAGDVITIDKATGKISKLGRSFTRARDYDAMGSQTKFVQCPDGELQKRKEVV 239

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LH++DV N++ QG   F++L +    + K+++    R +IN  V+++ ++G AE++
Sbjct: 240 HTVSLHEIDVINSRTQG---FLALFSGDTGEIKSEV----REQINAKVAEWREEGKAEII 292

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE+F+FLNRALES +
Sbjct: 293 PGVLFIDEVHMLDIESFSFLNRALESDM 320
```

```
>gi|4929561|gb|AAD34041.1|AF151804_1  (AF151804) CGI-46 protein [Homo sapiens]
          Length = 442

 Score =  117 bits (294), Expect = 4e-26
 Identities = 62/148 (41%), Positives = 100/148 (66%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E+LT++KV  GDVI I+  +G + ++GR       YD    +  +V  P G++ K+KE+V
Sbjct: 180 ESLTKDKVQAGDVITIDKATGKISKLGRSFTRARNYDAMGSQTKFVQCPDGELQKRKEVV 239

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LH++DV N++ QG   F++L +    + K+++    R +IN  V+++ ++G AE++
Sbjct: 240 HTVSLHEIDVINSRTQG---FLALFSGDTGEIKSEV----REQINAKVAEWREEGKAEII 292

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE+F+FLNRALES +
Sbjct: 293 PGVLFIDEVHMLDIESFSFLNRALESDM 320
```

```
>gi|20093447|ref|NP_613294.1|  (NC_003551) DNA helicase TIP49, TBP-interacting protein
[Methanopyrus kandleri AV19]
 gi|19886266|gb|AAM01224.1|  (AE010302) DNA helicase TIP49, TBP-interacting protein
[Methanopyrus kandleri AV19]
          Length = 455

 Score =  115 bits (289), Expect = 1e-25
 Identities = 66/146 (45%), Positives = 92/146 (62%), Gaps = 2/146 (1%)
 Frame = -2

Query: 433 LTREKVTIGDVIYIEANSGNVKRVGRCD--IYNSEYDLEAEEYVPLPKGDVHKKKEIVQD 260
           L +  V  GDVI I+  SG+V ++GR    +   E +L     V LP+G V KKKEI +
Sbjct: 182 LVQAGVREGDVIQIDVESGHVTKLGRAKDALEEEEEELLGVHAVELPEGPVQKKKEIKRV 241

Query: 259 VTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPG 80
           VTLHDLD+AN +            +L   ++ +ITD++R ++++ V K +D+G A LVPG
Sbjct: 242 VTLHDLDMANVRA----------GRLLGFREEEITDEIRQKVDERVQKMVDEGEASLVPG 291

Query: 79  VIFIDECHMLDIEAFTFLNRALESTL 2
           V+FIDE HMLDIEAF FLNR+LE  +
Sbjct: 292 VLFIDEAHMLDIEAFAFLNRSLEEEI 317
```

```
>gi|6325021|ref|NP_015089.1|  (NC_001148) RUVB-like protein, TIP49b Homologue; Rvb2p
[Saccharomyces cerevisiae]
 gi|2132242|pir||S61029  hypothetical protein YPL235w - yeast (Saccharomyces cerevisiae)
 gi|1061254|emb|CAA91609.1|  (Z67751) putative protein [Saccharomyces cerevisiae]
 gi|1181253|emb|CAA64252.1|  (X94561) ATP/GTP binding site motif A (P-loop) [Saccharomyces
cerevisiae]
 gi|1370486|emb|CAA97952.1|  (Z73591) ORF YPL235w [Saccharomyces cerevisiae]
          Length = 471

 Score =  115 bits (288), Expect = 2e-25
 Identities = 61/145 (42%), Positives = 96/145 (66%), Gaps = 2/145 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           + LT+EKV  GDVI I+  SG + ++GR    + +YD    +  +V  P+G++ K+K +V
Sbjct: 177 DGLTKEKVLAGDVISIDKASGKITKLGRSFARSRDYDAMGADTRFVQCPEGELQKRKTVV 236

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             V+LH++DV N++ QG   F++L          +I  ++R +IN  V+++ ++G AE+V
Sbjct: 237 HTVSLHEIDVINSRTQG---FLALFTG----DTGEIRSEVRDQINTKVAEWKEEGKAEIV 289

Query: 85  PGVIFIDECHMLDIEAFTFLNRALE 11
           PGV+FIDE HMLDIE F+F+NRALE
Sbjct: 290 PGVLFIDEVHMLDIECFSFINRALE 314



>gi|15240788|ref|NP_201564.1|  (NM_126163) RuvB DNA helicase-like protein [Arabidopsis
thaliana]
 gi|9757884|dbj|BAB08471.1|  (AB013390) RuvB DNA helicase-like protein [Arabidopsis
thaliana]
 gi|16974568|gb|AAL31257.1|  (AY061754) AT5g67630/K9I9_20 [Arabidopsis thaliana]
          Length = 469

 Score =  114 bits (285), Expect = 4e-25
 Identities = 62/148 (41%), Positives = 98/148 (65%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E L +EKV  GDVI I+  +G + ++GR    + +YD    +  +V  P+G++ K+KE+V
Sbjct: 177 EALNKEKVQSGDVIAIDKATGKITKLGRSFSRSRDYDAMGAQTKFVQCPEGELQKRKEVV 236

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             VTLH++DV N++ QG   F++L          +I  ++R +I+  V+++ ++G AE+V
Sbjct: 237 HCVTLHEIDVINSRTQG---FLALFTG----DTGEIRSEVREQIDTKVAEWREEGKAEIV 289

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE F+FLNRALE+ +
Sbjct: 290 PGVLFIDEVHMLDIECFSFLNRALENEM 317



>gi|11499401|ref|NP_070640.1|  (NC_000917) TBP-interacting protein TIP49 [Archaeoglobus
fulgidus]
 gi|7449962|pir||D69476  TBP-interacting protein TIP49 homolog - Archaeoglobus fulgidus
 gi|2648730|gb|AAB89434.1|  (AE000977) TBP-interacting protein TIP49 [Archaeoglobus
fulgidus]
          Length = 449
```

```
 Score =  114 bits (284), Expect = 5e-25
 Identities = 57/139 (41%), Positives = 89/139 (64%)
 Frame = -2

Query: 418 VTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLD 239
           + +GDVI I+  +G + ++GR +    +YDL  +E VP+P G V K+KE    VTLHDLD
Sbjct: 185 IEVGDVIVIDKETGRIGKLGRSEKAKKKYDLGDDEVVPVPSGKVEKEKEFTYVVTLHDLD 244

Query: 238 VANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVIFIDEC 59
            ANA+        + +  L      +I +++R  +++ V + +++G AELVPGV+FIDE
Sbjct: 245 EANARR-------TSIFSLFSPPSREIDNEVREAVDEQVKRLVEEGRAELVPGVLFIDET 297

Query: 58  HMLDIEAFTFLNRALESTL 2
           H++DIE F F+NRA+ES +
Sbjct: 298 HLMDIELFAFMNRAMESEM 316




>gi|17737635|ref|NP_524156.1|  (NM_079432) reptin [Drosophila melanogaster]
 gi|7243682|gb|AAF43412.1|AF233279_1  (AF233279) reptin [Drosophila melanogaster]
 gi|7293815|gb|AAF49182.1|  (AE003517) reptin gene product [Drosophila melanogaster]
 gi|16768562|gb|AAL28500.1|  (AY060952) GM08688p [Drosophila melanogaster]
 gi|16768968|gb|AAL28703.1|  (AY061155) LD12420p [Drosophila melanogaster]
          Length = 481

 Score =  113 bits (282), Expect = 9e-25
 Identities = 62/148 (41%), Positives = 94/148 (62%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E   +EK+  GDVI I+  SG V ++GR      +YD    +  +V  P+G++ K+KE+V
Sbjct: 176 ECFMKEKIQAGDVITIDKASGKVNKLGRSFTRARDYDATGAQTRFVQCPEGELQKRKEVV 235

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             VTLH++DV N++    H F++L +        +I  ++R +IN  V ++ ++G AE+
Sbjct: 236 HTVTLHEIDVINSRT---HGFLALFSG----DTGEIKQEVRDQINNKVLEWREEGKAEIN 288

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE F+FLNRALES +
Sbjct: 289 PGVLFIDEVHMLDIECFSFLNRALESDM 316




>gi|17542510|ref|NP_501067.1|  (NM_068666) T22D1.10.p [Caenorhabditis elegans]
 gi|7508215|pir||T32710  hypothetical protein T22D1.10 - Caenorhabditis elegans
 gi|9798160|gb|AAF98631.1|  (AF039052) Hypothetical protein T22D1.10 [Caenorhabditis
elegans]
          Length = 448

 Score =  110 bits (276), Expect = 5e-24
 Identities = 63/146 (43%), Positives = 93/146 (63%), Gaps = 4/146 (2%)
 Frame = -2

Query: 427 REKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAE----EYVPLPKGDVHKKKEIVQD 260
           +EKV  GDVI ++  SG V R+GR   +N  +D +A     + V  P G++ K++E V
Sbjct: 180 KEKVMPGDVIQVDKASGRVTRLGRS--FNRSHDYDAMGPKVKLVQCPDGEIQKRRETVHT 237
```

```
Query: 259 VTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPG 80
           V LHD+DV N++ QG   +V+L +    + K ++ D    +INK V ++ ++G A+ VPG
Sbjct: 238 VCLHDIDVINSRTQG---YVALFSGDTGEIKAEVRD----QINKKVLEWREEGKAKFVPG 290

Query: 79  VIFIDECHMLDIEAFTFLNRALESTL 2
           V+FIDE HMLDIE F+FLNRA+E  L
Sbjct: 291 VLFIDEAHMLDIECFSFLNRAIEGEL 316



>gi|18978219|ref|NP_579576.1|  (NC_003413) tbp-interacting protein tip49 [Pyrococcus
furiosus DSM 3638]
 gi|18894032|gb|AAL81971.1|  (AE010280) tbp-interacting protein tip49 [Pyrococcus furiosus
DSM 3638]
          Length = 441

 Score =  108 bits (271), Expect = 2e-23
 Identities = 62/136 (45%), Positives = 84/136 (61%)
 Frame = -2

Query: 409 GDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVAN 230
           GDVI I+A +G V R+G       E  L   + V +P G V K KE    VTLHDLDV N
Sbjct: 188 GDVIQIDAETGRVSRIGTT---KEEEGLFFRKKVEMPTGPVLKIKEFTYTVTLHDLDVVN 244

Query: 229 AKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVIFIDECHML 50
           A+  G       +  L    + +I D++R  +++ V +++++G A LVPGV+FIDECHML
Sbjct: 245 ARAGG-------IFSLLFGGRMEINDEIRERVDQTVKQWVEEGKATLVPGVLFIDECHML 297

Query: 49  DIEAFTFLNRALESTL 2
           DIEAF+FL RA+ES L
Sbjct: 298 DIEAFSFLARAMESEL 313



>gi|15920793|ref|NP_376462.1|  (NC_003106) 452aa long hypothetical TATA-binding protein-
interacting protein [Sulfolobus tokodaii]
 gi|15621577|dbj|BAB65571.1|  (AP000983) 452aa long hypothetical TATA-binding protein-
interacting protein [Sulfolobus tokodaii]
          Length = 452

 Score =  108 bits (271), Expect = 2e-23
 Identities = 63/151 (41%), Positives = 91/151 (59%), Gaps = 2/151 (1%)
 Frame = -2

Query: 448 STHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSE--YDLEAEEYVPLPKGDVHKKK 275
           S  E L +  V  GDVI+I+A +G V +VG+   +     YD+E    V +P G V K+K
Sbjct: 175 SIAEQLVQLNVKKGDVIWIDAQTGEVSKVGKAKGFEGAKTYDIETIRQVDIPTGPVKKEK 234

Query: 274 EIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVA 95
           E    VTLHDLD+  A        +S+    +   + +I  ++R  ++K+V   I++G A
Sbjct: 235 ETTITVTLHDLDLNVAARN-----ISITALFSFFTEREINSEIRESVDKLVKDMINRGEA 289

Query: 94  ELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           ELVPGV+FID+ HMLDIEAF+FL +ALE+ L
Sbjct: 290 ELVPGVLFIDDAHMLDIEAFSFLTKALEADL 320
```

```
>gi|15229623|ref|NP_190552.1|  (NM_114843) RuvB DNA helicase - like protein [Arabidopsis
thaliana]
 gi|11280158|pir||T46049  RuvB DNA helicase-like protein - Arabidopsis thaliana
 gi|6723428|emb|CAB66921.1|  (AL132965) RuvB DNA helicase-like protein [Arabidopsis
thaliana]
          Length = 473

 Score =  107 bits (268), Expect = 4e-23
 Identities = 57/148 (38%), Positives = 96/148 (64%), Gaps = 2/148 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEE--YVPLPKGDVHKKKEIV 266
           E L +EKV  GDVI ++   G + ++GR    + ++D+   +  +V  P+G++ K+KE++
Sbjct: 178 EPLDKEKVQSGDVIVLDRFCGKITKLGRSFTRSRDFDVMGSKTKFVQCPEGELEKRKEVL 237

Query: 265 QDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELV 86
             VTLH++DV N++ QG   +++L          +I  + R + +  V+++ ++G AE+V
Sbjct: 238 HSVTLHEIDVINSRTQG---YLALFTG----DTGEIRSETREQSDTKVAEWREEGKAEIV 290

Query: 85  PGVIFIDECHMLDIEAFTFLNRALESTL 2
           PGV+FIDE HMLDIE F+FLNRALE+ +
Sbjct: 291 PGVLFIDEVHMLDIECFSFLNRALENDM 318




>gi|14520579|ref|NP_126054.1|  (NC_000868) TBP-interacting protein TIP49 [Pyrococcus abyssi]
 gi|7449965|pir||F75150  tbp-interacting protein tip49 PAB2107 - Pyrococcus abyssi  (strain
Orsay)
 gi|5457795|emb|CAB49285.1|  (AJ248284) TBP-interacting protein TIP49 [Pyrococcus abyssi]
          Length = 441

 Score =  107 bits (268), Expect = 4e-23
 Identities = 62/136 (45%), Positives = 82/136 (59%)
 Frame = -2

Query: 409 GDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVAN 230
           GDVI I+A +G V R+G       E  L   + V LP G V K KE    VTLHDLDV N
Sbjct: 188 GDVIQIDAETGRVSRIGTT---KEEEGLFFRKKVELPSGPVLKIKEFTYTVTLHDLDVVN 244

Query: 229 AKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVIFIDECHML 50
           A+  G    +            +I D++R  +++ V ++I++G A LVPGV+FIDECHML
Sbjct: 245 ARAGGIFSLIFGGGM-------EINDEIRERVDQTVKQWIEEGKATLVPGVLFIDECHML 297

Query: 49  DIEAFTFLNRALESTL 2
           DIEAF+FL RA+E+ L
Sbjct: 298 DIEAFSFLARAMENEL 313




>gi|18312811|ref|NP_559478.1|  (NC_003364) TBP-interacting protein TIP49 [Pyrobaculum
aerophilum]
 gi|18160296|gb|AAL63660.1|  (AE009834) TBP-interacting protein TIP49 [Pyrobaculum
aerophilum]
          Length = 450

 Score =  107 bits (267), Expect = 5e-23
 Identities = 57/134 (42%), Positives = 92/134 (68%)
 Frame = -2
```

```
Query: 409 GDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVAN 230
           GD+I I+  +G V  +GR +    +YD+  +    LPKG V+K+KEIV+  TLHD+DV+
Sbjct: 187 GDIIMIDEETGAVSVLGRGE-GGEQYDVAIKRRAELPKGPVYKEKEIVRFFTLHDIDVSL 245

Query: 229 AKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVIFIDECHML 50
           A+ +G    +S M     ++  +I D++R + +++V K +++G AELVPGV+FID+ H+L
Sbjct: 246 ARQRG---LISAMIFGFAEEVKEIPDEVRRQSDEIVKKTVEEGKAELVPGVLFIDDAHLL 302

Query: 49  DIEAFTFLNRALES 8
           DIE+F+FL RA+E+
Sbjct: 303 DIESFSFLMRAMET 316




>gi|14591557|ref|NP_143639.1|  (NC_000961) hypothetical protein [Pyrococcus horikoshii]
 gi|7449963|pir||D71191  hypothetical protein PH1804 - Pyrococcus horikoshii
          Length = 441

 Score =  106 bits (265), Expect = 9e-23
 Identities = 61/136 (44%), Positives = 83/136 (60%)
 Frame = -2

Query: 409 GDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVAN 230
           GDVI I+A +G V ++G       E  L  ++ V LP G V K KE    VTLHDLDV N
Sbjct: 188 GDVIQIDAETGRVSKIGTT---KEEEGLFFKKKVELPTGPVLKIKEFTYTVTLHDLDVVN 244

Query: 229 AKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVIFIDECHML 50
           A+  G    +            +I D++R  +++ V ++I++G A LVPGV+FIDECHML
Sbjct: 245 ARAGGIFSLIFGGGM-------EINDEIRERVDQTVKQWIEEGKATLVPGVLFIDECHML 297

Query: 49  DIEAFTFLNRALESTL 2
           DIEAF+FL RA+E+ L
Sbjct: 298 DIEAFSFLARAMENEL 313




>gi|19074927|ref|NP_586433.1|  (NC_003237) similarity to HYPOTHETICAL PROTEIN YP59_MYCTU
[Encephalitozoon cuniculi]
 gi|19069652|emb|CAD26037.1|  (AL590450) similarity to HYPOTHETICAL PROTEIN YP59_MYCTU
[Encephalitozoon cuniculi]
          Length = 418

 Score =  103 bits (258), Expect = 6e-22
 Identities = 57/141 (40%), Positives = 90/141 (63%)
 Frame = -2


Query: 433 LTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVT 254
           L +EKV+ GDVI I    G V ++G   +  S+       +VP P+G++ +  E  Q+++
Sbjct: 158 LDKEKVSAGDVIRIVRERGRVYKIGTSMVKRSDVVGTDTRFVPCPEGELIRITEETQEIS 217

Query: 253 LHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPGVI 74
           LHD+DV N+K +G   +++L +     +  +I  + R E+NK V  +I++G AE+V GV+
Sbjct: 218 LHDIDVVNSKAEG---YLALFSG----ETGEIRAETRDEVNKKVWGWINEGKAEIVRGVL 270

Query: 73  FIDECHMLDIEAFTFLNRALE 11
           FIDE HMLDIE+F FLN+A+E
Sbjct: 271 FIDEVHMLDIESFAFLNKAVE 291
```

```
>gi|14600633|ref|NP_147150.1|  (NC_000854) TATA-binding protein-interracing protein 49
[Aeropyrum pernix]
 gi|7449964|pir||E72723  probable TATA-binding protein-interracing protein 49 APE0326 -
Aeropyrum pernix (strain K1)
 gi|5103965|dbj|BAA79281.1|  (AP000059) 473aa long hypothetical TATA-binding protein-
interracing protein 49 [Aeropyrum pernix]
          Length = 473

 Score =  102 bits (253), Expect = 2e-21
 Identities = 57/150 (38%), Positives = 91/150 (60%)
 Frame = -2

Query: 451 PSTHENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKE 272
           P   E L    V  GDVI I+  +G V+ VGR       +D++    V LP+G V K KE
Sbjct: 197 PEVAEQLVALGVRRGDVIVIDLETGVVRVVGRGKGRGQSFDIDVVREVELPEGPVRKVKE 256

Query: 271 IVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAE 92
            V+ +TLHD+D + A  +    F  L++    ++   +T + R + +++V K++ +G AE
Sbjct: 257 FVRTLTLHDIDASIAAQRVA--FTGLLSMFEAER--GVTSEDRKKTDELVKKWVGEGKAE 312

Query: 91  LVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           +V GVIFID+ H+LD+E+F+FL++A+ES L
Sbjct: 313 IVAGVIFIDDAHLLDMESFSFLSKAMESDL 342



>gi|15899195|ref|NP_343800.1|  (NC_002754) TATA binding protein (TBP)-interacting protein
(TIP49-like), putative [Sulfolobus solfataricus]
 gi|13815753|gb|AAK42590.1|  (AE006844) TATA binding protein (TBP)-interacting protein
(TIP49-like), putative [Sulfolobus solfataricus]
          Length = 476

 Score = 95.1 bits (235), Expect = 3e-19
 Identities = 58/155 (37%), Positives = 93/155 (59%), Gaps = 6/155 (3%)
 Frame = -2

Query: 448 STHENLTREKVTIG----DVIYIEANSGNVKRVGRCDIYNSE--YDLEAEEYVPLPKGDV 287
           S  E + ++ V++G    DVI I+A +G V   G+   +     YD+E  + + +P G V
Sbjct: 195 SVGEEIAQQLVSLGVKKGDVIMIDAQTGQVIVEGKAKGFEGAKTYDIETTKVLEMPTGPV 254

Query: 286 HKKKEIVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVVSKYID 107
            K+KEI   +TL+DLD+  A        +++    +   + +I + +R E++++V  +I+
Sbjct: 255 RKEKEITTTLTLNDLDLNLAARN-----LAVTAIFSFFTEREINEDVRKEVDRLVKDWIN 309


Query: 106 QGVAELVPGVIFIDECHMLDIEAFTFLNRALESTL 2
           QG AELV GV+FID+ H LD+EAF+FL RALES L
Sbjct: 310 QGRAELVVGVLFIDDAHTLDLEAFSFLTRALESEL 344



>gi|17977678|ref|NP_524220.1|  (NM_079496) abstrakt; abstract [Drosophila melanogaster]
 gi|12643521|sp|Q9V3C0|ABS_DROME  DEAD-box protein abstrakt
 gi|6635811|gb|AAF19985.1|AF212866_1  (AF212866) abstrakt protein [Drosophila
melanogaster]
```

```
 gi|7296891|gb|AAF52165.1| ☐ (AE003607) abs gene product [Drosophila melanogaster]
 gi|15291815|gb|AAK93176.1| ☐ (AY051752) LD28839p [Drosophila melanogaster]
          Length = 619

 Score = 35.0 bits (79), Expect = 0.32
 Identities = 35/130 (26%), Positives = 54/130 (40%), Gaps = 30/130 (23%)
 Frame = -2

Query: 376 NVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVANAKP-------- 221
           +VKR  R    + E DL+ E+YVP       KK+ +++   +  L    A+P
Sbjct: 3   HVKRYRRSSKSSEEGDLDNEDYVPYVPVKERKKQHMIKLGRIVQLVSETAQPKSSSENEN 62

Query: 220 ----QGGHDF--------VSLMNQ------LNKKKKTDITDKLRGEINKVVSKYIDQ--- 104
               QG HD         +SL++Q      + + KK    +K   E  K++     Q
Sbjct: 63  EDDSQGAHDVETWGRKYNISLLDQHTELKKIAEAKKLSAVEKQLREEEKIMESIAQQKAL 122

Query: 103 -GVAELVPGV 77
            GVAEL  G+
Sbjct: 123 MGVAELAKGI 132



>gi|453389|gb|AAA73419.1|  (U05814) cytochrome C oxidase subunit III [Blastocrithidia
culicis]
          Length = 288

 Score = 34.3 bits (77), Expect = 0.55
 Identities = 18/68 (26%), Positives = 34/68 (49%), Gaps = 3/68 (4%)
 Frame = +2

Query: 164 SFLFFIQLIHQ*NKIMTSLWFSICYIQIMQCNILNNFFLFMYIAFGQWNILFCLQII--- 334
           S LF I L+H  + ++  +   IC+ +I   + ++  F+F+Y+    W+ + C+
Sbjct: 215 SILFVIDLLHFSHVLLGCILLFICFGRIFNFSCMDTRFIFLYLVCFYWHFVDCVWFFLLR 274

Query: 335 FTIVDVAS 358
           F   DV S
Sbjct: 275 FVYFDVLS 282



>gi|6682808|dbj|BAA88886.1|  (AB018404) similar to vaccinia A23R [Yaba monkey tumor virus]
          Length = 382

 Score = 33.9 bits (76), Expect = 0.72
 Identities = 27/112 (24%), Positives = 46/112 (40%), Gaps = 4/112 (3%)
 Frame = -2

Query: 445 THENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIV 266
           T +N    +    D++  +    N+K++  CDI+ +++ +    Y     G+V  K
Sbjct: 36  TMKNKISSETMFEDIVKCDDVRRNIKKLVYCDIHITKHIINQSLYPVYNVGNVFNKYSHF 95

Query: 265 QDVTLHDLDVA----NAKPQGGHDFVSLMNQLNKKKKTDITDKLRGEINKVV 122
            D+     DV+        +     +S +   NKKKK D      GEI K V
Sbjct: 96  FDINSSQSDVSLRTVEIFERDKSSLISYVKTTNKKKKVDY-----GEIKKTV 142
```

```
>gi|18311394|ref|NP_563328.1|  (NC_003366) RNA polymerase beta' subunit [Clostridium
perfringens]
 gi|20139578|sp|Q93R87|RPOC_CLOPE  DNA-directed RNA polymerase beta' chain (Transcriptase
beta' chain) (RNA polymerase beta' subunit)
 gi|15186725|dbj|BAB62885.1|  (AB055810) RNA polymerase beta' subunit [Clostridium
perfringens]
 gi|18146078|dbj|BAB82118.1|  (AP003194) RNA polymerase beta' subunit [Clostridium
perfringens]
          Length = 1178

 Score = 32.7 bits (73), Expect = 1.6
 Identities = 35/140 (25%), Positives = 61/140 (43%), Gaps = 28/140 (20%)
 Frame = -2

Query: 355 CDIYNSEYD-----------LEAE---EYVPLPKGDVHKKKE------IVQDVTLHD--- 245
           C  YN+++D           +EA+    ++ L  G++ K  +        QD+ L
Sbjct: 444 CTAYNADFDGDQMAVHVPLSVEAQAEARFLMLAAGNIMKPSDGRPVCVPTQDMVLGSYYL 503

Query: 244 -LDVANAKPQGGH----DFVSLMNQLNKKKKTDITDKLRGEINKVVSKYIDQGVAELVPG 80
            +D   AK +G +    D V +   QL   K+ DI  K+  ++ K +   +  G+ +  PG
Sbjct: 504 TMDKDGAKGEGKYFASFDEVIMAYQL---KEVDIHAKINVKVTKEIDGELKSGIIKTTPG 560

Query: 79  VIFIDECHMLDIEAFTFLNR 20
            I  +EC   D+    F+NR
Sbjct: 561 FIIFNECIPQDL---GFVNR 577




>gi|11466188|ref|NP_066511.1|  (NC_002573) NADH dehydrogenase subunit 5 [Naegleria gruberi]
 gi|10444223|gb|AAG17789.1|AF288092_14  (AF288092) NADH dehydrogenase subunit 5 [Naegleria
gruberi]
          Length = 666

 Score = 32.3 bits (72), Expect = 2.1
 Identities = 23/98 (23%), Positives = 45/98 (45%), Gaps = 2/98 (2%)
 Frame = +2

Query: 137 FTSQLVSDISFLFFIQL--IHQ*NKIMTSLWFSICYIQIMQCNILNNFFLFMYIAFGQWN 310
           + S L+  +S + F  L   +   KI+  L+F++ Y+Q     +LN FF F + AF
Sbjct: 376 YISILIGSLSLIGFPFLSGFYSKEKIV-QLFFNLFYLQFDSYGLLNLFFFFYFFAF---- 430

Query: 311 ILFCLQIIFTIVDVASAYSFNITTIGFNIDNISNGYFF 424
           +      ++++  +   +        +N+ NI  G F+
Sbjct: 431 VSIICTTLYSVKLLTFVFFVKYNGFKYNLHNIQYGSFY 468




>gi|7021194|dbj|BAA91402.1|   (AK000868) unnamed protein product [Homo sapiens]
          Length = 489

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 24/93 (25%), Positives = 44/93 (46%), Gaps = 1/93 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQD 260
           + L+ +K    D+   ++ SGN +     DI+    D E EE+    + D+ ++K   Q
Sbjct: 239 KELSDKKNEEKDLFGSDSESGNEEENLIADIFGESGDEEEEEFTGFNQEDLEEEKGETQV 298
```

```
Query: 259 VTLHDLDVANAKPQGGH-DFVSLMNQLNKKKKT 164
               D D  +   +G H DF+S    + ++KK+
Sbjct: 299 KEAEDSDSDDNIKRGKHMDFLSDFGMMLQRKKS 331
```

```
>gi|267421|sp|P14629|XPG_XENLA  DNA-repair protein complementing XP-G cells homolog
(Xeroderma pigmentosum group G complementing protein homolog)
 gi|422627|pir||S35994  DNA repair protein XPGC - African clawed frog
 gi|312433|emb|CAA49597.1|  (X69977) XP-G related factor [Xenopus laevis]
          Length = 1196

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 22/76 (28%), Positives = 33/76 (42%), Gaps = 2/76 (2%)
 Frame = -2

Query: 340 SEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLDVANAKPQGGHDF--VSLMNQLNKKKK 167
           S  D+E+EE+  LP    H+    ++D T     +  A P+   DF    L   L K
Sbjct: 188 SSVDIESEEFKSLPPEVKHEILTDMKDFTKRRRTLFEAMPEDSSDFSQYQLKGLLKKNDL 247

Query: 166 TDITDKLRGEINKVVS 119
               D +R E+N+  S
Sbjct: 248 NKCIDNVRKELNQQYS 263
```

```
>gi|16552402|dbj|BAB71301.1|  (AK056881) unnamed protein product [Homo sapiens]
          Length = 494

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 24/93 (25%), Positives = 44/93 (46%), Gaps = 1/93 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQD 260
           + L+ +K    D+   ++ SGN +     DI+    D E EE+    + D+ ++K   Q
Sbjct: 121 KELSDKKNEEKDLFGSDSESGNEEENLIADIFGESGDEEEEEFTGFNQEDLEEEKGETQV 180

Query: 259 VTLHDLDVANAKPQGGH-DFVSLMNQLNKKKKT 164
               D D  +   +G H DF+S    + ++KK+
Sbjct: 181 KEAEDSDSDDNIKRGKHMDFLSDFEMMLQRKKS 213
```

```
>gi|14042323|dbj|BAB55198.1|  (AK027561) unnamed protein product [Homo sapiens]
          Length = 819

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 24/93 (25%), Positives = 44/93 (46%), Gaps = 1/93 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQD 260
           + L+ +K    D+   ++ SGN +     DI+    D E EE+    + D+ ++K   Q
Sbjct: 446 KELSDKKNEEKDLFGSDSESGNEEENLIADIFGESGDEEEEEFTGFNQEDLEEEKGETQV 505

Query: 259 VTLHDLDVANAKPQGGH-DFVSLMNQLNKKKKT 164
               D D  +   +G H DF+S    + ++KK+
Sbjct: 506 KEAEDSDSDDNIKRGKHMDFLSDFEMMLQRKKS 538
```

```
>gi|17864498|ref|NP_524849.1|   (NM_080110) Ecdysone-inducible gene E1 [Drosophila
melanogaster]
 gi|6746588|gb|AAF27637.1|AF217281_1   (AF217281) ecdysone-inducible gene E1 [Drosophila
melanogaster]
          Length = 1616

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 26/97 (26%), Positives = 42/97 (42%), Gaps = 9/97 (9%)
 Frame = -2

Query: 424 EKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEY---------VPLPKGDVHKKKE 272
           E + I     +E+ S NV  VG+ +   +E ++E E+          V LP  +     E
Sbjct: 594 EALKIEQPAQVESESQNVDEVGKPEEPQAEMEIEKEKEETAATEAAEVELPATEAKPDME 653

Query: 271 IVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTD 161
              D      D+   KP  G D ++  NQLN+ + T+
Sbjct: 654 SEIDQVEQSADI---KPTSGEDDLNEANQLNQDQVTE 687
```

```
>gi|14737288|ref|XP_041928.1|   (XM_041928) similar to unnamed protein product [Homo
sapiens]
 gi|18551452|ref|XP_087073.1|   (XM_087073) hypothetical protein FLJ10006 [Homo sapiens]
          Length = 819

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 24/93 (25%), Positives = 44/93 (46%), Gaps = 1/93 (1%)
 Frame = -2

Query: 439 ENLTREKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEYVPLPKGDVHKKKEIVQD 260
           + L+ +K    D+   ++ SGN +     DI+    D E EE+    + D+ ++K   Q
Sbjct: 446 KELSDKKNEEKDLFGSDSESGNEEENLIADIFGESGDEEEEEFTGFNQEDLEEEKGETQV 505

Query: 259 VTLHDLDVANAKPQGGH-DFVSLMNQLNKKKKT 164
               D D  +   +G H DF+S    + ++KK+
Sbjct: 506 KEAEDSDSDDNIKRGKHMDFLSDFEMMLQRKKS 538
```

```
>gi|7295103|gb|AAF50429.1|   (AE003555) ImpE1 gene product [Drosophila melanogaster]
          Length = 1468

 Score = 32.0 bits (71), Expect = 2.7
 Identities = 26/97 (26%), Positives = 42/97 (42%), Gaps = 9/97 (9%)
 Frame = -2

Query: 424 EKVTIGDVIYIEANSGNVKRVGRCDIYNSEYDLEAEEY---------VPLPKGDVHKKKE 272
           E + I     +E+ S NV  VG+ +   +E ++E E+          V LP  +     E
Sbjct: 446 EALKIEQPAQVESESQNVDEVGKPEEPQAEMEIEKEKEETAATEAAEVELPATEAKPDME 505

Query: 271 IVQDVTLHDLDVANAKPQGGHDFVSLMNQLNKKKKTD 161
              D      D+   KP  G D ++  NQLN+ + T+
Sbjct: 506 SEIDQVEQSADI---KPTSGEDDLNEANQLNQDQVTE 539
```

```
>gi|3152919|gb|AAC17172.1|  (AF065141) ABC transporter ATP-binding protein homolog
[Streptococcus mutans]
          Length = 293

 Score = 31.6 bits (70), Expect = 3.6
 Identities = 16/64 (25%), Positives = 34/64 (53%), Gaps = 2/64 (3%)
 Frame = -2

Query: 412 IGDVIYIEANSGNVKRVGRC--DIYNSEYDLEAEEYVPLPKGDVHKKKEIVQDVTLHDLD 239
           + D +Y  A    +   GR   +I  +++D ++E+Y+ L    +++  +I+QD   + +
Sbjct: 194 LSDELYQVATKFGIIENGRIIKEISKADFDTQSEDYIVLKTSHINEASQILQDQMNYRIK 253

Query: 238 VANA 227
           V NA
Sbjct: 254 VTNA 257



>gi|6321327|ref|NP_011404.1|  (NC_001139) Nop seven associated; Nsa1p [Saccharomyces
cerevisiae]
 gi|1723904|sp|P53136|YGL1_YEAST  Hypothetical 51.9 kDa protein in TAF60-MLC1 intergenic
region
 gi|2131591|pir||S64119  hypothetical protein YGL111w - yeast (Saccharomyces cerevisiae)
 gi|1310713|emb|CAA66241.1|  (X97644) unknown ORF [Saccharomyces cerevisiae]
 gi|1322659|emb|CAA96818.1|  (Z72633) ORF YGL111w [Saccharomyces cerevisiae]
          Length = 463

 Score = 31.6 bits (70), Expect = 3.6
 Identities = 14/28 (50%), Positives = 17/28 (60%)
 Frame = +1

Query: 223 VQHLLHPNHVMQHLKQFLSFYVHRLWVM 306
           VQ  L P HV  HL + L  YV R+WV+
Sbjct: 26  VQSALQPFHVAPHLAEGLKAYVDRMWVI 53




>gi|11023511|gb|AAG26468.1|AF285774_8  (AF285774) putative flippase [Bacteroides fragilis]
          Length = 497

 Score = 31.2 bits (69), Expect = 4.7
 Identities = 23/72 (31%), Positives = 36/72 (49%)
 Frame = +2

Query: 32  CKCLYIKHVTFIYKYNSWDQFSHSLINIFAHNLVDFTSQLVSDISFLFFIQLIHQ*NKIM 211
           C   Y+ +  FIY +   + FS  +IN F   L+ F+S       FLF +  I + NKI+
Sbjct: 319 CIGFYLLNEQFIYLWLGENNFSGKIINFFIIILLFFSSINKVCFQFLFALGDIKKSNKIL 378

Query: 212 TSLWFSICYIQI 247
              + S+ Y+ I
Sbjct: 379 --FYQSLLYLPI 388
```

```
>gi|6689106|emb|CAB65393.1|  (AJ235845) NADH dehydrogenase, putative [Rondeletia odorata]
          Length = 734

 Score = 31.2 bits (69), Expect = 4.7
 Identities = 17/52 (32%), Positives = 27/52 (51%)
 Frame = +2

Query: 263 LNNFFLFMYIAFGQWNILFCLQIIFTIVDVASAYSFNITTIGFNIDNISNGY 418
           +NN F++ Y+     N  FCL+   + I   + S  S  ITT+G  +   S+ Y
Sbjct: 63  INNNFIYQYVWSWIINNDFCLEFGYLIDPLTSIMSILITTVGIMVLIYSDNY 114




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 252,326,721
Number of Sequences: 919285
Number of extensions: 5045995
Number of successful extensions: 18178
Number of sequences better than 10.0: 132
Number of HSP's better than 10.0 without gapping: 17390
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 18108
length of database: 288,558,979
effective HSP length: 111
effective length of database: 186,518,344
effective search space used: 10258508920
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020955450-015277-3947
```

7.1.13 Query= **hy-13_20235003**

```
         (313 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 150 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|400138|sp|Q02595|KPK2_PLAFK  Probable serine/threonine-pr...    52   2e-06
gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. prot...    51   3e-06
gi|18034789|ref|NP_542151.1|  (NM_080584) phosphorylase kina...    48   2e-05
gi|5453882|ref|NP_006204.1|  (NM_006213) phosphorylase kinas...    48   2e-05
gi|13276200|emb|CAC34070.1|  (AJ409105) putative serine/thre...    47   3e-05
gi|125534|sp|P00518|KPBG_RABIT  Phosphorylase B kinase gamma...    47   3e-05
gi|4389105|pdb|2PHK|A  Chain A, The Crystal Structure Of A P...    47   3e-05
gi|6730463|pdb|1QL6|A  Chain A, The Catalytic Mechanism Of P...    47   3e-05
gi|1827741|pdb|1PHK|   Two Structures Of The Catalytic Domai...    47   3e-05
gi|17433140|sp|Q9DB30|KPBH_MOUSE  Phosphorylase B kinase gam...    46   9e-05
gi|17436077|ref|XP_060231.1|  (XM_060231) similar to axonema...    45   1e-04
gi|226008|prf||1405343A  phosphorylase kinase gamma [Rattus ...    45   1e-04
gi|91179|pir||A29872  phosphorylase kinase (EC 2.7.1.38) cat...    45   1e-04
gi|13928776|ref|NP_113761.1|  (NM_031573) Phosphorylase kina...    45   1e-04
gi|6755054|ref|NP_035209.1|  (NM_011079) phosphorylase kinas...    45   1e-04
gi|14017839|dbj|BAB47440.1|  (AB058714) KIAA1811 protein [Ho...    45   2e-04
gi|19401871|gb|AAL87697.1|AF479826_1  (AF479826) putative se...    45   2e-04
gi|15302255|ref|XP_035920.2|  (XM_035920) KIAA1811 protein [...    45   2e-04
gi|19401874|gb|AAL87698.1|AF479827_1  (AF479827) protein kin...    45   2e-04
gi|7299437|gb|AAF54626.1|  (AE003691) KP78b gene product [Dr...    44   5e-04
gi|2564680|gb|AAB81837.1|  (AF023484) putative KP78 protein ...    44   5e-04
gi|19527891|gb|AAL90060.1|  (AY089322) AT13327p [Drosophila ...    44   5e-04
gi|4505785|ref|NP_000285.1|  (NM_000294) phosphorylase kinas...    44   5e-04
gi|9966875|ref|NP_065130.1|  (NM_020397) CamKI-like protein ...    43   8e-04
```

```
gi|19683092|gb|AAL92711.1|AC115685_15  (AC115685) Putative S...     43   8e-04
gi|804819|gb|AAB97113.1|  (L42007) protein serine/threonine ...     42   0.001
gi|5734778|gb|AAD50043.1|AC007980_8  (AC007980) Hypothetical...     42   0.001
gi|18076255|emb|CAC79947.1|  (AJ278957) protein kinase [Nyct...     42   0.001
gi|17569429|ref|NP_510253.1|  (NM_077852) serine/threonine k...     42   0.001
gi|6325116|ref|NP_015184.1|  (NC_001148) Hypothetical ORF; Y...     42   0.001
gi|3172111|dbj|BAA28663.1|  (AB014885) HrPOPK-1 [Halocynthia...     42   0.001
gi|6760436|gb|AAF28351.1|AF219232_1  (AF219232) qin-induced ...     42   0.001
gi|7499143|pir||T20941  hypothetical protein F15A2.6 - Caeno...     42   0.001
gi|7294217|gb|AAF49569.1|  (AE003529) CG6114 gene product [D...     42   0.001
gi|15222909|ref|NP_175441.1|  (NM_103907) hypothetical prote...     42   0.001
gi|19075441|ref|NP_587941.1|  (NC_003421) cds1 checkpoint ki...     42   0.001
gi|7434372|pir||T37321  Ca2+/calmodulin-dependent protein ki...     42   0.002
gi|17539480|ref|NP_500139.1|  (NM_067738) ser/thr protein ki...     42   0.002
gi|6324807|ref|NP_014876.1|  (NC_001147) protein kinase; Kin...     42   0.002
gi|12484153|gb|AAG53993.1|AF333958_1  (AF333958) calmodulin-...     41   0.002
gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falci...     41   0.002
gi|2854042|gb|AAC02532.1|  (AF043629) protein kinase 4 [Toxo...     41   0.002
gi|14133215|dbj|BAA34501.2|  (AB018324) KIAA0781 protein [Ho...     41   0.003
gi|12053045|emb|CAB66698.1|  (AL136764) hypothetical protein...     41   0.003
gi|18420882|ref|NP_568466.1|  (NM_122420) serine/threonine p...     40   0.005
gi|15215664|gb|AAK91377.1|  (AY050360) AT5g25110/T11H3_120 [...     40   0.005
gi|6331315|dbj|BAA86592.1|  (AB033104) KIAA1278 protein [Hom...     40   0.007
gi|9664225|gb|AAF97028.1|AF200815_1  (AF200815) FUSED serine...     40   0.007
gi|18552621|ref|XP_050803.3|  (XM_050803) KIAA1278 protein [...     40   0.007
gi|20072860|gb|AAH26158.1|  (BC026158) serine/threonine kina...     40   0.007
gi|18958196|emb|CAD24070.1|  (AJ431365) SNF1-related protein...     39   0.009
gi|1076633|pir||A56009  serine/threonine-specific protein ki...     39   0.009
gi|4567091|gb|AAD23582.1|AF128443_1  (AF128443) SNF-1-like s...     39   0.009
gi|18416414|ref|NP_568241.1|  (NM_121131) serine/threonine p...     39   0.011
gi|18425163|ref|NP_569046.1|  (NM_126120) receptor protein k...     39   0.011
gi|10177607|dbj|BAB10954.1|  (AB020742) receptor protein kin...     39   0.011
gi|11346410|pir||T50802  serine/threonine protein kinase-lik...     39   0.011
gi|6716522|gb|AAF26675.1|AF155821_1  (AF155821) CPG16 [Mus m...     39   0.011
gi|17985955|ref|NP_445795.1|  (NM_053343) double cortin and ...     39   0.011
gi|9910164|ref|NP_064362.1|  (NM_019978) double cortin and c...     39   0.011
gi|4758128|ref|NP_004725.1|  (NM_004734) doublecortin and Ca...     39   0.015
gi|13811953|ref|NP_113082.1|  (NC_002751) SNF1-related prote...     39   0.015
gi|12857094|dbj|BAB30889.1|  (AK017713) data source:SPTR, so...     39   0.015
gi|7299322|gb|AAF54516.1|  (AE003686) CG11870 gene product [...     39   0.015
gi|6225242|sp|O15075|DCK1_HUMAN  Serine/threonine-protein ki...     39   0.015
gi|15242507|ref|NP_199393.1|  (NM_123949) serine/threonine p...     38   0.019
gi|4507281|ref|NP_003150.1|  (NM_003159) serine/threonine ki...     38   0.019
gi|6521217|dbj|BAA88064.1|  (AB018002) Death-associated prot...     38   0.019
gi|4490842|emb|CAB38870.1|  (Z92542) dJ245G19.2 (serine/thre...     38   0.019
gi|12844658|dbj|BAB26448.1|  (AK009701) DEATH ASSOCIATED PRO...     38   0.019
gi|7446373|pir||T14822  probable serine/threonine protein ki...     38   0.025
gi|14758806|ref|XP_028271.1|  (XM_028271) DKFZP586E0820 prot...     38   0.025
gi|19923468|ref|NP_057541.2|  (NM_016457) protein kinase D2 ...     38   0.025
gi|7434268|pir||T08777  probable protein kinase C (EC 2.7.1....     38   0.025
gi|20161899|dbj|BAB90811.1|  (AP004614) putative protein kin...     38   0.025
gi|575292|emb|CAA57898.1|  (X82548) SNF1-related protein kin...     38   0.025
gi|2130048|pir||S59941  serine/threonine-specific protein ki...     38   0.025
gi|3217028|emb|CAA07196.1|  (AJ006701) putative serine/threo...     38   0.025
gi|18406082|ref|NP_566843.1|  (NM_113839) Snf1-related prote...     38   0.025
gi|1742967|emb|CAA64382.1|  (X94755) ser/thr protein kinase ...     38   0.025
gi|4678722|emb|CAB41259.1|  (AL049688) hypothetical protein ...     38   0.025
gi|4107009|dbj|BAA36298.1|  (D82039) OSK1 [Oryza sativa]            38   0.025
gi|15233981|ref|NP_193605.1|  (NM_117986) putative protein k...     38   0.025
gi|16755792|gb|AAL28100.1|AF428261_1  (AF428261) calcium/cal...     38   0.025
gi|14196445|ref|NP_065172.1|  (NM_020439) calcium/calmodulin...     38   0.025
```

```
gi|4007153|emb|CAA19296.1|  (AL023754) dJ272L16.1 (Rat Ca2+/...   38   0.025
gi|19263755|gb|AAH25307.1|  (BC025307) Unknown (protein for ...   38   0.025
gi|4742001|gb|AAD28798.1|AF146688_7  (AF146688) serine-threo...   37   0.033
gi|15226241|ref|NP_180965.1|  (NM_128969) putative protein k...   37   0.033
gi|20340969|ref|XP_110303.1|  (XM_110303) similar to calcium...   37   0.033
gi|11968142|ref|NP_071991.1|  (NM_022546) Death-associated l...   37   0.033
gi|6681133|ref|NP_031854.1|  (NM_007828) death-associated ki...   37   0.033
gi|10086466|gb|AAG12526.1|AC015446_7  (AC015446) Putative Pr...   37   0.033
gi|6730645|gb|AAF27066.1|AC008262_15  (AC008262) F4N2.17 [Ar...   37   0.033
gi|3560543|gb|AAC35001.1|  (AF052941) DAP-kinase related pro...   37   0.033
gi|15218505|ref|NP_174673.1|  (NM_103134) hypothetical prote...   37   0.033
gi|14670383|ref|NP_055141.2|  (NM_014326) death-associated p...   37   0.033
gi|2077934|dbj|BAA19880.1|  (D86557) Protein Kinase [Rattus ...   37   0.033
gi|18381097|gb|AAH22165.1|AAH22165  (BC022165) death-associa...   37   0.033
gi|18409220|ref|NP_564955.1|  (NM_105589) putative serine th...   37   0.033
```

Alignments

```
>gi|400138|sp|Q02595|KPK2_PLAFK  Probable serine/threonine-protein kinase 2
 gi|283478|pir||S23466  protein kinase homolog - malaria parasite (Plasmodium falciparum)
 gi|9938|emb|CAA45179.1|  (X63648) putative protein kinase [Plasmodium falciparum]
          Length = 510

 Score = 51.6 bits (122), Expect = 2e-06
 Identities = 22/59 (37%), Positives = 37/59 (62%)
 Frame = -1

Query: 256 DFEQCKELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSN 80
           + ++ K  E  VY EI++LR + H+ +++L   YE   ++++V EY KG  LF  LN+N
Sbjct: 141 EVDKSKVKESNVYTEIEVLRKVMHKYIIKLISAYEQEGFVYLVLEYLKGGELFEYLNNN 199



>gi|16805062|ref|NP_473091.1|  (NC_000910) calcium-dept. protein kinase (C-term. EF hand)
[Plasmodium falciparum]
 gi|422320|pir||A45472  protein kinase (EC 2.7.1.37) - malaria parasite (Plasmodium
falciparum)
 gi|9878|emb|CAA47704.1|  (X67288) protein kinase [Plasmodium falciparum]
 gi|3845284|gb|AAC71952.1|  (AE001419) calcium-dept. protein kinase (C-term. EF hand)
[Plasmodium falciparum 3D7]
          Length = 524

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 20/64 (31%), Positives = 42/64 (65%)
 Frame = -1

Query: 238 ELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQ 59
           ++ + +YNEI +L+S+ H ++++L  V+E+ KY ++V E+Y+G  LF  + +  +  E
Sbjct: 108 KIHEEIYNEISLLKSLDHPNIIKLFDVFEDKKYFYLVTEFYEGGELFEQIINRHKFDECD 167

Query: 58  IASV 47
            A++
Sbjct: 168 AANI 171



>gi|18034789|ref|NP_542151.1|  (NM_080584) phosphorylase kinase, gamma 2 (testis) [Rattus
norvegicus]
```

```
 gi|400122|sp|P31325|KPBH_RAT  Phosphorylase B kinase gamma catalytic chain, testis/liver
isoform (PHK-gamma-T) (Phosphorylase kinase gamma subunit 2)
 gi|285382|pir||A42034  phosphorylase kinase (EC 2.7.1.38) catalytic chain, gamma T - rat
 gi|206164|gb|AAA41863.1|   (M73808) phosphorylase kinase catalytic subunit [Rattus
norvegicus]
          Length = 406

 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 26/82 (31%), Positives = 46/82 (55%), Gaps = 1/82 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           EQ +E+ DA   E+ ILR +  H  ++ L   YE+S ++F+V++  +   LF+ L   +
Sbjct: 65  EQLEEVRDATRREMHILRQVAGHPHIITLIDSYESSSFMFLVFDLMRKGELFDYLTEKVA 124

Query: 73  LHEVQIASVYNSHLKIDNILNL 8
           L E +   S+  S L+  N L++
Sbjct: 125 LSEKETRSIMRSLLEAVNFLHV 146



>gi|5453882|ref|NP_006204.1|   (NM_006213) phosphorylase kinase, gamma 1 (muscle) [Homo
sapiens]
 gi|2833281|sp|Q16816|KPBG_HUMAN  Phosphorylase B kinase gamma catalytic chain, skeletal
muscle isoform (Phosphorylase kinase gamma subunit 1)
 gi|1147567|emb|CAA56681.1|   (X80590) phosphorylase kinase [Homo sapiens]
 gi|17224425|gb|AAL36972.1|  (AF254253) phosphorylase kinase gamma [Homo sapiens]
          Length = 387

 Score = 47.8 bits (112), Expect = 2e-05
 Identities = 22/76 (28%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ ILR +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 62  EEVRELREATLKEVDILRKVSGHPNIIQLKDTYETNTFFFLVFDLMKRGELFDYLTEKVT 121

Query: 73  LHEVQIASVYNSHLKI 26
           L E +    +  + L++
Sbjct: 122 LSEKETRKIMRALLEV 137



>gi|13276200|emb|CAC34070.1|  (AJ409105) putative serine/threonine kinase 2 [Entamoeba
histolytica]
          Length = 238

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 19/53 (35%), Positives = 39/53 (72%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNS 83
           ++++  +Y EI  ++ IR+  +++L +VYENS +LFIV +Y +G  L++L+++
Sbjct: 53  EKIKSQIYREIIFMKIIRNSYIVQLHQVYENSTHLFIVLDYMEGGELYDLIHT 105
```

```
>gi|125534|sp|P00518|KPBG_RABIT  Phosphorylase B kinase gamma catalytic chain, skeletal
muscle isoform (Phosphorylase kinase gamma subunit 1)
 gi|66740|pir||KIRBFG  phosphorylase kinase (EC 2.7.1.38) catalytic chain, skeletal muscle -
rabbit
 gi|1660|emb|CAA68682.1|  (Y00684) phosphorylase kinase gamma-subunit (AA 1-387)
[Oryctolagus cuniculus]
          Length = 387

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 22/76 (28%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ ILR +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 62  EEVQELREATLKEVDILRKVSGHPNIIQLKDTYETNTFFFLVFDLMKKGELFDYLTEKVT 121

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 122 LSEKETRKIMRALLEV 137
```

```
>gi|4389105|pdb|2PHK|A  Chain A, The Crystal Structure Of A Phosphorylase Kinase Peptide
Substrate Complex: Kinase Substrate Recognition
          Length = 277

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 22/76 (28%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ ILR +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 48  EEVQELREATLKEVDILRKVSGHPNIIQLKDTYETNTFFFLVFDLMKKGELFDYLTEKVT 107

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 108 LSEKETRKIMRALLEV 123
```

```
>gi|6730463|pdb|1QL6|A  Chain A, The Catalytic Mechanism Of Phosphorylase Kinase Probed By
Mutational Studies
          Length = 298

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 22/76 (28%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ ILR +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 61  EEVQELREATLKEVDILRKVSGHPNIIQLKDTYETNTFFFLVFDLMKKGELFDYLTEKVT 120

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 121 LSEKETRKIMRALLEV 136
```

```
>gi|1827741|pdb|1PHK|  Two Structures Of The Catalytic Domain Of Phosphorylase, Kinase: An
Active Protein Kinase Complexed With Nucleotide, Substrate-Analogue And Product
          Length = 298

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 22/76 (28%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ ILR +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 61  EEVQELREATLKEVDILRKVSGHPNIIQLKDTYETNTFFFLVFDLMKKGELFDYLTEKVT 120

Query: 73  LHEVQIASVYNSHLKI 26
           L E +    +  + L++
Sbjct: 121 LSEKETRKIMRALLEV 136



>gi|17433140|sp|Q9DB30|KPBH_MOUSE  Phosphorylase B kinase gamma catalytic chain,
testis/liver isoform (PHK-gamma-T) (Phosphorylase kinase gamma subunit 2)
 gi|12837721|dbj|BAB23926.1|  (AK005277) data source:SPTR, source key:P31325,
evidence:ISS~homolog to PHOSPHORYLASE B KINASE GAMMA CATALYTIC CHAIN, TESTIS ISOFORM (EC
2.7.1.38) (PHK-GAMMA-T) (PHOSPHORYLASE KINASE GAMMA SUBUNIT)~putative [Mus musculus]
          Length = 406

 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 25/81 (30%), Positives = 45/81 (54%), Gaps = 1/81 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           EQ +E+ DA   E+ ILR +  H  ++ L   YE+S ++F+V++  +   LF+ L   +
Sbjct: 65  EQLEEVRDATRREMHILRQVAGHPHIITLIDSYESSSFMFLVFDLMRKGELFDYLTEKVA 124

Query: 73  LHEVQIASVYNSHLKIDNILN 11
           L E +   S+  S L+  + L+
Sbjct: 125 LSEKETRSIMRSLLEAVSFLH 145



>gi|17436077|ref|XP_060231.1|  (XM_060231) similar to axonemal dynein heavy chain 7 [Homo
sapiens]
          Length = 5777

 Score = 45.4 bits (106), Expect = 1e-04
 Identities = 20/57 (35%), Positives = 35/57 (61%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQL 71
           K+  + V N +Q+LRS+ H +++E    YE + +L+I+ EY  G  +  +L SN+ L
Sbjct: 330 KKRRERVVNCVQLLRSMHHSNVIEFHNWYETNNHLWIITEYCTGGDMSTILRSNINL 386



>gi|226008|prf||1405343A  phosphorylase kinase gamma [Rattus norvegicus]
          Length = 387

 Score = 45.4 bits (106), Expect = 1e-04
 Identities = 21/76 (27%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
```

```
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ IL+ +  H ++++LK  YE + + F+V++  K   LF+ L    +
Sbjct: 61  EEVQELREATLKEVDILQKVSGHPNIIQLKDTYETNTFFFLVFDLMKRGELFDYLTEKVT 120

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 121 LTEKETRKIMRALLEV 136




>gi|91179|pir||A29872  phosphorylase kinase (EC 2.7.1.38) catalytic chain, skeletal muscle -
mouse
 gi|200341|gb|AAA39926.1|  ☐ (M16216) phosphorylase kinase, gamma-subunit [Mus musculus]
          Length = 388

 Score = 45.4 bits (106), Expect = 1e-04
 Identities = 21/76 (27%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ IL+ +  H ++++LK  YE + + F+V++  K   LF+ L    +
Sbjct: 62  EEVQELREATLKEVDILQKVSGHPNIIQLKDTYETNTFFFLVFDLMKRGELFDYLTEKVT 121

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 122 LTEKETRKIMRALLEV 137




>gi|13928776|ref|NP_113761.1|  ☐ (NM_031573) Phosphorylase kinase, gamma 1; phosphorylase
kinase gamma [Rattus norvegicus]
 gi|125535|sp|P13286|KPBG_RAT  Phosphorylase B kinase gamma catalytic chain, skeletal muscle
isoform (Phosphorylase kinase gamma subunit 1)
 gi|92615|pir||S00731  phosphorylase kinase (EC 2.7.1.38) catalytic chain [similarity] - rat
 gi|56927|emb|CAA30280.1|  ☐ (X07320) phosphorylase kinase gamma-subunit (AA 1 - 388)
[Rattus norvegicus]
          Length = 388

 Score = 45.4 bits (106), Expect = 1e-04
 Identities = 21/76 (27%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ IL+ +  H ++++LK  YE + + F+V++  K   LF+ L    +
Sbjct: 62  EEVQELREATLKEVDILQKVSGHPNIIQLKDTYETNTFFFLVFDLMKRGELFDYLTEKVT 121

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 122 LTEKETRKIMRALLEV 137




>gi|6755054|ref|NP_035209.1|  ☐ (NM_011079) phosphorylase kinase gamma [Mus musculus]
 gi|2507189|sp|P07934|KPBG_MOUSE  Phosphorylase B kinase gamma catalytic chain, skeletal
muscle isoform (Phosphorylase kinase gamma subunit 1)
```

```
 gi|2117794|pir||A46124  phosphorylase kinase (EC 2.7.1.38) gamma chain - mouse
 gi|200339|gb|AAA39925.1|   (J03293) phosphorylase kinase, gamma-subunit [Mus musculus]
 gi|499651|gb|AAB59721.1|   (L08059) gamma phosphorylase kinase [Mus musculus]
          Length = 388

 Score = 45.4 bits (106), Expect = 1e-04
 Identities = 21/76 (27%), Positives = 43/76 (55%), Gaps = 1/76 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           E+ +EL +A   E+ IL+ +  H ++++LK  YE + + F+V++  K   LF+ L   +
Sbjct: 62  EEVQELREATLKEVDILQKVSGHPNIIQLKDTYETNTFFFLVFDLMKRGELFDYLTEKVT 121

Query: 73  LHEVQIASVYNSHLKI 26
           L E +   +  + L++
Sbjct: 122 LTEKETRKIMRALLEV 137



>gi|14017839|dbj|BAB47440.1|   (AB058714) KIAA1811 protein [Homo sapiens]
          Length = 715

 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 22/45 (48%), Positives = 29/45 (63%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI IL+ I H  +L+L  VYEN KYL++V E+  G  LF+ L
Sbjct: 15  VEREIAILKLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 59



>gi|19401871|gb|AAL87697.1|AF479826_1   (AF479826) putative serine/threonine protein
kinase [Homo sapiens]
          Length = 794

 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 22/45 (48%), Positives = 29/45 (63%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI IL+ I H  +L+L  VYEN KYL++V E+  G  LF+ L
Sbjct: 94  VEREIAILKLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 138



>gi|15302255|ref|XP_035920.2|   (XM_035920) KIAA1811 protein [Homo sapiens]
          Length = 703

 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 22/45 (48%), Positives = 29/45 (63%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI IL+ I H  +L+L  VYEN KYL++V E+  G  LF+ L
Sbjct: 3   VEREIAILKLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 47
```

```
>gi|19401874|gb|AAL87698.1|AF479827_1  ☐ (AF479827) protein kinase-like protein [Homo
sapiens]
          Length = 778

 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 22/45 (48%), Positives = 29/45 (63%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI IL+ I H  +L+L  VYEN KYL++V E+  G  LF+ L
Sbjct: 78  VEREIAILKLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 122



>gi|7299437|gb|AAF54626.1|  ☐ (AE003691) KP78b gene product [Drosophila melanogaster]
          Length = 604

 Score = 43.5 bits (101), Expect = 5e-04
 Identities = 18/53 (33%), Positives = 33/53 (61%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           +Y E+ I++ + H +++ L +V E+ + L++V EY  G  LFN L  N ++ E
Sbjct: 107 LYREVNIMKKLNHPNIVRLLQVIESERTLYLVMEYVSGGELFNYLVKNGRMRE 159



>gi|2564680|gb|AAB81837.1|  ☐ (AF023484) putative KP78 protein kinase [Drosophila
melanogaster]
          Length = 604

 Score = 43.5 bits (101), Expect = 5e-04
 Identities = 18/53 (33%), Positives = 33/53 (61%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           +Y E+ I++ + H +++ L +V E+ + L++V EY  G  LFN L  N ++ E
Sbjct: 107 LYREVNIMKKLNHPNIVRLLQVIESERTLYLVMEYVSGGELFNYLVKNGRMRE 159



>gi|19527891|gb|AAL90060.1|  ☐ (AY089322) AT13327p [Drosophila melanogaster]
          Length = 604

 Score = 43.5 bits (101), Expect = 5e-04
 Identities = 18/53 (33%), Positives = 33/53 (61%)
 Frame = -1


Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           +Y E+ I++ + H +++ L +V E+ + L++V EY  G  LFN L  N ++ E
Sbjct: 107 LYREVNIMKKLNHPNIVRLLQVIESERTLYLVMEYVSGGELFNYLVKNGRMRE 159
```

```
>gi|4505785|ref|NP_000285.1| ☐ (NM_000294) phosphorylase kinase, gamma 2 (testis);
Phosphorylase kinase, gamma 2 (testis/liver) [Homo sapiens]
 gi|18586175|ref|XP_083946.1| ☐ (XM_083946) phosphorylase kinase, gamma 2 (testis) [Homo
sapiens]
 gi|125536|sp|P15735|KPBH_HUMAN  Phosphorylase B kinase gamma catalytic chain, testis/liver
isoform (PHK-gamma-T) (Phosphorylase kinase gamma subunit 2) (PSK-C3)
 gi|66739|pir||KIHUCT  phosphorylase kinase (EC 2.7.1.38) catalytic chain, testis - human
 gi|189941|gb|AAA36442.1| ☐ (M31606) phosphorylase kinase [Homo sapiens]
 gi|2832753|emb|CAA72694.1| ☐ (Y11950) phosphorylase kinase gamma subunit [Homo sapiens]
 gi|12803435|gb|AAH02541.1|AAH02541 ☐ (BC002541) phosphorylase kinase, gamma 2 (testis)
[Homo sapiens]
          Length = 406

 Score = 43.5 bits (101), Expect = 5e-04
 Identities = 24/81 (29%), Positives = 44/81 (53%), Gaps = 1/81 (1%)
 Frame = -1

Query: 250 EQCKELEDAVYNEIQILRSIR-HQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQ 74
           EQ +E+ +A   E  ILR +  H  ++ L   YE+S ++F+V++  +   LF+ L   +
Sbjct: 65  EQLEEVREATRRETHILRQVAGHPHIITLIDSYESSSFMFLVFDLMRKGELFDYLTEKVA 124

Query: 73  LHEVQIASVYNSHLKIDNILN 11
           L E +  S+  S L+  + L+
Sbjct: 125 LSEKETRSIMRSLLEAVSFLH 145



>gi|9966875|ref|NP_065130.1| ☐ (NM_020397) CamKI-like protein kinase [Homo sapiens]
 gi|9837341|gb|AAG00534.1|AF286366_1 ☐ (AF286366) CamKI-like protein kinase [Homo sapiens]
          Length = 357

 Score = 42.7 bits (99), Expect = 8e-04
 Identities = 17/49 (34%), Positives = 33/49 (66%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFN 95
           K  E ++ NEI +LR I+H++++ L+ +YE+  +L++V +   G  LF+
Sbjct: 60  KGKESSIENEIAVLRKIKHENIVALEDIYESPNHLYLVMQLVSGGELFD 108



>gi|19683092|gb|AAL92711.1|AC115685_15  (AC115685) Putative Serine/threonine protein kinase
[Dictyostelium discoideum]
          Length = 833

 Score = 42.7 bits (99), Expect = 8e-04
 Identities = 18/45 (40%), Positives = 30/45 (66%)
 Frame = -1


Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           +  EI +++ I H + +++  VYE SKYLF++ EY +G   LF+ L
Sbjct: 55  IEREIVLMKLIDHPNAMKMYEVYETSKYLFLILEYVEGGELFDYL 99
```

```
>gi|804819|gb|AAB97113.1|  (L42007) protein serine/threonine kinase [Toxoplasma gondii]
          Length = 509

 Score = 42.4 bits (98), Expect = 0.001
 Identities = 20/67 (29%), Positives = 36/67 (52%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIAS 50
           D V+ EIQ LR + H+ ++++K V +N KY+  + EY     L N ++   +L E +
Sbjct: 80  DRVFVEIQALRDLSHKHVIKMKDVVDNPKYICFIMEYATNGELRNYVSKKTRLKEDEARQ 139

Query: 49  VYNSHLK 29
            +   +K
Sbjct: 140 FFEQIIK 146



>gi|5734778|gb|AAD50043.1|AC007980_8  (AC007980) Hypothetical protein [Arabidopsis thaliana]
          Length = 1097

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/62 (32%), Positives = 40/62 (64%)
 Frame = -1

Query: 214 EIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIASVYNSH 35
           EI+ILR ++H++++E+   +EN++   +V E+ +GE LF +L  +  L E Q+ ++
Sbjct: 53  EIEILRKLKHENIIEMLDSFENAREFCVVTEFAQGE-LFEILEDDKCLPEEQVQAIAKQL 111

Query: 34  LK 29
           +K
Sbjct: 112 VK 113



>gi|18076255|emb|CAC79947.1|  (AJ278957) protein kinase [Nyctotherus ovalis]
          Length = 465

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/70 (28%), Positives = 39/70 (55%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIAS 50
           + + +E+ IL ++ H ++++  + YE+ +YL+IV EY  G  LF  +    + +E + A
Sbjct: 77  EKLMSEVDILSAVDHPNIVKYMKHYESKRYLYIVMEYCPGGDLFQKVIKQNKFNEAEAAI 136

Query: 49  VYNSHLKIDN 20
           V    L+  N
Sbjct: 137 VMEEVLRAIN 146



>gi|17569429|ref|NP_510253.1|  (NM_077852) serine/threonine kinase (KIN1/SNF1/Nim1
subfamily) [Caenorhabditis elegans]
 gi|12276122|gb|AAG50270.1|AF316542_1  (AF316542) serine/threonine kinase SAD-1
[Caenorhabditis elegans]
 gi|14530414|emb|CAA94127.2|  (Z70207) predicted using Genefinder~Similarity to Arabidopsis
serine/threonine protein kinase (PIR Acc. No. JC1446), contains similarity to Pfam domain:
```

```
PF00069 (Eukaryotic protein kinase domain), Score=316.9, E-value=7.8e-92, N=1~cDNA EST
yk41f2.3 com>
          Length = 914

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/45 (44%), Positives = 28/45 (61%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI I++ I H  +L L  VYEN KYL+++ E+  G  LF+ L
Sbjct: 91  VEREIAIMKLIEHPHVLHLYDVYENKKYLYLLLEHVSGGELFDYL 135




>gi|6325116|ref|NP_015184.1|  (NC_001148) Hypothetical ORF; Ypl141cp [Saccharomyces
cerevisiae]
 gi|2132203|pir||S69044  hypothetical protein YPL141c - yeast (Saccharomyces cerevisiae)
 gi|1244774|gb|AAB68219.1|  (U43703) Lpi5p [Saccharomyces cerevisiae]
          Length = 865

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 18/64 (28%), Positives = 36/64 (56%)
 Frame = -1

Query: 232 EDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIA 53
           E  +Y EI  L+ + H ++++L+ V +NS+Y+ IV EY  G   +  +    +L E+
Sbjct: 94  EVKIYREINALKHLSHPNIVKLEEVLQNSRYIGIVLEYACGGEFYKYIQKKRRLKEMNAC 153

Query: 52  SVYN 41
            +++
Sbjct: 154 RLFS 157




>gi|3172111|dbj|BAA28663.1|  (AB014885) HrPOPK-1 [Halocynthia roretzi]
          Length = 698

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 22/55 (40%), Positives = 33/55 (60%), Gaps = 4/55 (7%)
 Frame = -1

Query: 241 KELEDAVYN----EIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           ++L ++V N    EI I++ I H  +L L  VYEN KYL+++ E   G  LF+ L
Sbjct: 48  EKLSESVINKVEREIAIMKLIEHPHILGLHDVYENKKYLYLILELVSGGELFDYL 102




>gi|6760436|gb|AAF28351.1|AF219232_1  (AF219232) qin-induced kinase [Gallus gallus]
          Length = 798

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 19/55 (34%), Positives = 33/55 (59%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           + +Y E+QI++ + H  +++L +V E    L+IV E+ K   +F+ L SN  L E
Sbjct: 68  EKIYREVQIMKLLNHPHIIKLYQVMETKDMLYIVTEFAKNGEMFDHLTSNGHLSE 122
```

```
>gi|7499143|pir||T20941  hypothetical protein F15A2.6 - Caenorhabditis elegans
          Length = 887

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/45 (44%), Positives = 28/45 (61%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI I++ I H  +L L  VYEN KYL+++ E+  G  LF+ L
Sbjct: 64  VEREIAIMKLIEHPHVLHLYDVYENKKYLYLLLEHVSGGELFDYL 108



>gi|7294217|gb|AAF49569.1|   (AE003529) CG6114 gene product [Drosophila melanogaster]
          Length = 851

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/45 (44%), Positives = 28/45 (61%)
 Frame = -1

Query: 223 VYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLL 89
           V  EI I++ I H  +L L  VYEN KYL+++ E+  G  LF+ L
Sbjct: 62  VEREIAIMKLIDHPHVLGLSDVYENKKYLYLILEHVSGGELFDYL 106



>gi|15222909|ref|NP_175441.1|  (NM_103907) hypothetical protein [Arabidopsis thaliana]
          Length = 269

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/62 (32%), Positives = 40/62 (64%)
 Frame = -1

Query: 214 EIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIASVYNSH 35
           EI+ILR ++H++++E+   +EN++   +V E+ +GE LF +L  +  L E Q+ ++
Sbjct: 53  EIEILRKLKHENIIEMLDSFENAREFCVVTEFAQGE-LFEILEDDKCLPEEQVQAIAKQL 111

Query: 34  LK 29
           +K
Sbjct: 112 VK 113



>gi|19075441|ref|NP_587941.1|  (NC_003421) cds1 checkpoint kinase. [Schizosaccharomyces
pombe]
 gi|12644396|sp|Q09170|CDS1_SCHPO  Protein kinase cds1 (Checkpoint kinase cds1)
 gi|7490200|pir||T41204  cds1 kinase - fission yeast (Schizosaccharomyces pombe)
 gi|2689197|emb|CAA11019.1|  (AJ222869) Cds1 kinase [Schizosaccharomyces pombe]
 gi|5701966|emb|CAB52158.1|  (AL109736) cds1 checkpoint kinase. [Schizosaccharomyces pombe]
          Length = 460

 Score = 42.0 bits (97), Expect = 0.001
 Identities = 20/59 (33%), Positives = 32/59 (53%)
 Frame = -1
```

```
Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           K   +    EI IL+S+ H  +++   ++EN   LFIV EY +G  L + L +N  + E
Sbjct: 210 KRATEMFQREIDILKSLHHPGVVQCHEIFENDDELFIVMEYVEGGDLMDFLIANGSIDE 268



>gi|7434372|pir||T37321  Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I -
Caenorhabditis elegans
 gi|5672678|dbj|BAA82674.1|  (AB021864) Ca2+/calmodulin-dependent protein kinase I
[Caenorhabditis elegans]
          Length = 348

 Score = 41.6 bits (96), Expect = 0.002
 Identities = 17/54 (31%), Positives = 35/54 (64%)
 Frame = -1

Query: 256 DFEQCKELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFN 95
           D +  K  E+++ NEI++LR +RH ++++L   Y+  +++++V E   G  LF+
Sbjct: 55  DKKALKGKEESLENEIKVLRKLRHNNIVQLFDTYDEKQFVYLVMELVTGGELFD 108



>gi|17539480|ref|NP_500139.1|  (NM_067738) ser/thr protein kinase [Caenorhabditis elegans]
 gi|6671812|gb|AAF23187.1|AF099924_2  (AF099924) Hypothetical protein K07A9.2
[Caenorhabditis elegans]
          Length = 348

 Score = 41.6 bits (96), Expect = 0.002
 Identities = 17/54 (31%), Positives = 35/54 (64%)
 Frame = -1

Query: 256 DFEQCKELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFN 95
           D +  K  E+++ NEI++LR +RH ++++L   Y+  +++++V E   G  LF+
Sbjct: 55  DKKALKGKEESLENEIKVLRKLRHNNIVQLFDTYDEKQFVYLVMELVTGGELFD 108



>gi|6324807|ref|NP_014876.1|  (NC_001147) protein kinase; Kin4p [Saccharomyces cerevisiae]
 gi|729900|sp|Q01919|KIN4_YEAST  Serine/threonine-protein kinase KIN4
 gi|322953|pir||S29344  protein kinase KIN3 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
 gi|5515|emb|CAA48115.1|  (X67916) protein kinase [Saccharomyces cerevisiae]
 gi|1420535|emb|CAA99453.1|  (Z75141) ORF YOR233w [Saccharomyces cerevisiae]
          Length = 800

 Score = 41.6 bits (96), Expect = 0.002
 Identities = 19/65 (29%), Positives = 35/65 (53%)
 Frame = -1

Query: 259 DDFEQCKELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSN 80
           D  ++  + E  +Y EI  L+ + H +++ L+ V +NSKY+ IV E+  G   +  +
Sbjct: 85  DTIKKDADKEIKIYREINALKHLTHPNIIYLEEVLQNSKYIGIVLEFVSGGEFYKYIQRK 144

Query: 79  LQLHE 65
            +L E
Sbjct: 145 RRLKE 149
```

```
>gi|12484153|gb|AAG53993.1|AF333958_1  (AF333958) calmodulin-domain protein kinase 1
[Toxoplasma gondii]
          Length = 507

 Score = 41.2 bits (95), Expect = 0.002
 Identities = 17/62 (27%), Positives = 37/62 (59%)
 Frame = -1

Query: 232 EDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIA 53
           ++++  E+Q+L+ + H ++++L   +E+  Y ++V E Y G  LF+ + S  +  EV  A
Sbjct: 93  KESLLREVQLLKQLDHPNIMKLYEFFEDKGYFYLVGEVYTGGELFDEIISRKRFSEVDAA 152

Query: 52  SV 47
            +
Sbjct: 153 RI 154




>gi|2315243|emb|CAA68090.1|  (X99763) CDPK2 [Plasmodium falciparum]
          Length = 513

 Score = 41.2 bits (95), Expect = 0.002
 Identities = 16/58 (27%), Positives = 33/58 (56%)
 Frame = -1


Query: 220 YNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIASV 47
           + EI+I++ + H ++++L   YEN  Y++++ E   G  LF+ +  N    E   A++
Sbjct: 115 FQEIEIMKKLDHPNIVKLYETYENDNYIYLIMELCSGRELFDSIIENGSFTEKNAATI 172




>gi|2854042|gb|AAC02532.1|  (AF043629) protein kinase 4 [Toxoplasma gondii]
          Length = 501

 Score = 41.2 bits (95), Expect = 0.002
 Identities = 17/62 (27%), Positives = 37/62 (59%)
 Frame = -1

Query: 232 EDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEVQIA 53
           ++++  E+Q+L+ + H ++++L   +E+  Y ++V E Y G  LF+ + S  +  EV  A
Sbjct: 93  KESLLREVQLLKQLDHPNIMKLYEFFEDKGYFYLVGEVYTGGELFDEIISRKRFSEVDAA 152

Query: 52  SV 47
            +
Sbjct: 153 RI 154



>gi|14133215|dbj|BAA34501.2|   (AB018324) KIAA0781 protein [Homo sapiens]
          Length = 896

 Score = 40.8 bits (94), Expect = 0.003
 Identities = 17/55 (30%), Positives = 35/55 (62%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           + +Y E+QI++ + H  +++L +V E    L++V EY K   +F+ L ++ +L+E
Sbjct: 32  EKIYREVQIMKMLDHPHIIKLYQVMETKSMLYLVTEYAKNGEIFDYLANHGRLNE 86
```

```
>gi|12053045|emb|CAB66698.1|  (AL136764) hypothetical protein [Homo sapiens]
          Length = 926

 Score = 40.8 bits (94), Expect = 0.003
 Identities = 17/55 (30%), Positives = 35/55 (62%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHE 65
           + +Y E+QI++ + H  +++L +V E    L++V EY K   +F+ L ++ +L+E
Sbjct: 62  EKIYREVQIMKMLDHPHIIKLYQVMETKSMLYLVTEYAKNGEIFDYLANHGRLNE 116


>gi|18420882|ref|NP_568466.1|  (NM_122420) serine/threonine protein kinase-like protein
           [Arabidopsis thaliana]
 gi|17646697|gb|AAL41008.1|AF448226_1  (AF448226) CBL-interacting protein kinase CIPK25
[Arabidopsis thaliana]
          Length = 488

 Score = 40.0 bits (92), Expect = 0.005
 Identities = 17/45 (37%), Positives = 27/45 (59%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFN 95
           + +  EI I+R +RH +++ELK V      +F + EY KG  LF+
Sbjct: 86  EQIKREISIMRLVRHPNIVELKEVMATKTKIFFIMEYVKGGELFS 130


>gi|15215664|gb|AAK91377.1|  (AY050360) AT5g25110/T11H3_120 [Arabidopsis thaliana]
          Length = 487

 Score = 40.0 bits (92), Expect = 0.005
 Identities = 17/45 (37%), Positives = 27/45 (59%)
 Frame = -1

Query: 229 DAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFN 95
           + +  EI I+R +RH +++ELK V      +F + EY KG  LF+
Sbjct: 86  EQIKREISIMRLVRHPNIVELKEVMATKTKIFFIMEYVKGGELFS 130


>gi|6331315|dbj|BAA86592.1|  (AB033104) KIAA1278 protein [Homo sapiens]
          Length = 1311

 Score = 39.7 bits (91), Expect = 0.007
 Identities = 21/65 (32%), Positives = 41/65 (62%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEV 62
           KEL + +  EI+I+R +RH +++ +   +E  K + +V +Y +GE LF +L  + +L E
Sbjct: 60  KELRN-LQREIEIMRGLRHPNIVHMLDSFETDKEVVVVTDYAEGE-LFQILEDDGKLPED 117
```

```
Query: 61  QIASV 47
           Q+ ++
Sbjct: 118 QVQAI 122



>gi|9664225|gb|AAF97028.1|AF200815_1  (AF200815) FUSED serine/threonine kinase [Homo
sapiens]
          Length = 1315

 Score = 39.7 bits (91), Expect = 0.007
 Identities = 21/65 (32%), Positives = 41/65 (62%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEV 62
           KEL + +  EI+I+R +RH +++ +   +E  K + +V +Y +GE LF +L  + +L E
Sbjct: 43  KELRN-LQREIEIMRGLRHPNIVHMLDSFETDKEVVVVTDYAEGE-LFQILEDDGKLPED 100

Query: 61  QIASV 47
           Q+ ++
Sbjct: 101 QVQAI 105



>gi|18552621|ref|XP_050803.3|  (XM_050803) KIAA1278 protein [Homo sapiens]
          Length = 1315

 Score = 39.7 bits (91), Expect = 0.007
 Identities = 21/65 (32%), Positives = 41/65 (62%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEV 62
           KEL + +  EI+I+R +RH +++ +   +E  K + +V +Y +GE LF +L  + +L E
Sbjct: 43  KELRN-LQREIEIMRGLRHPNIVHMLDSFETDKEVVVVTDYAEGE-LFQILEDDGKLPED 100

Query: 61  QIASV 47
           Q+ ++
Sbjct: 101 QVQAI 105



>gi|20072860|gb|AAH26158.1|  (BC026158) serine/threonine kinase 36 (fused homolog,
Drosophila) [Homo sapiens]
          Length = 1315

 Score = 39.7 bits (91), Expect = 0.007
 Identities = 21/65 (32%), Positives = 41/65 (62%)
 Frame = -1

Query: 241 KELEDAVYNEIQILRSIRHQSLLELKRVYENSKYLFIVYEYYKGETLFNLLNSNLQLHEV 62
           KEL + +  EI+I+R +RH +++ +   +E  K + +V +Y +GE LF +L  + +L E
Sbjct: 43  KELRN-LQREIEIMRGLRHPNIVHMLDSFETDKEVVVVTDYAEGE-LFQILEDDGKLPED 100

Query: 61  QIASV 47
           Q+ ++
Sbjct: 101 QVQAI 105
```

```
  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 147,064,199
Number of Sequences: 919285
Number of extensions: 2645529
Number of successful extensions: 14513
Number of sequences better than 10.0: 689
Number of HSP's better than 10.0 without gapping: 14297
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 14511
length of database: 288,558,979
effective HSP length: 79
effective length of database: 215,935,464
effective search space used: 5182451136
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020955661-019454-31363
```

7.1.14 Query= **hy-14_20240016**

```
          (374 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 50 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                        (bits) Value

gi|20380139|gb|AAH28035.1|  (BC028035) Unknown (protein for ...      78   2e-14
gi|20343349|ref|XP_110091.1|  (XM_110091) similar to bM150J2...      78   2e-14
gi|7427509|ref|NP_001576.2|  (NM_001585) chromosome 22 open ...      78   2e-14
gi|4502485|ref|NP_001575.1|  (NM_001584) chromosome 11 open ...      76   6e-14
gi|12849370|dbj|BAB28313.1|  (AK012553) data source:SPTR, so...      76   6e-14
gi|7294897|gb|AAF50227.1|  (AE003551) CG16717 gene product [...      67   3e-11
gi|17558234|ref|NP_505349.1|  (NM_072948) C25E10.12.p [Caeno...      62   1e-09
gi|15131236|emb|CAC48257.1|  (AL356240) dJ873F21.1 (brain pr...      58   2e-08
gi|3025256|sp|Q21268|YMSO_CAEEL  Hypothetical 30.9 kDa prote...      56   8e-08
gi|17562344|ref|NP_505116.1|  (NM_072715) K07C11.7.p [Caenor...      56   8e-08
gi|6562617|emb|CAB62592.1|  (AL049708) pryM736A3a.1 (239AB) ...      52   9e-07
gi|17485345|ref|XP_066345.1|  (XM_066345) similar to bM150J2...      52   9e-07
gi|17536097|ref|NP_495889.1|  (NM_063488) T07D4.2.p [Caenorh...      50   6e-06
gi|16944410|emb|CAC28758.2|  (AL513463) conserved hypothetic...      39   0.010
gi|15793029|ref|NP_282852.1|  (NC_002163) putative homoserin...      36   0.067
gi|15187101|gb|AAK91292.1|AF395860_1  (AF395860) adult brain...      34   0.33
gi|145130|gb|AAC64901.1|  (M57500) phenol monooxygenase [Pla...      33   0.56
gi|400766|sp|P31020|PHEA_PSESP  PHENOL 2-MONOOXYGENASE (PHEN...      33   0.56
gi|1084273|pir||S41581  probable phosphoesterase (EC 3.1.-.-...      32   0.96
gi|5918760|gb|AAD56144.1|AF154675_4  (AF154675) MutB [Strept...      32   1.3
gi|15828727|ref|NP_326087.1|  (NC_002771) conserved hypothet...      32   1.6
gi|18976654|ref|NP_578011.1|  (NC_003413) hypothetical prote...      31   2.1
gi|13346908|gb|AAK19913.1|  (U61226) unknown [Leptospira int...      31   2.1
```

```
gi|17940021|gb|AAL49428.1|AF316500_15  (AF316500) unknown [L...   31   2.1
gi|17940046|gb|AAL49446.1|AF316556_1  (AF316556) unknown [Le...   31   2.1
gi|17940058|gb|AAL49454.1|AF316560_1  (AF316560) unknown [Le...   31   2.1
gi|17940055|gb|AAL49452.1|AF316559_1  (AF316559) unknown [Le...   31   2.1
gi|17940070|gb|AAL49462.1|AF316564_1  (AF316564) unknown [Le...   31   2.1
gi|6978034|gb|AAF34252.1|AF168003_7  (AF168003) putative asp...   31   2.1
gi|14521860|ref|NP_127336.1|  (NC_000868) hypothetical prote...   31   2.8
gi|16118348|gb|AAL12717.1|  (AY037277) envelope glycoprotein...   31   2.8
gi|17562430|ref|NP_504114.1|  (NM_071713) K09D9.10.p [Caenor...   31   2.8
gi|12082659|gb|AAG48566.1|AF238860_4  (AF238860) MutB [Strep...   31   2.8
gi|16118336|gb|AAL12707.1|  (AY037275) envelope glycoprotein...   30   3.7
gi|7162115|emb|CAB76673.1|  (AJ133765) invertase, putative [...   30   3.7
gi|1076927|pir||S22569  swi4 protein - fission yeast (Schizo...   30   4.8
gi|19114864|ref|NP_593952.1|  (NC_003424) mating-type switch...   30   4.8
gi|135075|sp|P26359|SWI4_SCHPO  Mating-type switching protei...   30   4.8
gi|18558414|ref|XP_106237.1|  (XM_106237) hypothetical prote...   30   4.8
gi|19703909|ref|NP_603471.1|  (NC_003454) Hypothetical cytos...   30   4.8
gi|16126513|ref|NP_421077.1|  (NC_002696) methylmalonate-sem...   30   6.2
gi|17019507|gb|AAA63583.2|  (U03969) dynein heavy chain isot...   30   6.2
gi|5817763|gb|AAD52906.1|AF142735_1  (AF142735) maturase-lik...   30   6.2
gi|4239911|dbj|BAA74745.1|  (AB016860) phenol hydroxylase al...   30   6.2
gi|1184955|gb|AAA87586.1|  (U46756) CLE7 [Gallus gallus]          29   8.2
gi|15616758|ref|NP_239970.1|  (NC_002528) 50S ribosomal prot...   29   8.2
gi|84040|pir||E22845  hypothetical protein 4 - Trypanosoma b...   29   8.2
gi|10443060|emb|CAC10467.1|  (AL353699) dJ710M3.1 (chromosom...   29   8.2

                             Alignments


>gi|20380139|gb|AAH28035.1|  (BC028035) Unknown (protein for MGC:40027)  [Homo sapiens]
          Length = 326

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 47/139 (33%), Positives = 72/139 (50%), Gaps = 19/139 (13%)
 Frame = +2

Query: 5   AGDY------KEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQ-------- 142
           AGD+       EV+ FN WL     ++YKI+IAGNH+L+FD      L+
Sbjct: 116 AGDFTELGLPSEVKKFNEWLGSLP-YEYKIVIAGNHELTFDQEFMADLIKQDFYYFPSVS 174

Query: 143 ----ELQQEIQYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDES-ENIWKN 307
               E  + +Q L  N IYL++S+V + G++I+GSP    ++   F +   ++    W
Sbjct: 175 KLKPENYENVQSLLTNCIYLQDSEVTVRGFRIYGSPWQPWFYGWGFNLPRGQALLEKWNL 234

Query: 308 IHDQTDIVLTHGPPYGHGD 364
           I +  DI++THGPP G  D
Sbjct: 235 IPEGVDILITHGPPLGFLD 253




>gi|20343349|ref|XP_110091.1|  (XM_110091) similar to bM150J22.1 (novel protein (ortholog of
human C22orf1)) [Mus musculus]
 gi|15982075|emb|CAC88864.1|  (AL513354) bM150J22.1 (novel protein (ortholog of human
C22orf1)) [Mus musculus]
          Length = 326

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 47/139 (33%), Positives = 72/139 (50%), Gaps = 19/139 (13%)
 Frame = +2
```

```
Query: 5   AGDY------KEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQ-------- 142
           AGD+       EV+ FN WL     ++YKI+IAGNH+L+FD      L+
Sbjct: 116 AGDFTELGLPSEVKKFNEWLGSLP-YEYKIVIAGNHELTFDQEFMADLIKQDFYYFPSVS 174

Query: 143 ----ELQQEIQYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDES-ENIWKN 307
               E  + +Q L  N IYL++S+V + G++I+GSP    ++   F +   ++    W
Sbjct: 175 KLKPENYENVQSLLTNCIYLQDSEVTVRGFRIYGSPWQPWFYGWGFNLPRGQALLEKWNL 234

Query: 308 IHDQTDIVLTHGPPYGHGD 364
           I +  DI++THGPP G  D
Sbjct: 235 IPEGVDILITHGPPLGFLD 253



>gi|7427509|ref|NP_001576.2| ☐ (NM_001585) chromosome 22 open reading frame 1; 239AB [Homo
sapiens]
 gi|14916945|sp|O15442|239A_HUMAN  Adult brain protein 239 (239AB)
 gi|7374112|gb|AAC51673.2| ☐ (U84894) 239AB [Homo sapiens]
          Length = 302

 Score = 77.8 bits (190), Expect = 2e-14
 Identities = 47/139 (33%), Positives = 72/139 (50%), Gaps = 19/139 (13%)
 Frame = +2

Query: 5   AGDY------KEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQ-------- 142
           AGD+       EV+ FN WL     ++YKI+IAGNH+L+FD      L+
Sbjct: 92  AGDFTELGLPSEVKKFNEWLGSLP-YEYKIVIAGNHELTFDQEFMADLIKQDFYYFPSVS 150

Query: 143 ----ELQQEIQYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDES-ENIWKN 307
               E  + +Q L  N IYL++S+V + G++I+GSP    ++   F +   ++    W
Sbjct: 151 KLKPENYENVQSLLTNCIYLQDSEVTVRGFRIYGSPWQPWFYGWGFNLPRGQALLEKWNL 210

Query: 308 IHDQTDIVLTHGPPYGHGD 364
           I +  DI++THGPP G  D
Sbjct: 211 IPEGVDILITHGPPLGFLD 229



>gi|4502485|ref|NP_001575.1| ☐ (NM_001584) chromosome 11 open reading frame 8; 239FB [Homo
sapiens]
 gi|3023214|sp|Q15777|239F_HUMAN  Fetal brain protein 239 (239FB)
 gi|1405360|gb|AAC50564.1| ☐ (U57911) 239FB gene product [Homo sapiens]
 gi|1584082|prf||2122285A  239FB gene [Homo sapiens]
          Length = 294

 Score = 76.3 bits (186), Expect = 6e-14
 Identities = 44/130 (33%), Positives = 75/130 (56%), Gaps = 15/130 (11%)
 Frame = +2

Query: 20  EVQNFNRWLDEQKEFKYKIIIAGNHDLSFDT--------RKYPQLLNYQELQQE----IQ 163
           EV+ FN WL     ++YKI+IAGNH+L+FD         + Y +  +  +L+ E    +Q
Sbjct: 95  EVKKFNDWLGNLP-YEYKIVIAGNHELTFDKEFMADLVKQDYYRFPSVSKLKPEDFDNVQ 153


Query: 164 YLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENI---WKNIHDQTDIVL 334
            L  N IYL++S+V ++G++I+G+P +   W+  +  +    +++   W  I +  DI++
Sbjct: 154 SLLTNSIYLQDSEVTVKGFRIYGAPWT--PWFNGWGFNLPRGQSLLDKWNLIPEGIDILM 211
```

```
Query: 335 THGPPYGHGD 364
           THGPP G  D
Sbjct: 212 THGPPLGFRD 221




>gi|12849370|dbj|BAB28313.1| ☐ (AK012553) data source:SPTR, source key:Q15777,
evidence:ISS~homolog to FETAL BRAIN PROTEIN 239 (239FB)~putative [Mus musculus]
          Length = 294

 Score = 76.3 bits (186), Expect = 6e-14
 Identities = 44/130 (33%), Positives = 75/130 (56%), Gaps = 15/130 (11%)
 Frame = +2

Query: 20  EVQNFNRWLDEQKEFKYKIIIAGNHDLSFDT--------RKYPQLLNYQELQQE----IQ 163
           EV+ FN WL     ++YKI+IAGNH+L+FD         + Y +  +  +L+ E    +Q
Sbjct: 95  EVKKFNDWLGNLP-YEYKIVIAGNHELTFDKEFMADLVKQDYYRFPSVSKLKPEDFDNVQ 153

Query: 164 YLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENI---WKNIHDQTDIVL 334
            L  N IYL++S+V ++G++I+G+P +   W+  +  +    +++   W  I +  DI++
Sbjct: 154 SLLTNSIYLQDSEVTVKGFRIYGAPWT--PWFNGWGFNLPRGQSLLDKWNLIPEGIDILM 211

Query: 335 THGPPYGHGD 364
           THGPP G  D
Sbjct: 212 THGPPLGFRD 221




>gi|7294897|gb|AAF50227.1| ☐ (AE003551) CG16717 gene product [Drosophila melanogaster]
          Length = 300

 Score = 67.4 bits (163), Expect = 3e-11
 Identities = 47/151 (31%), Positives = 71/151 (46%), Gaps = 31/151 (20%)
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTR------------------------ 115
           G  +EV+ FN W+      ++KI+IAGNH+LSFD
Sbjct: 77  GQLEEVEEFNTWIGALPH-RHKIVIAGNHELSFDRTFTHPFQKSKSHASSSKHTGMSILD 135

Query: 116 KYPQLLNYQE------LQQEIQYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQIS 277
             P L N +E        Q ++ +  N  YLE+  ++I G +I+GSP   E+   AF +
Sbjct: 136 DLPTLGNAKENLESAVQTQNVRDVLTNCRYLEDELLEIWGIQIYGSPWQPEFCRWAFNVP 195

Query: 278 PDES-ENIWKNIHDQTDIVLTHGPPYGHGDM 367
              +  + W  I +  DI++TH PP GHGD+
Sbjct: 196 RGTACLDKWNQIPEGIDILVTHTPPVGHGDL 226




>gi|17558234|ref|NP_505349.1| ☐ (NM_072948) C25E10.12.p [Caenorhabditis elegans]
 gi|3025037|sp|Q18161|YBPT_CAEEL  Hypothetical 31.8 kDa protein C25E10.12 in chromosome V
 gi|7496484|pir||T15613  hypothetical protein C25E10.12 - Caenorhabditis elegans
 gi|1226307|gb|AAA92317.1|  (U50311) Hypothetical protein C25E10.12 [Caenorhabditis elegans]
          Length = 281

 Score = 61.6 bits (148), Expect = 1e-09
 Identities = 42/125 (33%), Positives = 60/125 (47%), Gaps = 4/125 (3%)
```

```
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYP---QLLNYQELQQEIQY-LKK 175
           GD  EV  FN  +      K+KI+IAGNH+L F+  +     QL     L     Y L
Sbjct: 81  GDIGEVIKFNAEIGSLPH-KHKIVIAGNHELGFEDGEEMSERQLAGLNMLGINKAYELLS 139

Query: 176 NFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHDQTDIVLTHGPPYG 355
           N  YL +   +  G KI+G+P      Y  F+    +  + W  I  + D+++TH PP G
Sbjct: 140 NCTYLCDKSYEAYGLKIYGAPWHSMPGYSFFRQRGQKILHKWNQIPAKIDVLMTHTPPLG 199

Query: 356 HGDMS 370
           HGD +
Sbjct: 200 HGDFN 204




>gi|15131236|emb|CAC48257.1|  (AL356240) dJ873F21.1 (brain protein 239) [Homo sapiens]
          Length = 135

 Score = 58.2 bits (139), Expect = 2e-08
 Identities = 32/85 (37%), Positives = 53/85 (61%), Gaps = 12/85 (14%)
 Frame = +2

Query: 20  EVQNFNRWLDEQKEFKYKIIIAGNHDLSFDT--------RKYPQLLNYQELQQE----IQ 163
           EV+ FN WL     ++YKI+IAGNH+L+FD         + Y +  +  +L+ E    +Q
Sbjct: 52  EVKKFNDWLGNLP-YEYKIVIAGNHELTFDKEFMADLVKQDYYRFPSVSKLKPEDFDNVQ 110

Query: 164 YLKKNFIYLENSDVDIEGYKIWGSP 238
            L  N IYL++S+V ++G++I+G+P
Sbjct: 111 SLLTNSIYLQDSEVTVKGFRIYGAP 135



>gi|3025256|sp|Q21268|YMSO_CAEEL  Hypothetical 30.9 kDa protein K07C11.7 in chromosome V
          Length = 271

 Score = 55.8 bits (133), Expect = 8e-08
 Identities = 38/133 (28%), Positives = 63/133 (46%), Gaps = 14/133 (10%)
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEF--KYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKK-- 175
           G  +E+  FN   +E   F  KYK+++AGNH+L FD         + E Q E Q   K
Sbjct: 100 GKREELIKFN---EEMTRFPHKYKLVVAGNHELGFD---------HDENQGERQDADKGL 147

Query: 176 ----------NFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHDQTD 325
                     N  YL++  V I+G   +GS +     +  ++   ++    WK + + T+
Sbjct: 148 GTEDGYNILTNVTYLQDKGVTIDGVTFFGSSYHPLRGFPFYRNRAEQLAECWKAVPNDTN 207

Query: 326 IVLTHGPPYGHGD 364
           +++TH PP G+ D
Sbjct: 208 VLITHTPPLGYLD 220



>gi|17562344|ref|NP_505116.1|  (NM_072715) K07C11.7.p [Caenorhabditis elegans]
 gi|15055392|gb|AAA96177.2|  (U53336) Hypothetical protein K07C11.7 [Caenorhabditis elegans]
          Length = 290

 Score = 55.8 bits (133), Expect = 8e-08
```

```
 Identities = 38/133 (28%), Positives = 63/133 (46%), Gaps = 14/133 (10%)
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEF--KYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKK-- 175
           G  +E+  FN   +E   F  KYK+++AGNH+L FD         + E Q E Q   K
Sbjct: 100 GKREELIKFN---EEMTRFPHKYKLVVAGNHELGFD---------HDENQGERQDADKGL 147

Query: 176 ----------NFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHDQTD 325
                     N  YL++  V I+G   +GS +     +  ++   ++    WK + + T+
Sbjct: 148 GTEDGYNILTNVTYLQDKGVTIDGVTFFGSSYHPLRGFPFYRNRAEQLAECWKAVPNDTN 207

Query: 326 IVLTHGPPYGHGD 364
           +++TH PP G+ D
Sbjct: 208 VLITHTPPLGYLD 220



>gi|6562617|emb|CAB62592.1|  (AL049708) pryM736A3a.1 (239AB) [Homo sapiens]
          Length = 74

 Score = 52.4 bits (124), Expect = 9e-07
 Identities = 26/71 (36%), Positives = 41/71 (57%), Gaps = 12/71 (16%)
 Frame = +2

Query: 62  FKYKIIIAGNHDLSFDTRKYPQLLNYQ------------ELQQEIQYLKKNFIYLENSDV 205
           ++YKI+IAGNH+L+FD      L+               E  + +Q L  N IYL++S+V
Sbjct: 4   YEYKIVIAGNHELTFDQEFMADLIKQDFYYFPSVSKLKPENYENVQSLLTNCIYLQDSEV 63

Query: 206 DIEGYKIWGSP 238
            + G++I+GSP
Sbjct: 64  TVRGFRIYGSP 74



>gi|17485345|ref|XP_066345.1|  (XM_066345) similar to bM150J22.1 (novel protein (ortholog
of human C22orf1)) [Homo sapiens]
          Length = 285

 Score = 52.4 bits (124), Expect = 9e-07
 Identities = 26/71 (36%), Positives = 41/71 (57%), Gaps = 12/71 (16%)
 Frame = +2

Query: 62  FKYKIIIAGNHDLSFDTRKYPQLLNYQ------------ELQQEIQYLKKNFIYLENSDV 205
           ++YKI+IAGNH+L+FD      L+               E  + +Q L  N IYL++S+V
Sbjct: 82  YEYKIVIAGNHELTFDQEFMADLIKQDFYYFPSVSKLKPENYENVQSLLTNCIYLQDSEV 141

Query: 206 DIEGYKIWGSP 238
            + G++I+GSP
Sbjct: 142 TVRGFRIYGSP 152



 Score = 31.6 bits (70), Expect = 1.6
 Identities = 11/16 (68%), Positives = 16/16 (99%)
 Frame = +2
```

```
Query: 62  FKYKIIIAGNHDLSFD 109
           ++YKI+IAGNH+L+FD
Sbjct: 38  YEYKIVIAGNHELTFD 53



>gi|17536097|ref|NP_495889.1|   (NM_063488) T07D4.2.p [Caenorhabditis elegans]
 gi|3025317|sp|Q22306|YW12_CAEEL  Hypothetical 38.7 kDa protein T07D4.2 in chromosome II
 gi|7507383|pir||T24660  hypothetical protein T07D4.2 - Caenorhabditis elegans
 gi|3879542|emb|CAA90405.1|  (Z50071) cDNA EST EMBL:M89044 comes from this gene
[Caenorhabditis elegans]
          Length = 342

 Score = 49.7 bits (117), Expect = 6e-06
 Identities = 46/145 (31%), Positives = 64/145 (43%), Gaps = 24/145 (16%)
 Frame = +2

Query: 5   AGDY------KEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTR--KYPQLLNYQE--LQQ 154
           AGD+       EV NFN+ L + K + YK++I GNH+ +FD    K  Q    +E  L+Q
Sbjct: 122 AGDFTSCGLPNEVHNFNKLLGKLK-YSYKVVIGGNHECTFDDTFLKLKQESEPKEMALKQ 180

Query: 155 EI--------------QYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESE 292
            +              + L  N IYLE++        I     SL +     Q+ PD
Sbjct: 181 ALLSAIHSDSKGGISAKDLLSNAIYLEDNATKSRQLGI----QSLSWTTTIGQMEPDPGG 236

Query: 293 NIWKNIHDQTDIVLTHGPPYGHGDM 367
                     D++LTH PP GHGDM
Sbjct: 237 ---------VDVLLTHTPPLGHGDM 252



>gi|16944410|emb|CAC28758.2|  (AL513463) conserved hypothetical protein [Neurospora crassa]
          Length = 448

 Score = 38.9 bits (89), Expect = 0.010
 Identities = 14/27 (51%), Positives = 20/27 (73%)
 Frame = +2

Query: 287 SENIWKNIHDQTDIVLTHGPPYGHGDM 367
           ++ IW +I   TDI++TH PPYGH D+
Sbjct: 217 AKEIWSSIPPDTDIIVTHTPPYGHCDL 243



 Score = 32.3 bits (72), Expect = 0.96
 Identities = 16/50 (32%), Positives = 27/50 (54%)
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQE 157
           G   E+    +WL E+ +F+ K++IAGNHD + +     +   +Q L  E
Sbjct: 45  GSLSELTKAIQWL-EKADFEAKVVIAGNHDKALNPAPSSESSQHQALSPE 93



>gi|15793029|ref|NP_282852.1|  (NC_002163) putative homoserine O-succinyltransferase
[Campylobacter jejuni]
```

```
 gi|12230304|sp|Q9PLV2|META_CAMJE  Homoserine O-succinyltransferase (Homoserine O-
transsuccinylase) (HTS)
 gi|11257724|pir||B81271  probable homoserine O-succinyltransferase (EC 2.3.1.46) Cj1726c
[imported] - Campylobacter jejuni (strain NCTC 11168)
 gi|6969141|emb|CAB73712.1|  (AL139079) putative homoserine O-succinyltransferase
[Campylobacter jejuni]
          Length = 293

 Score = 36.2 bits (82), Expect = 0.067
 Identities = 29/111 (26%), Positives = 55/111 (49%), Gaps = 2/111 (1%)
 Frame = +2

Query: 14  YKEVQNFNRWLDEQKEFKYK-IIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYL 190
           +  ++ F + L+E K+ K+   I+ G      D  K   +  ++EL +   +LK+N
Sbjct: 81  FTHLEKFYKGLEEVKKHKFDGAIVTGAPVEQMDFEK---VAYWEELLEIFDFLKQN---- 133

Query: 191 ENSDVDIEGYKIWGSPHSLEYWYGAFQISPDES-ENIWKNIHDQTDIVLTH 340
               V    Y  WG+  +L+Y+YG  +IS D+    ++K+     D++LT+
Sbjct: 134 ----VTSSMYICWGAMAALKYFYGVDKISLDKKIFGVYKHDKVSPDLLLTN 180




>gi|15187101|gb|AAK91292.1|AF395860_1  (AF395860) adult brain protein 239-like protein
[Branchiostoma belcheri]
          Length = 87

 Score = 33.9 bits (76), Expect = 0.33
 Identities = 15/32 (46%), Positives = 21/32 (64%)
 Frame = +2

Query: 8   GDYKEVQNFNRWLDEQKEFKYKIIIAGNHDLS 103
           G  ++V+ FN WL       YK++IAGNHDL+
Sbjct: 57  GHLQKVKEFNTWLGTLPH-PYKVVIAGNHDLT 87



>gi|145130|gb|AAC64901.1|  (M57500) phenol monooxygenase [Plasmid pEST1226]
          Length = 610

 Score = 33.1 bits (74), Expect = 0.56
 Identities = 20/67 (29%), Positives = 34/67 (49%), Gaps = 1/67 (1%)
 Frame = +2

Query: 38  RWLDEQ-KEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           RW D   + +   +++ G     FD  + P  LN  E  Q +    +N + +E+ DV+I
Sbjct: 254 RWFDRMVRPWNEWLVVWG-----FDINQEPPKLNDDEAIQIV----RNLVGIEDLDVEIL 304

Query: 215 GYKIWGS 235
           GY +WG+
Sbjct: 305 GYSLWGN 311




>gi|400766|sp|P31020|PHEA_PSESP  PHENOL 2-MONOOXYGENASE (PHENOL HYDROXYLASE)
 gi|77810|pir||JQ1132  phenol 2-monooxygenase (EC 1.14.13.7) - Pseudomonas sp. Plasmid
EST1412
          Length = 607

 Score = 33.1 bits (74), Expect = 0.56
```

```
 Identities = 20/67 (29%), Positives = 34/67 (49%), Gaps = 1/67 (1%)
 Frame = +2

Query: 38  RWLDEQ-KEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
            RW D   + +   +++ G     FD  + P  LN  E  Q +    +N + +E+ DV+I
Sbjct: 254 RWFDRMVRPWNEWLVVWG-----FDINQEPPKLNDDEAIQIV----RNLVGIEDLDVEIL 304

Query: 215 GYKIWGS 235
            GY +WG+
Sbjct: 305 GYSLWGN 311



>gi|1084273|pir||S41581  probable phosphoesterase (EC 3.1.-.-) U - Methanothermus fervidus
          Length = 212

 Score = 32.3 bits (72), Expect = 0.96
 Identities = 22/63 (34%), Positives = 32/63 (49%), Gaps = 2/63 (3%)
 Frame = +2

Query: 170 KKNFIYLENSDVDIEGYKIWGSPHSLEYWYGA-FQISPDESENIWKNIHDQ-TDIVLTHG 343
            K   I + N+ ++I+G KI G   S    +   F+    E  N  KNI  + T I++TH
Sbjct: 73  KSGTINIHNNAIEIKGIKICGFGGSNPTPFDTPFEFQEKEIYNSLKNIISKDTTILVTHA 132

Query: 344 PPY 352
            PPY
Sbjct: 133 PPY 135



>gi|5918760|gb|AAD56144.1|AF154675_4  (AF154675) MutB [Streptococcus mutans]
          Length = 990

 Score = 32.0 bits (71), Expect = 1.3
 Identities = 17/53 (32%), Positives = 29/53 (54%)
 Frame = +2

Query: 50  EQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVD 208
            ++K F  KI +      SF  ++YP LLNY +L + +Q+    +I  E+ + D
Sbjct: 720 KEKYFYVKIYVNKQRQTSFLQKEYPLLLNYLKLPENLQWF---YIRYEDDEKD 769



>gi|15828727|ref|NP_326087.1|  (NC_002771) conserved hypothetical protein [Mycoplasma
pulmonis]
 gi|14089669|emb|CAC13429.1|  (AL445563) conserved hypothetical protein [Mycoplasma
pulmonis]
          Length = 787

 Score = 31.6 bits (70), Expect = 1.6
 Identities = 22/71 (30%), Positives = 35/71 (48%), Gaps = 5/71 (7%)
 Frame = -1

Query: 332 IQYLFNRVCFSKYFQIHLEISKML--HTSILRNVDCPIFYILQYQHQNFLNKQ---NSFL 168
            ++YL+      +Y  IH +  K+   HT I+ N+D  I+ +L+  +Q  LN +   N
Sbjct: 452 LEYLYKE----EYADIHWKQKKISENHTKIISNLDSKIYNLLERNNQVILNIKKLINEAS 507
```

```
Query: 167 NIKSLIRAPDN 135
           NI   I   DN
Sbjct: 508 NINEAIEKKDN 518



>gi|18976654|ref|NP_578011.1|  (NC_003413) hypothetical protein [Pyrococcus furiosus DSM
3638]
 gi|18892226|gb|AAL80406.1|  (AE010152) hypothetical protein [Pyrococcus furiosus DSM 3638]
          Length = 213

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 22/77 (28%), Positives = 36/77 (46%), Gaps = 11/77 (14%)
 Frame = +2

Query: 158 IQYLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHD------- 316
           + +L++  + + +  V+I+G  + G        +G   I+P  +  IW+   D
Sbjct: 73  LDFLEERGVSVHDKRVEIKGVGVVG--------FGGSNITPFST--IWEFSDDVIYSSLI 122

Query: 317 ----QTDIVLTHGPPYG 355
               Q DIVLTH PPYG
Sbjct: 123 RNYRQGDIVLTHAPPYG 139



>gi|13346908|gb|AAK19913.1|  (U61226) unknown [Leptospira interrogans]
 gi|17940049|gb|AAL49448.1|AF316557_1  (AF316557) unknown [Leptospira interrogans]
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  + D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165



>gi|17940021|gb|AAL49428.1|AF316500_15  (AF316500) unknown [Leptospira interrogans]
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  + D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165
```

```
>gi|17940046|gb|AAL49446.1|AF316556_1  (AF316556) unknown [Leptospira interrogans]
 gi|17940052|gb|AAL49450.1|AF316558_1  (AF316558) unknown [Leptospira interrogans]
 gi|17940064|gb|AAL49458.1|AF316562_1  (AF316562) unknown [Leptospira interrogans]
 gi|17940067|gb|AAL49460.1|AF316563_1  (AF316563) unknown [Leptospira interrogans]
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  + D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165



>gi|17940058|gb|AAL49454.1|AF316560_1  (AF316560) unknown [Leptospira interrogans]
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  + D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165



>gi|17940055|gb|AAL49452.1|AF316559_1  (AF316559) unknown [Leptospira interrogans]
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  + D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165



>gi|17940070|gb|AAL49462.1|AF316564_1  (AF316564) unknown [Leptospira interrogans]
```

```
          Length = 265

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 18/62 (29%), Positives = 30/62 (48%)
 Frame = +2

Query: 35  NRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQYLKKNFIYLENSDVDIE 214
           +++LD+  +  Y  I+  +  D     RKYP  LNY      I+ L    +++  S  D+
Sbjct: 106 SKFLDQDVDLVYGNILIDSKDNGIIERKYPDRLNY--FYWSIKSLCHQAVFIRKSLFDLY 163

Query: 215 GY 220
           GY
Sbjct: 164 GY 165




>gi|6978034|gb|AAF34252.1|AF168003_7  (AF168003) putative asparagine synthetase
[Desulfovibrio gigas]
          Length = 643

 Score = 31.2 bits (69), Expect = 2.1
 Identities = 12/36 (33%), Positives = 24/36 (66%)
 Frame = +2

Query: 125 QLLNYQELQQEIQYLKKNFIYLENSDVDIEGYKIWG 232
           ++ NYQE+++E+Q L + F    +++V ++ Y  WG
Sbjct: 81  EIYNYQEVRRELQALGRTFQTRSDTEVLLQAYLEWG 116




>gi|14521860|ref|NP_127336.1|  (NC_000868) hypothetical protein [Pyrococcus abyssi]
 gi|7447470|pir||H75015  probable phosphoesterase (EC 3.1.-.-) PAB1249 - Pyrococcus abyssi
(strain Orsay)
 gi|5459080|emb|CAB50566.1|  (AJ248288) hypothetical protein [Pyrococcus abyssi]
          Length = 213

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 14/29 (48%), Positives = 18/29 (61%)
 Frame = +2

Query: 266 FQISPDESENIWKNIHDQTDIVLTHGPPY 352
           ++ S DE  +  K  + Q DIVLTH PPY
Sbjct: 110 WEFSDDEIYSSLKKNYKQGDIVLTHSPPY 138




>gi|16118348|gb|AAL12717.1|  (AY037277) envelope glycoprotein [Human immunodeficiency virus
type 1]
          Length = 871

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 16/60 (26%), Positives = 31/60 (51%)
 Frame = +2

Query: 164 YLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHDQTDIVLTHG 343
           +L +NF+ +    V++ G++ W     L+YW+   Q    E +N    ++ + T IV+  G
Sbjct: 784 HLLRNFLLIVTRTVELLGHRGW---EILKYWWSLLQYWSQELKNSAISLLNTTTIVVAEG 840
```

```
>gi|17562430|ref|NP_504114.1|   (NM_071713) K09D9.10.p [Caenorhabditis elegans]
 gi|7206769|gb|AAF39928.1|  (AC006673) Hypothetical protein K09D9.10 [Caenorhabditis
elegans]
          Length = 300

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 23/66 (34%), Positives = 37/66 (55%)
 Frame = -1

Query: 278 EISKMLHTSILRNVDCPIFYILQYQHQNFLNKQNSFLNIKSLIRAPDNSVAMDIFQYQSL 99
           E+S M+H +I  N  C ++  + YQ+ N  N++NS + I  LI     SVA  I+ Y+ L
Sbjct: 92  ELSVMIHLAISLNRFCAVW--IPYQYHNIFNEKNSKILI-ILIWLFTGSVA--IYLYEIL 146

Query: 98  NHDFQQ 81
            H + +
Sbjct: 147 CHIYYE 152



>gi|12082659|gb|AAG48566.1|AF238860_4  (AF238860) MutB [Streptococcus mutans]
          Length = 990

 Score = 30.8 bits (68), Expect = 2.8
 Identities = 14/39 (35%), Positives = 23/39 (58%)
 Frame = +2

Query: 50  EQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQEIQY 166
           ++K F  KI +      SF  ++YP LLNY +L + +Q+
Sbjct: 720 KEKYFYAKIYVNKQRQTSFLQKEYPLLLNYLKLPENLQW 758



>gi|16118336|gb|AAL12707.1|  (AY037275) envelope glycoprotein [Human immunodeficiency virus
type 1]
          Length = 870

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 16/60 (26%), Positives = 31/60 (51%)
 Frame = +2

Query: 164 YLKKNFIYLENSDVDIEGYKIWGSPHSLEYWYGAFQISPDESENIWKNIHDQTDIVLTHG 343
           +L +NF+ +    V++ G++ W     L+YW+   Q    E +N   ++ + T IV+  G
Sbjct: 783 HLLRNFLLIVTRTVELLGHRGW---EILKYWWSLLQYWSQELKNSAISLLNTTAIVVAEG 839



>gi|7162115|emb|CAB76673.1|  (AJ133765) invertase, putative [Solanum tuberosum]
          Length = 585

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 21/91 (23%), Positives = 42/91 (46%)
 Frame = -1

Query: 371 WTYPHGHMEVHALIQYLFNRVCFSKYFQIHLEISKMLHTSILRNVDCPIFYILQYQHQNF 192
           W    G M  H  +  L+    F K+ +    +    HT    N +CP F+ +  ++ N
Sbjct: 204 WRIVIGSMRKHRGMALLYRSRDFIKWAKAQHPLHSSPHTG---NWECPDFFPVSLKNTNG 260
```

```
Query: 191 LNKQNSFLNIKSLIRAPDNSVAMDIFQYQSL 99
           L+      N+K +++   NS+ ++ F+Y ++
Sbjct: 261 LDASYRGKNVKYVLK---NSLDVNRFEYYTI 288



>gi|1076927|pir||S22569  swi4 protein - fission yeast (Schizosaccharomyces pombe)
          Length = 859

 Score = 30.0 bits (66), Expect = 4.8
 Identities = 19/83 (22%), Positives = 41/83 (48%), Gaps = 6/83 (7%)
 Frame = -1

Query: 332 IQYLFNRVCF---SKYFQIHLEISKMLHTSILRNVDCPIFYILQYQHQNFL---NKQNSF 171
           +Q  F+  C    SK  ++H+E  K LH+  +  +D  I Y++++  ++     N    F
Sbjct: 245 LQDFFSSKCIMSGSKIIELHMEKVKSLHSLSIICLDMAISYLMEFSLEDLFVASNFYQPF 304

Query: 170 LNIKSLIRAPDNSVAMDIFQYQS 102
            +I S++ +      +++F  Q+
Sbjct: 305 DSISSMVLSKQALEGLELFVNQT 327



>gi|19114864|ref|NP_593952.1|  (NC_003424) mating-type switching protein swi4
[Schizosaccharomyces pombe]
 gi|11230451|emb|CAB52164.2|  (AL109738) mating-type switching protein swi4
[Schizosaccharomyces pombe]
          Length = 1004

 Score = 30.0 bits (66), Expect = 4.8
 Identities = 19/83 (22%), Positives = 41/83 (48%), Gaps = 6/83 (7%)
 Frame = -1

Query: 332 IQYLFNRVCF---SKYFQIHLEISKMLHTSILRNVDCPIFYILQYQHQNFL---NKQNSF 171
           +Q  F+  C    SK  ++H+E  K LH+  +  +D  I Y++++  ++     N    F
Sbjct: 337 LQDFFSSKCIMSGSKIIELHMEKVKSLHSLSIICLDMAISYLMEFSLEDLFVASNFYQPF 396

Query: 170 LNIKSLIRAPDNSVAMDIFQYQS 102
            +I S++ +      +++F  Q+
Sbjct: 397 DSISSMVLSKQALEGLELFVNQT 419



>gi|135075|sp|P26359|SWI4_SCHPO  Mating-type switching protein swi4
 gi|101087|pir||S21964  mating-type switching protein swi4 - fission yeast
(Schizosaccharomyces pombe)
 gi|7492169|pir||T39178  mating-type switching protein swi4 - fission yeast
(Schizosaccharomyces pombe)
 gi|5113|emb|CAA43603.1|  (X61306) Swi4 [Schizosaccharomyces pombe]
          Length = 993

 Score = 30.0 bits (66), Expect = 4.8
 Identities = 19/83 (22%), Positives = 41/83 (48%), Gaps = 6/83 (7%)
 Frame = -1

Query: 332 IQYLFNRVCF---SKYFQIHLEISKMLHTSILRNVDCPIFYILQYQHQNFL---NKQNSF 171
           +Q  F+  C    SK  ++H+E  K LH+  +  +D  I Y++++  ++     N    F
Sbjct: 337 LQDFFSSKCIMSGSKIIELHMEKVKSLHSLSIICLDMAISYLMEFSLEDLFVASNFYQPF 396
```

```
Query: 170 LNIKSLIRAPDNSVAMDIFQYQS 102
            +I S++ +      +++F  Q+
Sbjct: 397 DSISSMVLSKQALEGLELFVNQT 419



>gi|18558414|ref|XP_106237.1|  (XM_106237) hypothetical protein XP_106237 [Homo sapiens]
          Length = 164

 Score = 30.0 bits (66), Expect = 4.8
 Identities = 26/86 (30%), Positives = 39/86 (45%), Gaps = 3/86 (3%)
 Frame = -1

Query: 365 YPHGHMEVHALIQYLFNRVCFSKYFQIHLEISKMLHTSILRNVDCPIFYILQYQ---HQN 195
           YPH   E H L + L N +C   + +  L +S  L ++    VDC I ++   +    QN
Sbjct: 62  YPHRKQEAHHLKKALSNVLCSVDHRKYELYLS--LESTDWNKVDCLIEFLRLSKILYSQN 119

Query: 194 FLNKQNSFLNIKSLIRAPDNSVAMDI 117
           FL   +   N K LI     SV + +
Sbjct: 120 FLKADSP--NSKKLILFHQKSVELKV 143



>gi|19703909|ref|NP_603471.1|  (NC_003454) Hypothetical cytosolic protein [Fusobacterium
nucleatum subsp. nucleatum ATCC 25586]
 gi|19714075|gb|AAL94770.1|  (AE010568) Hypothetical cytosolic protein [Fusobacterium
nucleatum subsp. nucleatum ATCC 25586]
          Length = 560

 Score = 30.0 bits (66), Expect = 4.8
 Identities = 33/128 (25%), Positives = 57/128 (43%), Gaps = 14/128 (10%)
 Frame = +3

Query: 3   VQVITKKFKILTDGLTNRKNLNIKSSLLEIMIQALILENIHSY*II-------RSSNKRF 161
           +++ +KK+  + D L  + N+NI+ +LL+  IQ  + EN +S  +           N +
Sbjct: 303 IEIRSKKYNYIEDVLNLKDNINIEKNLLKEKIQ--LNENSYSKLVTLLEKYDEEIKNNKK 360

Query: 162 NIQKRILF------IQKILMLILKDIKYGAVHIPQNTGMEHFRYLQMNLKIFG-KTYTIK 320
            I   I F      + KI   IL ++K       +N  ME  + ++  LK    KT  I+
Sbjct: 361 EIDNYIKFRGLNNLVDKINADILSEVK------KENEAMELLKPIKKQLKNANEKTKKIE 414

Query: 321 QILYQRMD 344
              Y  +D
Sbjct: 415 DSYYNMID 422



>gi|16126513|ref|NP_421077.1|  (NC_002696) methylmalonate-semialdehyde dehydrogenase,
putative [Caulobacter crescentus CB15]
 gi|13423787|gb|AAK24245.1|  (AE005897) methylmalonate-semialdehyde dehydrogenase, putative
[Caulobacter crescentus CB15]
          Length = 498

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 15/40 (37%), Positives = 23/40 (57%), Gaps = 1/40 (2%)
 Frame = -3
```

```
Query: 336 VNTIS-VQSCMFFQIFSDSSGDIQNAPYQYSKECGLPHIL 220
           +NT++ + S    ++ +DS GDIQ         CG+PHIL
Sbjct: 80  MNTLAEILSSEHGKVIADSKGDIQRGLEVIEFACGIPHIL 119




>gi|17019507|gb|AAA63583.2|  (U03969) dynein heavy chain isotype 1B [Tripneustes gratilla]
          Length = 4318

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 25/99 (25%), Positives = 39/99 (39%)
 Frame = +3

Query: 66   NIKSSLLEIMIQALILENIHSY*IIRSSNKRFNIQKRILFIQKILMLILKDIKYGAVHIP 245
            N    +LE+ ++ALIL+ IHS  +++   K           QK L   L + K   + +
Sbjct: 1594 NAAIHVLELKLKALILDTIHSLDVVQLLQKENVTSLDNWLWQKQLRYYLDNTKTAIMRMV 1653

Query: 246  QNTGMEHFRYLQMNLKIFGKTYTIKQILYQRMDLHMAMG 362
                   + Y     K+     T K  L     +HM MG
Sbjct: 1654 DAEFQYTYEYQGNAAKLVHTPLTDKCYLTLTQGMHMGMG 1692




>gi|5817763|gb|AAD52906.1|AF142735_1  (AF142735) maturase-like protein [Swainsona
pterostylis]
          Length = 506

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 21/72 (29%), Positives = 37/72 (51%), Gaps = 7/72 (9%)
 Frame = -1

Query: 344 VHALIQYLFNRVCFSKY-------FQIHLEISKMLHTSILRNVDCPIFYILQYQHQNFLN 186
           +H++  +L ++  +  Y       + IHLEI   +    ++  D PIF++L+    NF N
Sbjct: 129 IHSIFPFLEDKFPYLNYVSDIRIPYPIHLEILVQILRYWVK--DAPIFHLLRLFLYNFCN 186

Query: 185 KQNSFLNIKSLI 150
           + NSF+  K  I
Sbjct: 187 R-NSFITPKKSI 197




>gi|4239911|dbj|BAA74745.1|  (AB016860) phenol hydroxylase alpha subunit [Pseudomonas
putida]
          Length = 377

 Score = 29.6 bits (65), Expect = 6.2
 Identities = 15/49 (30%), Positives = 27/49 (54%)
 Frame = +2

Query: 11  DYKEVQNFNRWLDEQKEFKYKIIIAGNHDLSFDTRKYPQLLNYQELQQE 157
           +Y     F+ W+ +Q E  +   +AG +  +F+    P+L ++QE QQE
Sbjct: 281 NYAAAAGFHTWVPKQDEMAW---LAGKYPETFERYYKPRLDHWQERQQE 326




>gi|1184955|gb|AAA87586.1|  (U46756) CLE7 [Gallus gallus]
          Length = 239
```

```
 Score = 29.3 bits (64), Expect = 8.2
 Identities = 24/74 (32%), Positives = 33/74 (44%), Gaps = 10/74 (13%)
 Frame = +2

Query: 11  DYKEVQNFNRWLDEQKEFKYKIIIAGN------HDLSFDTRKYPQLLN----YQELQQEI 160
           D  E +NF  WL++QK   YKI   GN       D      KY + +N     QE Q+ +
Sbjct: 21  DETEFRNFIVWLEDQKIRHYKIEDRGNLRNIHSDDWPKSIEKYMKDVNCPFKMQERQETV 80

Query: 161 QYLKKNFIYLENSD 202
            +L    + LE  D
Sbjct: 81  DWLLGLAVRLEYGD 94



>gi|15616758|ref|NP_239970.1|  (NC_002528) 50S ribosomal protein L25 [Buchnera sp. APS]
 gi|11387018|sp|P57238|RL25_BUCAI  50S ribosomal protein L25
 gi|10038821|dbj|BAB12856.1|  (AP001118) 50S ribosomal protein L25 [Buchnera sp. APS]
          Length = 95

 Score = 29.3 bits (64), Expect = 8.2
 Identities = 13/31 (41%), Positives = 21/31 (66%)
 Frame = +3

Query: 135 IIRSSNKRFNIQKRILFIQKILMLILKDIKY 227
           +I   N  FN+QK+I F ++ L+L ++D KY
Sbjct: 40  LILDHNTTFNLQKKIEFYKENLLLCVQDKKY 70



>gi|84040|pir||E22845  hypothetical protein 4 - Trypanosoma brucei mitochondrion
          Length = 445

 Score = 29.3 bits (64), Expect = 8.2
 Identities = 18/62 (29%), Positives = 32/62 (51%)
 Frame = -1

Query: 335 LIQYLFNRVCFSKYFQIHLEISKMLHTSILRNVDCPIFYILQYQHQNFLNKQNSFLNIKS 156
           L+Q  +  +C   +F I L I  + +  ++ N+     +I  YQ   F   + SF+NIK+
Sbjct: 367 LLQLSYLCICIGFFFIIWLIIIYIFYFRLIVNI-----FIFSYQFLGFWVVKLSFINIKN 421

Query: 155 LI 150
           L+
Sbjct: 422 LL 423



>gi|10443060|emb|CAC10467.1|  (AL353699) dJ710M3.1 (chromosome 11 open reading frame 8(Fetal
brain protein 239)) [Homo sapiens]
          Length = 115

 Score = 29.3 bits (64), Expect = 8.2
 Identities = 12/40 (30%), Positives = 21/40 (52%), Gaps = 3/40 (7%)
 Frame = +2

Query: 254 WYGAFQISPDESENI---WKNIHDQTDIVLTHGPPYGHGD 364
           W+  +  +    +++   W  I +  DI++THGPP G  D
Sbjct: 3   WFNGWGFNLPRGQSLLDKWNLIPEGIDILMTHGPPLGFRD 42
```

```
  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 205,884,535
Number of Sequences: 919285
Number of extensions: 4079667
Number of successful extensions: 15015
Number of sequences better than 10.0: 96
Number of HSP's better than 10.0 without gapping: 14619
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 14994
length of database: 288,558,979
effective HSP length: 100
effective length of database: 196,630,479
effective search space used: 4719131496
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020955900-023914-19845
```

7.1.15 Query= hy-15_20240017

```
         (363 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 40 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|14133207|dbj|BAA25482.2|  (AB011128) KIAA0556 protein [Ho...    88   2e-17
gi|7513017|pir||T00330  hypothetical protein KIAA0556 - huma...    88   2e-17
gi|15318238|ref|XP_044632.2|  (XM_044632) KIAA0556 protein [...    88   2e-17
gi|17508039|ref|NP_491836.1|  (NM_059435) K04F10.2.p [Caenor...    63   7e-10
gi|19569994|gb|AAL92295.1|  (AC115593) hypothetical protein ...    50   6e-06
gi|13242506|ref|NP_077519.1|  (NC_002687) EsV-1-34 [Ectocarp...    34   0.34
gi|18309237|ref|NP_561171.1|  (NC_003366) hypothetical prote...    33   0.44
gi|11467581|ref|NP_043727.1|  (NC_001715) NADH dehydrogenase...    31   2.2
gi|9800496|gb|AAF99319.1|AF281324_1  (AF281324) NADH dehydro...    31   2.2
gi|17540118|ref|NP_501233.1|  (NM_068832) F32E10.1.p [Caenor...    31   2.2
gi|15897391|ref|NP_341996.1|  (NC_002754) Conserved hypothet...    31   2.2
gi|15894245|ref|NP_347594.1|  (NC_003030) Xre family DNA-bin...    31   2.9
gi|169491|gb|AAA69780.1|  (L06137) proteinase inhibitor I [S...    31   2.9
gi|15596169|ref|NP_249663.1|  (NC_002516) TolB protein [Pseu...    30   3.7
gi|15611885|ref|NP_223536.1|  (NC_000921) CYSTEINYL-TRNA SYN...    30   3.7
gi|100458|pir||S25159  proteinase inhibitor I - potato (cult...    30   3.7
gi|13123736|gb|AAK12957.1|AF343914_10  (AF343914) unknown [C...    30   3.7
gi|15213122|gb|AAK85734.1|  (AY043279) beta-G spectrin [Brug...    30   4.9
gi|3660690|gb|AAC72317.1|  (AF063428) period [Biston betularia]    30   4.9
gi|14520380|ref|NP_125855.1|  (NC_000868) hypothetical prote...    30   4.9
gi|14764011|ref|XP_035625.1|  (XM_035625) KIAA0876 protein [...    30   6.4
gi|3882281|dbj|BAA34500.1|  (AB018323) KIAA0780 protein [Hom...    30   6.4
gi|18044920|gb|AAH20180.1|AAH20180  (BC020180) Similar to ge...    30   6.4
gi|20348507|ref|XP_110245.1|  (XM_110245) similar to KIAA067...    30   6.4
```

```
gi|15611460|ref|NP_223111.1|  (NC_000921) RNA POLYMERASE SIG...    30   6.4
gi|14133223|dbj|BAA74899.2|  (AB020683) KIAA0876 protein [Ho...    30   6.4
gi|10567164|dbj|BAB16102.1|  (AB037901) gene amplified in sq...    30   6.4
gi|15606531|ref|NP_213911.1|  (NC_000918) hypothetical prote...    29   8.3
gi|15594935|ref|NP_212724.1|  (NC_001318) dimethyladenosine ...    29   8.3
gi|16760984|ref|NP_456601.1|  (NC_003198) putative propanol ...    29   8.3
gi|169542|gb|AAA72133.1|  (L06985) proteinase inhibitor I [S...    29   8.3
```

Alignments

```
>gi|14133207|dbj|BAA25482.2|     (AB011128) KIAA0556 protein [Homo sapiens]
          Length = 1625

 Score = 87.8 bits (216), Expect = 2e-17
 Identities = 49/123 (39%), Positives = 69/123 (55%), Gaps = 3/123 (2%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            D +YIGL G+E++D  GE++      + A P SV  L  +  DVRTPDKL++ VN T DG
Sbjct: 1419 DPYYIGLTGLELYDERGEKIPLSENNIAAFPDSVNSLEGVGGDVRTPDKLIDQVNDTSDG 1478

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPVRG 11
            + MWL+P +                      VN++ ++FD P  +S I  +NY+KTP RG
Sbjct: 1479 RHMWLAPILPGL-------------------VNRVYVIFDLPTTVSMIKLWNYAKTPHRG 1519

Query: 10   VKE 2
            VKE
Sbjct: 1520 VKE 1522




 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 41/125 (32%), Positives = 63/125 (49%), Gaps = 6/125 (4%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            DLHY+GL G+E+    G+ L     ++ A P  + ELP+   D RT DKL++  NIT++
Sbjct: 1229 DLHYLGLTGLEVVGKEGQALPIHLHQISASPRDLNELPEYSDDSRTLDKLIDGTNITMED 1288

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTP--- 20
            + MWL PF             S   D +      + +  D  + I+ + F+NY+K+P
Sbjct: 1289 EHMWLIPF-------------SPGLDHV------VTIRLDRAESIAGLRFWNYNKSPEDT 1329

Query: 19   VRGVK 5
             RG K
Sbjct: 1330 YRGAK 1334




 Score = 58.9 bits (141), Expect = 1e-08
 Identities = 38/124 (30%), Positives = 57/124 (45%), Gaps = 4/124 (3%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL-YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQ 185
            D HY+GLNGIEIF   GE +    + A P  +  LP    D R    L++ VN T D
Sbjct: 994  DRHYVGLNGIEIFSSKGEPVQISNIKADPPDINILPAYGKDPRVVTNLIDGVNRTQDDMH 1053
```

```
Query: 184  MWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPV---R 14
            +WL+PF                      + + + + F  P  ++ I  +NY+K+ +   R
Sbjct: 1054 VWLAPFTRG-------------------RSHSITIDFTHPCHVALIRIWNYNKSRIHSFR 1094

Query: 13   GVKE 2
            GVK+
Sbjct: 1095 GVKD 1098



>gi|7513017|pir||T00330  hypothetical protein KIAA0556 - human (fragment)
          Length = 1081

 Score = 87.8 bits (216), Expect = 2e-17
 Identities = 49/123 (39%), Positives = 69/123 (55%), Gaps = 3/123 (2%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            D +YIGL G+E++D  GE++      + A P SV  L  +  DVRTPDKL++ VN T DG
Sbjct: 875  DPYYIGLTGLELYDERGEKIPLSENNIAAFPDSVNSLEGVGGDVRTPDKLIDQVNDTSDG 934

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPVRG 11
            + MWL+P +                      VN++ ++FD P  +S I  +NY+KTP RG
Sbjct: 935  RHMWLAPILPGL-------------------VNRVYVIFDLPTTVSMIKLWNYAKTPHRG 975

Query: 10   VKE 2
            VKE
Sbjct: 976  VKE 978



 Score = 67.8 bits (164), Expect = 2e-11
 Identities = 41/125 (32%), Positives = 63/125 (49%), Gaps = 6/125 (4%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            DLHY+GL G+E+    G+ L     ++ A P  + ELP+   D RT DKL++  NIT++
Sbjct: 685  DLHYLGLTGLEVVGKEGQALPIHLHQISASPRDLNELPEYSDDSRTLDKLIDGTNITMED 744

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTP--- 20
            + MWL PF             S   D +      + +  D  + I+ + F+NY+K+P
Sbjct: 745  EHMWLIPF-------------SPGLDHV------VTIRLDRAESIAGLRFWNYNKSPEDT 785

Query: 19   VRGVK 5
             RG K
Sbjct: 786  YRGAK 790



 Score = 58.9 bits (141), Expect = 1e-08
 Identities = 38/124 (30%), Positives = 57/124 (45%), Gaps = 4/124 (3%)
 Frame = -3

Query: 361 DLHYIGLNGIEIFDPNGEQL-YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQ 185
           D HY+GLNGIEIF   GE +    + A P  +  LP    D R    L++ VN T D
Sbjct: 450 DRHYVGLNGIEIFSSKGEPVQISNIKADPPDINILPAYGKDPRVVTNLIDGVNRTQDDMH 509
```

```
Query: 184 MWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPV---R 14
           +WL+PF                      + + + + F  P  ++  I  +NY+K+ +   R
Sbjct: 510 VWLAPFTRG-------------------RSHSITIDFTHPCHVALIRIWNYNKSRIHSFR 550

Query: 13  GVKE 2
           GVK+
Sbjct: 551 GVKD 554



>gi|15318238|ref|XP_044632.2|  (XM_044632) KIAA0556 protein [Homo sapiens]
          Length = 1592

 Score = 87.8 bits (216), Expect = 2e-17
 Identities = 49/123 (39%), Positives = 69/123 (55%), Gaps = 3/123 (2%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            D +YIGL G+E++D  GE++      + A P SV  L  +  DVRTPDKL++ VN T DG
Sbjct: 1386 DPYYIGLTGLELYDERGEKIPLSENNIAAFPDSVNSLEGVGGDVRTPDKLIDQVNDTSDG 1445

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPVRG 11
            + MWL+P +                      VN++ ++FD P  +S I  +NY+KTP RG
Sbjct: 1446 RHMWLAPILPGL-------------------VNRVYVIFDLPTTVSMIKLWNYAKTPHRG 1486

Query: 10   VKE 2
            VKE
Sbjct: 1487 VKE 1489



 Score = 65.9 bits (159), Expect = 8e-11
 Identities = 40/125 (32%), Positives = 62/125 (49%), Gaps = 6/125 (4%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            DLHY+GL G+E+    G+ L     ++ A P  + ELP+   D R  DKL++   NIT++
Sbjct: 1196 DLHYLGLTGLEVVGKEGQALPIHLHQISASPRDLNELPEYSDDSRALDKLIDGTNITMED 1255

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTP--- 20
            + MWL PF             S   D +      + +  D  + I+ + F+NY+K+P
Sbjct: 1256 EHMWLIPF-------------SPGLDHV------VTIRLDRAESIAGLRFWNYNKSPEDT 1296

Query: 19   VRGVK 5
             RG K
Sbjct: 1297 YRGAK 1301



 Score = 58.9 bits (141), Expect = 1e-08
 Identities = 38/124 (30%), Positives = 57/124 (45%), Gaps = 4/124 (3%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL-YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQ 185
            D HY+GLNGIEIF   GE +    + A P  +  LP    D R    L++ VN T D
Sbjct: 961  DRHYVGLNGIEIFSSKGEPVQISNIKADPPDINILPAYGKDPRVVTNLIDGVNRTQDDMH 1020
```

```
Query: 184  MWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPV---R 14
            +WL+PF                      + + + + F  P  ++ I  +NY+K+ +   R
Sbjct: 1021 VWLAPFTRG-------------------RSHSITIDFTHPCHVALIRIWNYNKSRIHSFR 1061

Query: 13   GVKE 2
            GVK+
Sbjct: 1062 GVKD 1065



>gi|17508039|ref|NP_491836.1|  (NM_059435) K04F10.2.p [Caenorhabditis elegans]
 gi|2773240|gb|AAB96751.1|  (AF039719) Hypothetical protein K04F10.2 [Caenorhabditis
elegans]
          Length = 512

 Score = 62.8 bits (151), Expect = 7e-10
 Identities = 42/122 (34%), Positives = 59/122 (47%), Gaps = 3/122 (2%)
 Frame = -3

Query: 361 DLHYIGLNGIEIFDPNGEQLYPK---LFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
           D  YIGLNGIE+++  GE +  +   L A P SV  LP ++ D+RT + L+   N T
Sbjct: 376 DEFYIGLNGIELYNRKGELMKIREHNLAAFPESVNILPNIKNDLRTSNNLITQPNDTDIA 435

Query: 190 KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKTPVRG 11
           + MWL+  + +   R+                     +FD    IS I  YNY KTP RG
Sbjct: 436 RNMWLTALLPNRCARV-------------------FFVFDVQTYISKIVIYNYRKTPERG 476

Query: 10  VK 5
           V+
Sbjct: 477 VR 478



 Score = 31.2 bits (69), Expect = 2.2
 Identities = 19/53 (35%), Positives = 28/53 (51%)
 Frame = -3

Query: 361 DLHYIGLNGIEIFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNI 203
           D HYIGLN +EIF   GE+    +  K Y+   +   ++ VR  +  L+ V I
Sbjct: 54  DSHYIGLNSVEIFTSTGER---AVIDKNYNASRV-HAQIGVRYLEMYLDGVGI 102



>gi|19569994|gb|AAL92295.1|  (AC115593) hypothetical protein [Dictyostelium discoideum]
          Length = 2472

 Score = 49.7 bits (117), Expect = 6e-06
 Identities = 35/116 (30%), Positives = 54/116 (46%), Gaps = 3/116 (2%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQL---YPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDG 191
            D  Y+GL  I +FD   +Q+      + A+P  + ++P    D RT DKL++  N+T +
Sbjct: 2020 DRFYMGLTSIMVFDNEYKQIPIQASNITAQPKDINDVPGHSGDYRTLDKLIDGHNVTTND 2079

Query: 190  KQMWLSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDTPQQISAISFYNYSKT 23
            K MWL PF  S    L  ++  +VF                   IS I  +NY+K+
Sbjct: 2080 KHMWLIPFSPSKEHVLTIDL-GNVF------------------SISCIRIWNYNKS 2116
```

```
 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 41/153 (26%), Positives = 60/153 (38%), Gaps = 33/153 (21%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQLYPKLFAKPYSVKELP------------------------- 257
            D HY+GL+GI+IFD NG   Y KL      +K  P
Sbjct: 1575 DQHYVGLSGIDIFDSNGS--YVKLKDPKSQIKAFPSDLNDINNNNNNNNNNNNNNVNGHS 1632

Query: 256  -----QMEMDVRTPDKLLNNVNITLDGKQMWLSPFVNSYTDRLKTNINSSVFDQISYQVN 92
                 Q   D RT DK+++    T D   +WL+PF                    + QVN
Sbjct: 1633 NRNHYQHINDPRTIDKIMDGHPRTCDDNHLWLAPFT-------------------AGQVN 1673

Query: 91   KLILLFDTPQQISAISFYNYSKTPV---RGVKE 2
             + +  D+   +S I  +NY+K+ +   RG K+
Sbjct: 1674 TITITLDSKITLSLIRIWNYNKSRIHSSRGAKD 1706




 Score = 29.3 bits (64), Expect = 8.3
 Identities = 32/145 (22%), Positives = 58/145 (39%), Gaps = 25/145 (17%)
 Frame = -3

Query: 361  DLHYIGLNGIEIFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQM 182
            D +++GL+ IEI+D N   + P          +   +  DV+   KL+ N N++    +
Sbjct: 2238 DRNFVGLHAIEIYDQNDTIIKP---------NKSSVLSNDVKNSYKLIGN-NLSFQESEG 2287

Query: 181  WLSPFVN-SYTDRLKTNINSSVFDQ------------ISY------------QVNKLILL 77
            +L P+ + S+ +    N  S ++ Q            +SY            + N +
Sbjct: 2288 FLIPYFSVSFGNSNNNNGMSPLYHQQQQQQYQLGGNSVSYNSRGLEIPSFSKESNVIYFY 2347

Query: 76   FDTPQQISAISFYNYSKTPVRGVKE 2
            FD+P        +NY K     +KE
Sbjct: 2348 FDSPITGGKNKIWNYFKNISPAIKE 2372




>gi|13242506|ref|NP_077519.1|  (NC_002687) EsV-1-34 [Ectocarpus siliculosus virus]
 gi|13177316|gb|AAK14460.1|AF204951_34  (AF204951) EsV-1-34 [Ectocarpus siliculosus virus]
          Length = 742

 Score = 33.9 bits (76), Expect = 0.34
 Identities = 19/51 (37%), Positives = 31/51 (60%), Gaps = 3/51 (5%)
 Frame = -3

Query: 229 DKLLNNVNITLDGKQMWLSPFVNSYTDRLKTNINSSV---FDQISYQVNKL 86
           D+L+ N+N +++G    L   VN+YT  ++T IN SV    D I+  +NK+
Sbjct: 283 DELVGNLNKSINGMAFGLETAVNTYTGGIETVINKSVGGITDGINNNMNKV 333




>gi|18309237|ref|NP_561171.1|  (NC_003366) hypothetical protein [Clostridium perfringens]
 gi|18143913|dbj|BAB79961.1|  (AP003186) hypothetical protein [Clostridium perfringens]
```

```
          Length = 104

 Score = 33.5 bits (75), Expect = 0.44
 Identities = 15/33 (45%), Positives = 21/33 (63%)
 Frame = -3

Query: 112 QISYQVNKLILLFDTPQQISAISFYNYSKTPVR 14
           ++   + KL   +F TP  IS I+FYNY KT +R
Sbjct: 47  ELGNYIPKLFFIFFTPIVISLITFYNYMKTELR 79



>gi|11467581|ref|NP_043727.1|  (NC_001715) NADH dehydrogenase, subunit 5 [Allomyces
macrogynus]
 gi|1709391|sp|P50365|NU5M_ALLMA  NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5
 gi|2147767|pir||S63645  NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Allomyces
macrogynus mitochondrion
 gi|604921|gb|AAB05846.1|  (U17010) NADH dehydrogenase, subunit 5 [Allomyces macrogynus]
 gi|1236411|gb|AAC49228.1|  (U41288) NADH dehydrogenase, subunit 5 [Allomyces macrogynus]
          Length = 641

 Score = 31.2 bits (69), Expect = 2.2
 Identities = 24/74 (32%), Positives = 32/74 (42%), Gaps = 8/74 (10%)
 Frame = -3

Query: 355 HYIGLNGIEIFDPNGEQLYPKLFAKPYSVK------ELPQMEMDVRTPDKLLNNVNIT-- 200
           HY+G +   I    G Q YP LFA     +        + P+     +   P+ LLN VNIT
Sbjct: 499 HYLGFDIEFISTSFGVQFYP-LFASLLGISLALIVLKDPKKMHSIEKPEGLLNTVNITRW 557

Query: 199 LDGKQMWLSPFVNS 158
           L  K  W     N+
Sbjct: 558 LSSKSYWFDNVYNT 571



>gi|9800496|gb|AAF99319.1|AF281324_1  (AF281324) NADH dehydrogenase subunit 5 [Allomyces
arbuscula]
          Length = 212

 Score = 31.2 bits (69), Expect = 2.2
 Identities = 24/74 (32%), Positives = 32/74 (42%), Gaps = 8/74 (10%)
 Frame = -3

Query: 355 HYIGLNGIEIFDPNGEQLYPKLFAKPYSVK------ELPQMEMDVRTPDKLLNNVNIT-- 200
           HY+G +   I    G Q YP LFA     +        + P+     +   P+ LLN VNIT
Sbjct: 70  HYLGFDIEFISTSFGVQFYP-LFASLLGISLALIVLKDPKKMHSIEKPEGLLNTVNITRW 128

Query: 199 LDGKQMWLSPFVNS 158
           L  K  W     N+
Sbjct: 129 LSSKSYWFDNVYNT 142



>gi|17540118|ref|NP_501233.1|   (NM_068832) F32E10.1.p [Caenorhabditis elegans]
 gi|1118177|gb|AAA83359.1|  (U41992) Hypothetical protein F32E10.1 [Caenorhabditis elegans]
          Length = 754
```

```
 Score = 31.2 bits (69), Expect = 2.2
 Identities = 26/76 (34%), Positives = 36/76 (47%), Gaps = 1/76 (1%)
 Frame = -3

Query: 295 KLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQMWLSPFVNSYTDRLKTNINSSVF 116
           KL  +   V++  Q+  D   PD + N VNIT DGK +W S     +      N  S  F
Sbjct: 31  KLEQRNTDVRQRIQLIQDFEMPD-VSNTVNITPDGKYVWASGNYKPWLRCYDLNNLSLKF 89

Query: 115 DQ-ISYQVNKLILLFD 71
           ++ +   V KLI L D
Sbjct: 90  ERGLDADVIKLIPLSD 105




>gi|15897391|ref|NP_341996.1|  (NC_002754) Conserved hypothetical protein [Sulfolobus
solfataricus]
 gi|13813618|gb|AAK40786.1|  (AE006678) Conserved hypothetical protein [Sulfolobus
solfataricus]
          Length = 214

 Score = 31.2 bits (69), Expect = 2.2
 Identities = 18/48 (37%), Positives = 24/48 (49%)
 Frame = +2

Query: 5   LYSSHRCFTIIVKGYGRYLLRSIKQQNQFVDLVTNLIKNTRIYIRFQS 148
           LY    +   +++KGY  L  SI   N FVD + N I    IY+ F S
Sbjct: 60  LYYLEQLNYLVIKGYYSELFTSIFITNSFVDFIFNFISLYVIYLIFGS 107




>gi|15894245|ref|NP_347594.1|  (NC_003030) Xre family DNA-binding domain and TPR-repeat-
containing
           protein [Clostridium acetobutylicum]
 gi|15023862|gb|AAK78934.1|AE007611_2  (AE007611) Xre family DNA-binding domain and TPR-
repeat-containing protein [Clostridium acetobutylicum]
          Length = 421

 Score = 30.8 bits (68), Expect = 2.9
 Identities = 18/75 (24%), Positives = 38/75 (50%), Gaps = 3/75 (4%)
 Frame = +1

Query: 1   IPLLLSQVFYYNCKRIW-QIFIEEYQITKSIC*LSN--QSDQKHSNLYSFLIYLCMNQRK 171
           I +L+   +Y     I  + +   Y++ K +  L     +D +H+ +YS +++LC+  +K
Sbjct: 188 IKMLIESYYYETLSYILCEDYDNAYELEKELLELIEYVDTDNQHAKIYSVMMFLCL--KK 245


Query: 172 EITTFAYHLRLCQHY 216
            I  F Y+    ++Y
Sbjct: 246 NIDQFDYYKEKIENY 260





>gi|169491|gb|AAA69780.1|  (L06137) proteinase inhibitor I [Solanum tuberosum]
          Length = 107

 Score = 30.8 bits (68), Expect = 2.9
```

```
 Identities = 24/70 (34%), Positives = 36/70 (51%), Gaps = 1/70 (1%)
 Frame = -3

Query: 325 FDPNGEQLYPKLFAKPYSV-KELPQMEMDVRTPDKLLNNVNITLDGKQMWLSPFVNSYTD 149
           F+ NG+Q +P+L   P  + KE+ + E       + L+NNV I L+G  + +    N
Sbjct: 39  FECNGKQFWPELIGVPTKLAKEIIEKE------NSLINNVQILLNGSPVTMDYRCNRV-- 90

Query: 148 RLKTNINSSV 119
           RL  NI  SV
Sbjct: 91  RLFDNILGSV 100




>gi|15596169|ref|NP_249663.1|  (NC_002516) TolB protein [Pseudomonas aeruginosa]
 gi|12644681|sp|P50601|TOLB_PSEAE  TolB protein precursor
 gi|11352644|pir||F83525  TolB protein PA0972 [imported] - Pseudomonas aeruginosa (strain
PAO1)
 gi|5733837|gb|AAC44661.2|  (U39558) TolB [Pseudomonas aeruginosa]
 gi|9946879|gb|AAG04361.1|AE004530_14  (AE004530) TolB protein [Pseudomonas aeruginosa]
          Length = 432

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 22/91 (24%), Positives = 39/91 (42%), Gaps = 6/91 (6%)
 Frame = -3

Query: 355 HYIGLNGIEIFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITL------D 194
           ++ GLNG   F P+G +L            E+  M++  R   +L NN+ I        D
Sbjct: 240 NFEGLNGAPAFSPDGNRL--AFVLSRDGNPEIYVMDLGSRALRRLTNNLAIDTEPFWGKD 297

Query: 193 GKQMWLSPFVNSYTDRLKTNINSSVFDQISY 101
           G  ++ +          K N+NS   D++++
Sbjct: 298 GSTLYFTSDRGGKPQIYKMNVNSGAVDRVTF 328




>gi|15611885|ref|NP_223536.1|  (NC_000921) CYSTEINYL-TRNA SYNTHETASE [Helicobacter pylori
J99]
 gi|6831687|sp|Q9ZKW6|SYC_HELPJ  Cysteinyl-tRNA synthetase (Cysteine--tRNA ligase) (CysRS)
 gi|7437718|pir||D71884  cysteine--tRNA ligase (EC 6.1.1.16) - Helicobacter pylori  (strain
J99)
 gi|4155389|gb|AAD06399.1|  (AE001511) CYSTEINYL-TRNA SYNTHETASE [Helicobacter pylori J99]
          Length = 465

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 28/112 (25%), Positives = 53/112 (47%), Gaps = 7/112 (6%)
 Frame = -3


Query: 343 LNGIEI-----FDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQMW 179
           L+G E+     F    +++  K F +  S++EL  + ++  T D  LN +N+    KQ
Sbjct: 55  LSGYEVVLVRNFTDIDDKIINKAFKENKSIQELSSIYIESYTRD--LNALNV----KQPS 108

Query: 178 LSPFVNSYTDRLKTNINSSVFDQISYQVNKLILLFDT--PQQISAISFYNYS 29
           L P  + Y D +   I + +    +Y+V+   +  DT   +   ++S +N S
Sbjct: 109 LEPKASEYLDAMVRMIETLLEKNFAYRVSNGDIYLDTSKDKDYGSLSMHNSS 160



>gi|100458|pir||S25159  proteinase inhibitor I - potato (cultivar Ilam Hardy)
```

```
 gi|21528|emb|CAA47907.1|  (X67675) proteinase inhibitor I [Solanum tuberosum]
 gi|21530|emb|CAA48136.1|  (X67950) protease inhibitor I [Solanum tuberosum]
          Length = 107

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 24/70 (34%), Positives = 36/70 (51%), Gaps = 1/70 (1%)
 Frame = -3

Query: 325 FDPNGEQLYPKLFAKPYSV-KELPQMEMDVRTPDKLLNNVNITLDGKQMWLSPFVNSYTD 149
           F+ NG+Q +P+L   P  + KE+ + E      + L+NNV I L+G  + +    N
Sbjct: 39  FECNGKQFWPELIGVPTKLAKEIIEKE------NSLINNVQILLNGSPVAMDYRCNRV-- 90

Query: 148 RLKTNINSSV 119
           RL  NI  SV
Sbjct: 91  RLFDNILGSV 100



>gi|13123736|gb|AAK12957.1|AF343914_10  (AF343914) unknown [Campylobacter jejuni]
          Length = 272

 Score = 30.4 bits (67), Expect = 3.7
 Identities = 18/48 (37%), Positives = 25/48 (51%), Gaps = 4/48 (8%)
 Frame = -3

Query: 163 NSYTDRLKTNINSSVFDQ----ISYQVNKLILLFDTPQQISAISFYNY 32
           NSY D  +TNINSS F +    + +++N     F+   QI A  F  Y
Sbjct: 175 NSYKDLKQTNINSSYFSRKDINLVHKINFKKTSFEIHNQIRAFIFQEY 222



>gi|15213122|gb|AAK85734.1|  (AY043279) beta-G spectrin [Brugia malayi]
          Length = 2339

 Score = 30.0 bits (66), Expect = 4.9
 Identities = 18/58 (31%), Positives = 28/58 (48%)
 Frame = -3

Query: 337  GIEIFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLDGKQMWLSPFV 164
            G ++FD N +QLY +  +      E  Q +M V    + L   VN+ +  +QM  S  V
Sbjct: 1369 GQKLFDANRQQLYVQSISDMKDWAEQLQQQMTVEDTGQDLTTVNVAMQKQQMIESEMV 1426



>gi|3660690|gb|AAC72317.1|  (AF063428) period [Biston betularia]
          Length = 321

 Score = 30.0 bits (66), Expect = 4.9
 Identities = 20/66 (30%), Positives = 32/66 (48%)
 Frame = +1

Query: 121 HSNLYSFLIYLCMNQRKEITTFAYHLRLCQHYSITYLEF*HPFPFEAVPQLNMVLQTIQD 300
           H ++ +F  Y   N   E+ T      + +H +   LE+  P   E++P L  + Q I+D
Sbjct: 94  HLHIQAFYFYSAFNTANELITKPTSF-MFRHAANGILEYIDP---ESIPYLGYLPQDIKD 149

Query: 301 KVILHL 318
           K  LHL
Sbjct: 150 KSALHL 155
```

```
>gi|14520380|ref|NP_125855.1|  (NC_000868) hypothetical protein [Pyrococcus abyssi]
 gi|7517774|pir||G75204  hypothetical protein PAB0108 - Pyrococcus abyssi (strain Orsay)
 gi|5457595|emb|CAB49086.1|  (AJ248283) hypothetical protein [Pyrococcus abyssi]
          Length = 240

 Score = 30.0 bits (66), Expect = 4.9
 Identities = 21/55 (38%), Positives = 29/55 (52%)
 Frame = -3

Query: 358 LHYIGLNGIEIFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNVNITLD 194
           LHY+ LNG    +   E+L  +L  K   V+E   MEM    P KL+NN  + +D
Sbjct: 173 LHYVLLNG----EATVEELSDRLNMKEREVRE-KIMEMARFVPIKLINNEKVLID 222




>gi|14764011|ref|XP_035625.1|  (XM_035625) KIAA0876 protein [Homo sapiens]
          Length = 1096

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 184 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 213




>gi|3882281|dbj|BAA34500.1|  (AB018323) KIAA0780 protein [Homo sapiens]
          Length = 1100

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 238 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 267




>gi|18044920|gb|AAH20180.1|AAH20180  (BC020180) Similar to gene amplified in squamous cell
carcinoma 1 [Mus musculus]
          Length = 1054

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 185 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 214
```

```
>gi|20348507|ref|XP_110245.1|  (XM_110245) similar to KIAA0677 gene product [Mus musculus]
 gi|13938056|gb|AAH07145.1|AAH07145  (BC007145) Similar to KIAA0677 gene product [Mus
musculus]
          Length = 1086

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 184 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 213



>gi|15611460|ref|NP_223111.1|  (NC_000921) RNA POLYMERASE SIGMA FACTOR (sigma 28)
[Helicobacter pylori J99]
 gi|7451629|pir||F71938  RNA polymerase sigma factor (sigma 28) - Helicobacter pylori
(strain J99)
 gi|4154925|gb|AAD05973.1|  (AE001473) RNA POLYMERASE SIGMA FACTOR (sigma 28) [Helicobacter
pylori J99]
          Length = 255

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 27/92 (29%), Positives = 43/92 (46%), Gaps = 6/92 (6%)
 Frame = -3

Query: 268 KELPQMEMDVRTPDKLLNNVNITLDGKQMWLS----PFVNSYTDRLKTNINSSVF--DQI 107
           KE+ + E   ++ +K+LN  +      Q  L+    P V +   RLK  + SS+   D +
Sbjct: 11  KEIQKTEASEKSIEKVLNAYDKQQHHHQDALAIQYLPAVRAMAFRLKERLPSSIDFNDLV 70

Query: 106 SYQVNKLILLFDTPQQISAISFYNYSKTPVRG 11
           S    +LI L    +     SF+ Y+KT V G
Sbjct: 71  SIGTEELIKLARRYESALNDSFWGYAKTRVNG 102



>gi|14133223|dbj|BAA74899.2|  (AB020683) KIAA0876 protein [Homo sapiens]
          Length = 1119

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 207 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 236



>gi|10567164|dbj|BAB16102.1|  (AB037901) gene amplified in squamous cell carcinoma-1
[Homo sapiens]
          Length = 1056

 Score = 29.6 bits (65), Expect = 6.4
 Identities = 14/30 (46%), Positives = 17/30 (56%)
```

```
 Frame = +1

Query: 178 TTFAYHLRLCQHYSITYLEF*HPFPFEAVP 267
           TTFA+H      YSI YL F  P  + A+P
Sbjct: 185 TTFAWHTEDMDLYSINYLHFGEPKSWYAIP 214




>gi|15606531|ref|NP_213911.1|  (NC_000918) hypothetical protein [Aquifex aeolicus]
 gi|7430337|pir||H70414  conserved hypothetical protein aq_1328 - Aquifex aeolicus
 gi|2983748|gb|AAC07309.1|  (AE000734) hypothetical protein [Aquifex aeolicus]
          Length = 764

 Score = 29.3 bits (64), Expect = 8.3
 Identities = 17/51 (33%), Positives = 29/51 (56%), Gaps = 2/51 (3%)
 Frame = -3

Query: 355 HYIGLNGIE--IFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNNV 209
           H+I L   E  I  P+G+ +Y K+  +P + + +   E+D R P +LLN +
Sbjct: 3   HFIVLEKSEFKINPPSGKLVYKKVSGRPKTEELIKGKEVDERIPYELLNPI 53




>gi|15594935|ref|NP_212724.1|  (NC_001318) dimethyladenosine transferase (ksgA) [Borrelia
burgdorferi]
 gi|7433475|pir||E70173  dimethyladenosine transferase (ksgA) homolog - Lyme disease
spirochete
 gi|2688507|gb|AAB91517.1|  (AE001160) dimethyladenosine transferase (ksgA) [Borrelia
burgdorferi]
          Length = 281

 Score = 29.3 bits (64), Expect = 8.3
 Identities = 26/95 (27%), Positives = 36/95 (37%), Gaps = 14/95 (14%)
 Frame = -3

Query: 334 IEIFDPNGEQLYP---------KLFAKPYSVKELPQMEMDVRTP-----DKLLNNVNITL 197
           I+I D      YP         KL  K  ++K   +    VRT       KL N +   +
Sbjct: 184 IKILDIGENNFYPAPKVKSTTLKLIPKKNNIKNFKEFNKLVRTVFSNRRKKLKNTIINFI 243

Query: 196 DGKQMWLSPFVNSYTDRLKTNINSSVFDQISYQVN 92
             K      F+  Y D+   NI+   F QIS  +N
Sbjct: 244 TNKATLRENFLKEYLDKRPENISVEEFIQISNTLN 278




>gi|16760984|ref|NP_456601.1|  (NC_003198) putative propanol dehydrogenase [Salmonella
enterica subsp. enterica serovar Typhi]
 gi|16503282|emb|CAD02413.1|  (AL627273) putative propanol dehydrogenase [Salmonella
enterica subsp. enterica serovar Typhi]
          Length = 370

 Score = 29.3 bits (64), Expect = 8.3
 Identities = 20/82 (24%), Positives = 33/82 (39%), Gaps = 2/82 (2%)
 Frame = -3

Query: 328 IFDPNGEQLYPKLFAKPYSVKELPQMEMDVRTPDKLLNN--VNITLDGKQMWLSPFVNSY 155
           I DP+    YP      Y    +   E+ V  P ++  N  +++     + W+SP  + +
Sbjct: 133 ISDPDKGIKYPLFNNALYPDMAILDPELVVSVPPQITANTGMDVLTHALEAWVSPHASDF 192
```

```
Query: 154 TDRLKTNINSSVFDQISYQVNK 89
           TD L       VF  +   V K
Sbjct: 193 TDALAEKAAKLVFQYLPTAVEK 214




>gi|169542|gb|AAA72133.1|  (L06985) proteinase inhibitor I [Solanum tuberosum]
          Length = 106

 Score = 29.3 bits (64), Expect = 8.3
 Identities = 17/46 (36%), Positives = 27/46 (57%), Gaps = 1/46 (2%)
 Frame = -3

Query: 325 FDPNGEQLYPKLFAKPYSV-KELPQMEMDVRTPDKLLNNVNITLDG 191
           F+ NG+Q +P+L   P  + KE+ + E      + L+NNV I L+G
Sbjct: 38  FECNGKQFWPELIGVPTKLAKEIIEKE------NSLINNVQILLNG 77




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 202,955,033
Number of Sequences: 919285
Number of extensions: 3997447
Number of successful extensions: 13395
Number of sequences better than 10.0: 62
Number of HSP's better than 10.0 without gapping: 13055
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 13371
length of database: 288,558,979
effective HSP length: 96
effective length of database: 200,307,619
effective search space used: 4807382856
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

RID: 1020956160-028840-16581

7.1.16 Query= **hy-16_12169804**

(357 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
919,285 sequences; 288,558,979 total letters

If you have any problems or questions with the results of this search
please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 336 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|15221158|ref|NP_177555.1|  (NM_106075) hypothetical prote...    62   1e-09
gi|14495224|dbj|BAB60943.1|  (AP003578) contains EST C72594(...    62   1e-09
gi|7489906|pir||S68824  rngB protein, cytosolic - slime mold...    59   1e-08
gi|7243045|dbj|BAA92570.1|  (AB037753) KIAA1332 protein [Hom...    57   5e-08
gi|18543193|ref|NP_569869.1|  (NM_130513) EG:BACR7A4.19 gene...    55   2e-07
gi|7290098|gb|AAF45563.1|  (AE003419) EG:BACR7A4.19 gene pro...    55   2e-07
gi|19075851|ref|NP_588351.1|  (NC_003421) cell polarity prot...    54   2e-07
gi|7489916|pir||S72442  actin-fragmin kinase - slime mold (P...    54   3e-07
gi|15241009|ref|NP_198115.1|  (NM_122645) putative protein [...    54   3e-07
gi|20260514|gb|AAM13155.1|  (AY093156) unknown protein [Arab...    54   3e-07
gi|5032015|ref|NP_005824.1|  (NM_005833) Rab9 effector p40 [...    53   7e-07
gi|12653463|gb|AAH00503.1|AAH00503  (BC000503) Rab9 effector...    53   7e-07
gi|17539228|ref|NP_501279.1|  (NM_068878) C46A5.9.p [Caenorh...    52   9e-07
gi|6321677|ref|NP_011754.1|  (NC_001139) protein containing ...    52   9e-07
gi|7488151|pir||T00913  probable phosphoprotein phosphatase ...    52   1e-06
gi|15218816|ref|NP_171844.1|  (NM_100227) putative serine/th...    52   1e-06
gi|9758913|dbj|BAB09450.1|  (AB012248) contains similarity t...    51   2e-06
gi|15229986|ref|NP_187193.1|  (NM_111415) unknown protein [A...    51   2e-06
gi|18423130|ref|NP_568723.1|  (NM_124410) putative protein [...    51   2e-06
gi|9795595|gb|AAF98413.1|AC026238_5  (AC026238) Hypothetical...    51   3e-06
gi|13173410|gb|AAK14396.1|  (AF339882) attractin [Caenorhabd...    51   3e-06
gi|15221823|ref|NP_173296.1|  (NM_101719) hypothetical prote...    51   3e-06
gi|17567429|ref|NP_510443.1|  (NM_078042) perlecan like glyc...    51   3e-06
gi|18043893|gb|AAH19800.1|AAH19800  (BC019800) Similar to Ra...    50   4e-06
gi|7019405|ref|NP_037452.1|  (NM_013320) host cell factor 2 ...    50   5e-06
```

```
gi|16306876|gb|AAH06558.1|AAH06558  (BC006558) Similar to ho...    50   5e-06
gi|18277872|sp|Q39610|DYHA_CHLRE  Dynein alpha chain, flagel...    50   6e-06
gi|18398038|ref|NP_566316.1|  (NM_111651) expressed protein ...    49   1e-05
gi|15237715|ref|NP_196062.1|  (NM_120524) putative protein [...    49   1e-05
gi|15724206|gb|AAL06496.1|AF412043_1  (AF412043) AT5g04420/T...    49   1e-05
gi|18029285|gb|AAL56463.1|  (AF374376) similar to host cell ...    49   1e-05
gi|9280670|gb|AAF86539.1|AC002560_32  (AC002560) F21B7.7 [Ar...    48   2e-05
gi|6321952|ref|NP_012028.1|  (NC_001140) protein containing ...    48   2e-05
gi|7291384|gb|AAF46812.1|  (AE003456) CG6758 gene product [D...    47   3e-05
gi|15126782|gb|AAH12312.1|AAH12312  (BC012312) Similar to hy...    47   3e-05
gi|6680185|ref|NP_032250.1|  (NM_008224) host cell factor C1...    47   4e-05
gi|1293686|gb|AAB01163.1|  (U53925) transcription factor C1 ...    47   4e-05
gi|7304317|gb|AAF59349.1|  (AE003844) CG1710 gene product [D...    47   4e-05
gi|14970918|emb|CAC44472.1|  (AJ320236) host cell factor [Dr...    47   4e-05
gi|4885403|ref|NP_005325.1|  (NM_005334) host cell factor C1...    47   4e-05
gi|17864168|ref|NP_524621.1|  (NM_079882) Host cell factor [...    47   4e-05
gi|1362760|pir||A56088  host cell factor C1 precursor - human      47   4e-05
gi|14768289|ref|XP_048390.1|  (XM_048390) hypothetical prote...    47   4e-05
gi|1708193|sp|P51610|HFC1_HUMAN  Host cell factor C1 (HCF) (...    47   4e-05
gi|1708194|sp|P51611|HFC1_MESAU  Host cell factor C1 (HCF) (...    47   4e-05
gi|20348947|ref|XP_109547.1|  (XM_109547) similar to transcr...    47   5e-05
gi|20260248|gb|AAM13022.1|  (AY093023) unknown protein [Arab...    47   5e-05
gi|6651060|gb|AAF22156.1|AF133093_11  (AF133093) host cell f...    47   5e-05
gi|7290955|gb|AAF46395.1|  (AE003444) CG12081 gene product [...    46   7e-05
gi|3811109|gb|AAC69437.1|  (U88869) protein serine/threonine...    46   7e-05
gi|12698095|dbj|BAB21874.1|  (AB055250) hypothetical protein...    46   9e-05
gi|19115011|ref|NP_594099.1|  (NC_003424) coiled-coil protei...    46   9e-05
gi|17390341|gb|AAH18154.1|AAH18154  (BC018154) RIKEN cDNA 13...    45   1e-04
gi|13397913|emb|CAC34582.1|  (AJ278127) hypothetical protein...    45   1e-04
gi|17017965|ref|NP_082186.1|  (NM_027910) RIKEN cDNA 1300011...    45   1e-04
gi|14669812|dbj|BAB62016.1|  (AB053465) intercellular mediat...    45   1e-04
gi|15236244|ref|NP_192217.1|  (NM_116542) putative phospho-s...    45   1e-04
gi|6694745|gb|AAF25385.1|AF214574_1  (AF214574) myrosinase-b...    45   1e-04
gi|18401116|ref|NP_566546.1|  (NM_112511) putative lectin [A...    45   1e-04
gi|15228197|ref|NP_188262.1|  (NM_112512) putative lectin [A...    45   2e-04
gi|15225787|ref|NP_180866.1|  (NM_128867) putative myrosinas...    44   3e-04
gi|19262942|emb|CAD24864.1|  (AJ437480) ND2 protein [Paramec...    44   3e-04
gi|15079732|gb|AAH11680.1|AAH11680  (BC011680) Similar to hy...    44   4e-04
gi|12844307|dbj|BAB26317.1|  (AK009481) data source:SPTR, so...    44   4e-04
gi|13542753|gb|AAH05581.1|AAH05581  (BC005581) RIKEN cDNA 23...    44   4e-04
gi|14747228|ref|XP_052684.1|  (XM_052684) host cell factor h...    43   6e-04
gi|16945972|ref|NP_476502.1|  (NM_057161) testis intracellul...    43   6e-04
gi|14495697|gb|AAH09460.1|AAH09460  (BC009460) testis intrac...    43   6e-04
gi|7657301|ref|NP_055130.1|  (NM_014315) host cell factor ho...    43   6e-04
gi|8777408|dbj|BAA96998.1|  (AB023039) contains similarity t...    43   7e-04
gi|18422882|ref|NP_568692.1|  (NM_124193) putative protein [...    43   7e-04
gi|7497490|pir||T29816  hypothetical protein C46A5.9 - Caeno...    42   0.001
gi|12654437|gb|AAH01044.1|AAH01044  (BC001044) Similar to hy...    42   0.001
gi|10436213|dbj|BAB14756.1|  (AK023986) unnamed protein prod...    42   0.001
gi|18606155|gb|AAH22969.1|  (BC022969) hypothetical protein ...    42   0.001
gi|7294558|gb|AAF49898.1|  (AE003541) CG4069 gene product [D...    42   0.002
gi|18401112|ref|NP_566545.1|  (NM_112510) putative lectin [A...    41   0.002
gi|16118838|gb|AAL14622.1|AF416786_1  (AF416786) epithiospec...    41   0.002
gi|18491229|gb|AAL69516.1|  (AY074550) putative jasmonate in...    41   0.002
gi|4587545|gb|AAD25776.1|AC006577_12  (AC006577) Identical t...    41   0.002
gi|421840|pir||S33464  hypothetical protein (clone GBGa476) ...    41   0.002
gi|15221025|ref|NP_175806.1|  (NM_104281) jasmonate inducibl...    41   0.002
gi|18146791|dbj|BAB82454.1|  (AB022764) D-protein [Hordeum v...    40   0.004
gi|15225857|ref|NP_180289.1|  (NM_128279) putative phosphopr...    40   0.006
gi|2497945|sp|Q25386|SCRB_LIMPO  Beta-scruin >gi|1015535|emb...    40   0.006
gi|13160052|emb|CAC32457.1|  (AL132773) dJ741H3.1.2 (attract...    39   0.008
```

```
gi|13160051|emb|CAC32456.1|  (AL132773) dJ741H3.1.1 (attract...   39   0.008
gi|18412640|ref|NP_567268.1|  (NM_116705) coded for by A. th...   39   0.008
gi|6912258|ref|NP_036202.1|  (NM_012070) attractin; attracti...   39   0.008
gi|17064868|gb|AAL32588.1|  (AY062510) Unknown protein [Arab...   39   0.008
gi|8118082|gb|AAF72881.1|AAF72881  (AF218915) membrane attra...   39   0.008
gi|13431311|sp|O75882|ATRN_HUMAN  Attractin precursor (Mahog...   39   0.008
gi|19683003|gb|AAL92627.1|AC115679_6  (AC115679) hypothetica...   39   0.008
gi|17561284|ref|NP_506895.1|  (NM_074494) Human host cell fa...   39   0.008
gi|15227641|ref|NP_181177.1|  (NM_129193) unknown protein [A...   39   0.008
gi|4093196|gb|AAD03057.1|  (AF106861) attractin-2 [Homo sapi...   39   0.008
gi|8118083|gb|AAF72882.1|AAF72882  (AF218915) secreted attra...   39   0.008
gi|15223207|ref|NP_172318.1|  (NM_100715) protein serine/thr...   39   0.011
gi|15238883|ref|NP_197360.1|  (NM_121864) putative protein [...   39   0.011
gi|6664307|gb|AAF22889.1|AC006932_6  (AC006932) T27G7.10 [Ar...   39   0.011
```

Alignments

```
>gi|15221158|ref|NP_177555.1|  (NM_106075) hypothetical protein [Arabidopsis thaliana]
 gi|12323816|gb|AAG51875.1|AC079678_5  (AC079678) hypothetical protein; 26726-23758
[Arabidopsis thaliana]
          Length = 552

 Score = 62.0 bits (149), Expect = 1e-09
 Identities = 31/92 (33%), Positives = 46/92 (49%)
 Frame = +1

Query: 52  DQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK 231
           D     W+ PE  G PP PR  H+ + V +   + +FGG D ++   Y ND+     +
Sbjct: 56  DAETQIWIRPEINGVPPCPRDSHSCTTVGDN--LFVFGGTDGTK---YLNDVHILDTYSH 110

Query: 232 EWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            W++ DI G  PR R +HSA     ++  FGG
Sbjct: 111 TWIRPDIRGEGPRVREAHSAALVDKRLFIFGG 142



 Score = 52.8 bits (125), Expect = 7e-07
 Identities = 29/102 (28%), Positives = 51/102 (49%), Gaps = 3/102 (2%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG---RDDSRSNPYFNDIFAYK 219
           LD   ++ W+ P+  G+ P  R  H+A+ V +   + IFGG     DS    ++ND++
Sbjct: 105 LDTYSHTWIRPDIRGEGPRVREAHSAALVDKR--LFIFGGCGKSSDSDDEVFYNDLYILN 162

Query: 220 IFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGY 345
                W +    G  P  R SH+  A+K++++  GG  ++ Y
Sbjct: 163 TETYMWKRAVTSGKPPSARDSHTCSAWKNKIIVVGGEDLDDY 204



 Score = 51.6 bits (122), Expect = 2e-06
 Identities = 36/109 (33%), Positives = 49/109 (44%)
 Frame = +1

Query: 1   DDKNTQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           D  +  FYN    L  L+     W    T+GKPP  R  HT S  +  N II+ GG D
Sbjct: 148 DSDDEVFYND---LYILNTETYMWKRAVTSGKPPSARDSHTCSAWK--NKIIVVGGED-- 200
```

```
Query: 181 RSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
             + Y +D+      +  W +L   G +   PRA H  VA +   +   FGG
Sbjct: 201 LDDYYLSDVHILDTDKFVWKELKTSGQVLTPRAGHVTVALERNLFVFGG 249




 Score = 40.4 bits (93), Expect = 0.004
 Identities = 24/83 (28%), Positives = 36/83 (42%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
           P  R  HT + ++    + +FGG    R N   N +  +    + W++ +I G  P PR
Sbjct: 19  PGKRWGHTCNAIKGGRFLYVFGGF--GRDNCLTNQVHVFDAETQIWIRPEINGVPPCPRD 76

Query: 280 SHSAVAYKSQVLYFGGVSMNGYL 348
           SHS       +  FGG     YL
Sbjct: 77  SHSCTTVGDNLFVFGGTDGTKYL 99



>gi|14495224|dbj|BAB60943.1|  (AP003578) contains EST C72594(E1889)~similar to Arabidopsis
thaliana chromosome 1, F25I16.5~unknown protein [Oryza sativa (japonica cultivar-group)]
          Length = 624

 Score = 62.0 bits (149), Expect = 1e-09
 Identities = 31/94 (32%), Positives = 47/94 (49%), Gaps = 1/94 (1%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           LD     W  P   G PP PR  HT +     + +++FGG  +   N Y ND+    +
Sbjct: 102 LDLRTKEWTKPPCKGTPPSPRESHTVTACGGCDRLVVFGGSGEGEGN-YLNDVHVLDVAT 160

Query: 229 KEWVQLDIYGN-IPRPRASHSAVAYKSQVLYFGG 327
             W   ++ G+ +P PR SH AVA  S+++ +GG
Sbjct: 161 MTWSSPEVKGDVVPAPRDSHGAVAVGSRLVVYGG 194




 Score = 46.6 bits (109), Expect = 5e-05
 Identities = 32/102 (31%), Positives = 47/102 (45%), Gaps = 2/102 (1%)
 Frame = +1

Query: 28  HFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDI 207
           HF  +L L+     W    TTG  P  R  H A+ V     +++FGG + S+     ND+
Sbjct: 45  HFSDVLTLNLETMAWSSLATTGARPGTRDSHGAALVGHR--MMVFGGTNGSKK---VNDL 99

Query: 208 FAYKIFEKEWVQLDIYGNIPRPRASHSAVAYK--SQVLYFGG 327
               +  KEW +    G  P PR SH+  A     +++ FGG
Sbjct: 100 HVLDLRTKEWTKPPCKGTPPSPRESHTVTACGGCDRLVVFGG 141
```

```
 Score = 45.1 bits (105), Expect = 1e-04
 Identities = 28/93 (30%), Positives = 41/93 (43%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
            LD     W      G  P  R  H A  V     + + GG  D +   Y++D +   +
Sbjct: 207 LDMDAMAWSRFAVKGASPGVRAGHAAVGVGSK--VYVIGGVGDKQ---YYSDAWILDVAN 261

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            + W QL+I G  P+ R SHSAV   + +  +GG
Sbjct: 262 RSWTQLEICGQQPQGRFSHSAVVLNTDIAIYGG 294




 Score = 40.8 bits (94), Expect = 0.003
 Identities = 24/88 (27%), Positives = 39/88 (44%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
            W+ P+  G  P  R  H+A + +   V+ +FGG        +F+D+    +    W  L
Sbjct: 9   WLYPKVVGFNPPERWGHSACFFE--GVVYVFGG---CCGGLHFSDVLTLNLETMAWSSLA 63

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVS 333
              G  P  R SH A     +++ FGG +
Sbjct: 64  TTGARPGTRDSHGAALVGHRMMVFGGTN 91




 Score = 40.4 bits (93), Expect = 0.004
 Identities = 30/100 (30%), Positives = 41/100 (41%), Gaps = 1/100 (1%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKP-PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIF 225
            LD     W  PE  G   P PR  H A  V     ++++GG    R   Y  ++    +
Sbjct: 156 LDVATMTWSSPEVKGDVVPAPRDSHGAVAVGSR--LVVYGGDCGDR---YHGEVDVLDMD 210

Query: 226 EKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGY 345
               W +  + G  P  RA H+AV   S+V   GGV    Y
Sbjct: 211 AMAWSRFAVKGASPGVRAGHAAVGVGSKVYVIGGVGDKQY 250




>gi|7489906|pir||S68824  rngB protein, cytosolic - slime mold (Dictyostelium discoideum)
          Length = 943

 Score = 58.9 bits (141), Expect = 1e-08
 Identities = 33/85 (38%), Positives = 44/85 (50%), Gaps = 9/85 (10%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
            P PR  HT + +   +  I+FGG     SN  FNDI  Y IF   W +++  GN P  R
Sbjct: 14  PEPRWGHTGTTLPNGSGFIVFGGN----SNRAFNDIQYYNIFNNSWSKIEAVGNAPSERY 69

Query: 280 SHSAVAYKS---------QVLYFGG 327
             HSAV Y+S         Q+++FGG
Sbjct: 70  GHSAVLYQSQSRPYSDSYQIIFFGG 94
```

```
 Score = 47.8 bits (112), Expect = 2e-05
 Identities = 31/110 (28%), Positives = 55/110 (49%), Gaps = 1/110 (0%)
 Frame = +1

Query: 1   DDKNTQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           ++  +++YN   ++L     +N W      G  P  R  H+   V   N + IFGG D
Sbjct: 147 NNHKSKYYN---SVLLFSLESNEWRQQVCGGVIPSARATHSTFQVNN-NKMFIFGGYDGK 202

Query: 181 RSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSA-VAYKSQVLYFGG 327
           +   Y+NDI+   +    W +++  G  P+PR+ HSA +   ++++ FGG
Sbjct: 203 K---YYNDIYYLDLETWIWKKVEAKGTPPKPRSGHSATMIQNNKLMIFGG 249




 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 27/96 (28%), Positives = 46/96 (47%), Gaps = 1/96 (1%)
 Frame = +1

Query: 61  NNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWV 240
           N  ++  + T K    R  HTA   ++   +++FGG ++ +S  Y+N +  + +   EW
Sbjct: 112 NRSFIWKQVTTKSIEGRAGHTAVVYRQN--LVVFGGHNNHKSK-YYNSVLLFSLESNEWR 168

Query: 241 QLDIYGNIPRPRASHSAVAYKSQVLY-FGGVSMNGY 345
           Q    G IP  RA+HS     +  ++ FGG     Y
Sbjct: 169 QQVCGGVIPSARATHSTFQVNNNKMFIFGGYDGKKY 204




 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 32/111 (28%), Positives = 46/111 (40%), Gaps = 7/111 (6%)
 Frame = +1

Query: 16  QFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMN-------VIIIFGGRD 174
           Q+YN F         NN W   E  G  P  R  H+A   Q  +        II FGGR
Sbjct: 46  QYYNIF---------NNSWSKIEAVGNAPSERYGHSAVLYQSQSRPYSDSYQIIFFGGR- 95


Query: 175 DSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            + S P+ +    Y    + ++   +       RA H+AV Y+  ++ FGG
Sbjct: 96  -ATSKPFSDINILYVNSNRSFIWKQVTTKSIEGRAGHTAVVYRQNLVVFGG 145



>gi|7243045|dbj|BAA92570.1|  (AB037753) KIAA1332 protein [Homo sapiens]
          Length = 651

 Score = 56.6 bits (135), Expect = 5e-08
 Identities = 27/79 (34%), Positives = 40/79 (50%)
 Frame = +1
```

```
Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPR 270
           G P   R  H+A Y      + +FGG   S  N  FND++   +  KEW++    G+ P
Sbjct: 47  GTPITQRFSHSACYYDANQSMYVFGGCTQSSCNAAFNDLWRLDLNSKEWIRPLASGSYPS 106

Query: 271 PRASHSAVAYKSQVLYFGG 327
           P+A  + V YK  ++ FGG
Sbjct: 107 PKAGATLVVYKDLLVLFGG 125



 Score = 42.7 bits (99), Expect = 7e-04
 Identities = 30/107 (28%), Positives = 51/107 (47%), Gaps = 8/107 (7%)
 Frame = +1

Query: 31  FMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPY----- 195
           F  L +LD  +  W+ P  +G  P P+    A+ V   +++++FGG   +R +PY
Sbjct: 82  FNDLWRLDLNSKEWIRPLASGSYPSPKAG--ATLVVYKDLLVLFGGW--TRPSPYPLHQP 137

Query: 196 ---FNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
              F++I  Y    K W    +  + P P A HS+     +++ FGG
Sbjct: 138 ERFFDEIHTYSP-SKNWWNCIVTTHGPPPMAGHSSCVIDDKMIVFGG 183



 Score = 35.0 bits (79), Expect = 0.15
 Identities = 26/89 (29%), Positives = 41/89 (45%), Gaps = 1/89 (1%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W    TT  PP P   H++  + +   +I+FGG   SR     ND++   + +  W +
Sbjct: 152 NWWNCIVTTHGPP-PMAGHSSCVIDDK--MIVFGGSLGSRQMS--NDVWVLDLEQWAWSK 206

Query: 244 LDIYGNIPRPRASHSAVAY-KSQVLYFGG 327
            +I G  P PR   S +    + +L  GG
Sbjct: 207 PNISGPSPHPRGGQSQIVIDDATILILGG 235



>gi|18543193|ref|NP_569869.1| ☐ (NM_130513) EG:BACR7A4.19 gene product [Drosophila
melanogaster]
 gi|17862766|gb|AAL39860.1| ☐ (AY069715) LP01394p [Drosophila melanogaster]
          Length = 786

 Score = 54.7 bits (130), Expect = 2e-07
 Identities = 35/107 (32%), Positives = 49/107 (45%), Gaps = 4/107 (3%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-RDDSRSNPYF---NDI 207
           L++    +  W     TG PP PR  H+A  V   + + IFGG   D  SN      ND+
Sbjct: 279 LIRFGVKDKSWGRACATGTPPAPRYHHSA--VVAGSSMFIFGGYTGDIHSNSNLTNKNDL 336

Query: 208 FAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYL 348
           F YK     WV+    G  P PR++H A  Y +++  + G   N  L
Sbjct: 337 FEYKFQSAMWVEWKFSGRQPVPRSAHGAAVYDNKMWIYAGYDGNARL 383
```

```
 Score = 39.7 bits (91), Expect = 0.006
 Identities = 25/73 (34%), Positives = 36/73 (49%)
 Frame = +1

Query: 109 RIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHS 288
           R +HT   V   + + +FGG  D+  N   ND+  + + +K W +    G  P PR  HS
Sbjct: 252 RSKHTV--VAYKDAMFVFGG--DNGKN-MLNDLIRFGVKDKSWGRACATGTPPAPRYHHS 306

Query: 289 AVAYKSQVLYFGG 327
           AV   S +  FGG
Sbjct: 307 AVVAGSSMFIFGG 319



>gi|7290098|gb|AAF45563.1|  (AE003419) EG:BACR7A4.19 gene product [Drosophila
melanogaster]
 gi|10190798|emb|CAB65872.1|  (AL109630) /prediction=(method:''genefinder'',
version:''084'')~/match=(desc:''LZTR-1'', species:''Homo sapiens (Human)'',
ranges:(query:103239..103385, target:SPTREMBL::Q14776:49..1, score:''172.00''),
(query:102846..103262, target:SPTREMBL::Q14776:23>
          Length = 975

 Score = 54.7 bits (130), Expect = 2e-07
 Identities = 35/107 (32%), Positives = 49/107 (45%), Gaps = 4/107 (3%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-RDDSRSNPYF---NDI 207
           L++    +  W     TG PP PR  H+A  V   + + IFGG   D  SN      ND+
Sbjct: 279 LIRFGVKDKSWGRACATGTPPAPRYHHSA--VVAGSSMFIFGGYTGDIHSNSNLTNKNDL 336

Query: 208 FAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYL 348
           F YK     WV+    G  P PR++H A  Y +++  + G   N  L
Sbjct: 337 FEYKFQSAMWVEWKFSGRQPVPRSAHGAAVYDNKMWIYAGYDGNARL 383
```

```
 Score = 39.7 bits (91), Expect = 0.006
 Identities = 25/73 (34%), Positives = 36/73 (49%)
 Frame = +1

Query: 109 RIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHS 288
           R +HT   V   + + +FGG  D+  N   ND+  + + +K W +    G  P PR  HS
Sbjct: 252 RSKHTV--VAYKDAMFVFGG--DNGKN-MLNDLIRFGVKDKSWGRACATGTPPAPRYHHS 306

Query: 289 AVAYKSQVLYFGG 327
           AV   S +  FGG
Sbjct: 307 AVVAGSSMFIFGG 319
```

```
>gi|19075851|ref|NP_588351.1|  (NC_003421) cell polarity protein tea1.tip elongation
aberrant protein 1 [Schizosaccharomyces pombe]
 gi|6094446|sp|P87061|TEA1_SCHPO  Tip elongation aberrant protein 1 (Cell polarity protein
tea1)
```

```
 gi|7490217|pir||T40866  cell polarity protein tea1p - fission yeast  (Schizosaccharomyces
pombe)
 gi|2065436|emb|CAA73246.1|  (Y12709) tea1p [Schizosaccharomyces pombe]
 gi|3618212|emb|CAA20875.1|  (AL031579) cell polarity protein tea1.tip elongation aberrant
protein 1 [Schizosaccharomyces pombe]
          Length = 1147

 Score = 54.3 bits (129), Expect = 2e-07
 Identities = 32/106 (30%), Positives = 49/106 (46%)
 Frame = +1

Query: 25  NHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFND 204
           N F  L       + W   ET G  P PR  H AS V+   ++ +FGGR  +    + ND
Sbjct: 265 NFFNDLWCYHPKQSAWSKVETFGVAPNPRAGHAASVVE--GILYVFGGR--ASDGTFLND 320

Query: 205 IFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
           ++A+++  K W +L      P PR+SH+       ++  GG    G
Sbjct: 321 LYAFRLSSKHWYKLSDLPFTPSPRSSHTLSCSGLTLVLIGGKQGKG 366




 Score = 53.9 bits (128), Expect = 3e-07
 Identities = 29/101 (28%), Positives = 49/101 (47%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  L+  ++ W +      PP  R  H A    +   + IFGG D +    +FND++ Y
Sbjct: 220 LNNLNTSDSRWELASVVNDPPPARAGHVAFTFSDK--LYIFGGTDGAN---FFNDLWCYH 274

Query: 220 IFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
             +  W +++ +G  P PRA H+A   +  +  FGG + +G
Sbjct: 275 PKQSAWSKVETFGVAPNPRAGHAASVVEGILYVFGGRASDG 315




 Score = 38.1 bits (87), Expect = 0.018
 Identities = 23/79 (29%), Positives = 38/79 (47%), Gaps = 2/79 (2%)
 Frame = +1

Query: 103 LPRIQHTAS-YVQEMNVIIIFGG-RDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPR 276
           LPR  H +  Y +    I IFGG   DS+     ND++    +   ++  L   G  P PR
Sbjct: 78  LPRYSHASHLYAEGGQEIYIFGGVASDSQPK---NDLWVLNLATSQFTSLRSLGETPSPR 134

Query: 277 ASHSAVAYKSQVLYFGGVS 333
             H+++   +  + FGG++
Sbjct: 135 LGHASILIGNAFIVFGGLT 153




>gi|7489916|pir||S72442  actin-fragmin kinase - slime mold (Physarum polycephalum)
 gi|1553133|gb|AAB08728.1|  (U64722) actin-fragmin kinase [Physarum polycephalum]
          Length = 737

 Score = 53.9 bits (128), Expect = 3e-07
 Identities = 34/108 (31%), Positives = 48/108 (43%), Gaps = 3/108 (2%)
 Frame = +1
```

```
Query: 40  LLKLDQPNNPW--VIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFA 213
           L   D   + W  + P   G  P PR  H+A+ V     + IFGG +    + Y ND++
Sbjct: 498 LYTFDFAKSTWEKIEPTKDGPWPSPRYHHSATLVGAS--LYIFGGAE--HKSKYHNDVYV 553

Query: 214 YKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG-VSMNGYLEF 354
           YK    +W  L+  G  P PRA    V + + +  FGG     GY  F
Sbjct: 554 YKFDANQWELLNATGETPEPRAGQMTVEWNNSLFTFGGHGGEGGYTSF 601




 Score = 42.0 bits (97), Expect = 0.001
 Identities = 24/90 (26%), Positives = 40/90 (43%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W+      + P PR  H+    +    + +FGG     ++   ND + +   +K+W  +
Sbjct: 407 WLKVVVPNEKPAPRRYHSGVLYE--GKLYVFGGVCIKTAS---NDFYVFDFAKKKWSIVV 461

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVSMN 339
             G  P PR  HSA  Y  ++  FGG + N
Sbjct: 462 AQGEAPSPRCGHSATVYGGKMWIFGGHNNN 491




 Score = 40.4 bits (93), Expect = 0.004
 Identities = 27/94 (28%), Positives = 45/94 (47%), Gaps = 2/94 (2%)
 Frame = +1

Query: 52  DQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK 231
           D     W I    G+ P PR  H+A+       + IFGG ++++  PY +D++ +   +
Sbjct: 451 DFAKKKWSIVVAQGEAPSPRCGHSATVYG--GKMWIFGGHNNNK-QPY-SDLYTFDFAKS 506

Query: 232 EWVQLDIY--GNIPRPRASHSAVAYKSQVLYFGG 327
            W +++    G  P PR  HSA    + +  FGG
Sbjct: 507 TWEKIEPTKDGPWPSPRYHHSATLVGASLYIFGG 540




 Score = 35.0 bits (79), Expect = 0.15
 Identities = 29/110 (26%), Positives = 43/110 (38%), Gaps = 8/110 (7%)
 Frame = +1

Query: 22  YNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFN 201
           Y++ + + K D   N W +   TG+ P PR       V+  N +  FGG         F
Sbjct: 547 YHNDVYVYKFDA--NQWELLNATGETPEPRAGQMT--VEWNNSLFTFGGHGGEGGYTSFV 602

Query: 202 DIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAY--------KSQVLYFGG 327
           D   ++I    + ++D  G  PR     S V Y        +  V  FGG
Sbjct: 603 DAHVFEIATNTFHEVDCSGTFPRTARPLSYVPYYYGSGDKREGAVFSFGG 652




 Score = 31.2 bits (69), Expect = 2.2
```

```
 Identities = 13/35 (37%), Positives = 20/35 (57%)
 Frame = +1

Query: 226 EKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGV 330
           E  W+++ +    P PR  HS V Y+ ++  FGGV
Sbjct: 404 ECAWLKVVVPNEKPAPRRYHSGVLYEGKLYVFGGV 438



>gi|15241009|ref|NP_198115.1|  (NM_122645) putative protein [Arabidopsis thaliana]
          Length = 680

 Score = 53.9 bits (128), Expect = 3e-07
 Identities = 31/89 (34%), Positives = 49/89 (54%), Gaps = 1/89 (1%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W I +T GKPP+ R   + + V +   ++IFGG+D  RS    ND+    +    W ++D
Sbjct: 295 WSILKTYGKPPISRGGQSVTLVGKS--LVIFGGQDAKRS--LLNDLHILDLDTMTWEEID 350

Query: 250 IYGNIPRPRASHSAVAYKSQ-VLYFGGVS 333
             G+ P PR+ H+A  +  + +L FGG S
Sbjct: 351 AVGSPPTPRSDHAAAVHAERYLLIFGGGS 379




 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 28/93 (30%), Positives = 39/93 (41%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           LD     W   +  G PP PR  H A+ V     ++IFGG   +     F+D+    +
Sbjct: 339 LDLDTMTWEEIDAVGSPPTPRSDHAAA-VHAERYLLIFGGGSHATC---FDDLHVLDLQT 394

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            EW +    G+ P PRA H+ V         GG
Sbjct: 395 MEWSRHTQQGDAPTPRAGHAGVTIGENWYIVGG 427



 Score = 41.6 bits (96), Expect = 0.002
 Identities = 26/98 (26%), Positives = 46/98 (46%), Gaps = 10/98 (10%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W  P T+G+PP  R QH A+ +Q+   + ++GG  + R   Y  D+    +    W +
Sbjct: 182 NQWTAPRTSGQPPKARYQHGAAVIQDK--MYMYGGNHNGR---YLGDLHVLDLKNWTWSR 236

Query: 244 LD---IYGNIPRPR-------ASHSAVAYKSQVLYFGG 327
           ++   + G+             A HS + + +Q+L  GG
Sbjct: 237 VETKVVTGSQETSSPAKLTHCAGHSLIPWDNQLLSIGG 274



>gi|20260514|gb|AAM13155.1|  (AY093156) unknown protein [Arabidopsis thaliana]
          Length = 648
```

```
 Score = 53.9 bits (128), Expect = 3e-07
 Identities = 31/89 (34%), Positives = 49/89 (54%), Gaps = 1/89 (1%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W I +T GKPP+ R   + + V +   ++IFGG+D  RS    ND+    +    W ++D
Sbjct: 283 WSILKTYGKPPISRGGQSVTLVGKS--LVIFGGQDAKRS--LLNDLHILDLDTMTWEEID 338

Query: 250 IYGNIPRPRASHSAVAYKSQ-VLYFGGVS 333
             G+ P PR+ H+A  +  + +L FGG S
Sbjct: 339 AVGSPPTPRSDHAAAVHAERYLLIFGGGS 367




 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 28/93 (30%), Positives = 39/93 (41%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           LD     W   +  G PP PR  H A+ V     ++IFGG   +     F+D+    +
Sbjct: 327 LDLDTMTWEEIDAVGSPPTPRSDHAAA-VHAERYLLIFGGGSHATC---FDDLHVLDLQT 382

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            EW +    G+ P PRA H+ V         GG
Sbjct: 383 MEWSRHTQQGDAPTPRAGHAGVTIGENWYIVGG 415




 Score = 41.6 bits (96), Expect = 0.002
 Identities = 26/98 (26%), Positives = 46/98 (46%), Gaps = 10/98 (10%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W  P T+G+PP  R QH A+ +Q+   + ++GG  + R   Y  D+    +    W +
Sbjct: 170 NQWTAPRTSGQPPKARYQHGAAVIQDK--MYMYGGNHNGR---YLGDLHVLDLKNWTWSR 224

Query: 244 LD---IYGNIPRPR-------ASHSAVAYKSQVLYFGG 327
           ++   + G+             A HS + + +Q+L  GG
Sbjct: 225 VETKVVTGSQETSSPAKLTHCAGHSLIPWDNQLLSIGG 262




>gi|5032015|ref|NP_005824.1|  (NM_005833) Rab9 effector p40 [Homo sapiens]
 gi|2217970|emb|CAB09808.1|  (Z97074) p40 [Homo sapiens]
          Length = 372

 Score = 52.8 bits (125), Expect = 7e-07
 Identities = 33/94 (35%), Positives = 46/94 (48%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W  PET G PP PR  H    V     + I GG    R   +++D+    I + +W +L+
Sbjct: 180 WSQPETLGNPPSPRHGHVM--VAAGTKLFIHGGLAGDR---FYDDLHCIDISDMKWQKLN 234
```

```
Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYLE 351
             G  P   A+HSAVA    V  FGG++  G L+
Sbjct: 235 PTGAAPAGCAAHSAVAMGKHVYIFGGMTPAGALD 268




 Score = 49.3 bits (116), Expect = 8e-06
 Identities = 28/95 (29%), Positives = 41/95 (42%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           L+     W  PE T  PP PR  HT+S     N + +FGG +          +  +
Sbjct: 119 LNPETRTWTTPEVTSPPPSPRTFHTSS-AAIGNQLYVFGGGERGAQPVQDTKLHVFDANT 177

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
             W Q +  GN P PR   H  VA  +++   GG++
Sbjct: 178 LTWSQPETLGNPPSPRHGHVMVAAGTKLFIHGGLA 212




 Score = 29.3 bits (64), Expect = 8.4
 Identities = 25/83 (30%), Positives = 38/83 (45%), Gaps = 3/83 (3%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNV--IIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGN 261
           T K  LPR +H AS++       I +FGG + S +    N +       + W   ++
Sbjct: 79  TCKGLLPRYEH-ASFIPSCTPDRIWVFGGANQSGNR---NCLQVLNPETRTWTTPEVTSP 134

Query: 262 IPRPRASH-SAVAYKSQVLYFGG 327
            P PR  H S+ A  +Q+  FGG
Sbjct: 135 PPSPRTFHTSSAAIGNQLYVFGG 157
```

>gi|12653463|gb|AAH00503.1|AAH00503 ☐ (BC000503) Rab9 effector p40 [Homo sapiens]

```
          Length = 372

 Score = 52.8 bits (125), Expect = 7e-07
 Identities = 33/94 (35%), Positives = 46/94 (48%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W  PET G PP PR  H    V     + I GG    R   +++D+    I + +W +L+
Sbjct: 180 WSQPETLGNPPSPRHGHVM--VAAGTKLFIHGGLAGDR---FYDDLHCIDISDMKWQKLN 234

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYLE 351
             G  P   A+HSAVA    V  FGG++  G L+
Sbjct: 235 PTGAAPAGCAAHSAVAMGKHVYIFGGMTPAGALD 268




 Score = 49.3 bits (116), Expect = 8e-06
 Identities = 28/95 (29%), Positives = 41/95 (42%)
 Frame = +1
```

```
Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           L+     W  PE T  PP PR  HT+S     N + +FGG +          +  +
Sbjct: 119 LNPETRTWTTPEVTSPPPSPRTFHTSS-AAIGNQLYVFGGGERGAQPVQDTKLHVFDANT 177

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
             W Q +  GN P PR  H  VA  +++   GG++
Sbjct: 178 LTWSQPETLGNPPSPRHGHVMVAAGTKLFIHGGLA 212



 Score = 29.3 bits (64), Expect = 8.4
 Identities = 25/83 (30%), Positives = 38/83 (45%), Gaps = 3/83 (3%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNV--IIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGN 261
           T K  LPR +H AS++       I +FGG + S +    N +       + W   ++
Sbjct: 79  TCKGLLPRYEH-ASFIPSCTPDRIWVFGGANQSGNR---NCLQVLNPETRTWTTPEVTSP 134

Query: 262 IPRPRASH-SAVAYKSQVLYFGG 327
            P PR  H S+ A  +Q+  FGG
Sbjct: 135 PPSPRTFHTSSAAIGNQLYVFGG 157



>gi|17539228|ref|NP_501279.1|  (NM_068878) C46A5.9.p [Caenorhabditis elegans]
 gi|11359757|pir||T43277  host cell factor 1 - Caenorhabditis elegans
 gi|3851534|gb|AAD12580.1|  (AF072907) host cell factor 1 [Caenorhabditis elegans]
 gi|14550352|gb|AAB03147.2|  (U61948) Hypothetical protein C46A5.9 [Caenorhabditis elegans]
          Length = 782

 Score = 52.4 bits (124), Expect = 9e-07
 Identities = 26/90 (28%), Positives = 51/90 (55%)
 Frame = +1

Query: 73  VIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDI 252
           ++ ++TG  P PR  H A  ++E+  I+IFGG ++       +++ AY   ++EW
Sbjct: 32  IVQQSTGPNPKPRHGHRAVVLKEL--IVIFGGGNEGM----IDELHAYNTQKREWTAPQC 85

Query: 253 YGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
            G++P P A+  A++  +++  FGG++  G
Sbjct: 86  CGDVPTPAAAFGAISLGNKIYRFGGMTEYG 115



 Score = 42.0 bits (97), Expect = 0.001
 Identities = 32/106 (30%), Positives = 52/106 (48%), Gaps = 13/106 (12%)
 Frame = +1


Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIF 225
           NP V   + G  P PRI H+    Q+     +FGG  +  ++P      Y +D++   +
Sbjct: 135 NPRV--HSNGHLPCPRIGHSFVVSQKSQKAYVFGGLSNDLNDPKRNVPHYLDDLYVINLS 192

Query: 226 EKE---WVQLDIYGNIPRPRASHSAVAYK----SQVLYFGGVSMNG 342
             +   W +L++  G  P  R SH+AV Y+    S+++ +GG  MNG
Sbjct: 193 GPQHLIWEKLNATGPGPISRESHTAVIYEKDSISRMVVYGG--MNG 236
```

```
 Score = 36.6 bits (83), Expect = 0.052
 Identities = 26/114 (22%), Positives = 51/114 (43%), Gaps = 7/114 (6%)
 Frame = +1

Query: 7   KNTQFYNHFMALLKLDQPNNP-WVIPETTGKPPLPRIQHTASYVQEMNV--IIIFGGRDD 177
           +N   Y   + ++ L  P +  W     TG  P+ R  HTA   ++ ++  ++++GG +
Sbjct: 177 RNVPHYLDDLYVINLSGPQHLIWEKLNATGPGPISRESHTAVIYEKDSISRMVVYGGMNG 236

Query: 178 SRSNPYFNDIFAYKIFEKEWVQLDI----YGNIPRPRASHSAVAYKSQVLYFGG 327
            R      D++   +    W ++       G  P PR+ HS+V    ++  +GG
Sbjct: 237 VR----LGDLWYLNLNTLHWTEIKFDDPRTGIPPMPRSLHSSVLIGDKMFVYGG 286



 Score = 35.0 bits (79), Expect = 0.15
 Identities = 27/110 (24%), Positives = 45/110 (40%), Gaps = 15/110 (13%)
 Frame = +1

Query: 43  LKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----------RDDSRSN 189
           +K D P         TG PP+PR  H++  + +   + ++GG           +   +
Sbjct: 255 IKFDDPR--------TGIPPMPRSLHSSVLIGDK--MFVYGGWVPLLEHASTEQQTEKEW 304

Query: 190 PYFNDIFAYKIFEKEWVQLDIY----GNIPRPRASHSAVAYKSQVLYFGG 327
              + +  + I E  WV L +Y      IPR RA H A A   ++  + G
Sbjct: 305 KCTSSLGCWNITEDRWVPLHLYCSDEDTIPRGRAGHCAAAVGDRMYIWSG 354



>gi|6321677|ref|NP_011754.1|  (NC_001139) protein containing kelch repeats, similar to
YHR158c and YPL263c; Kel2p [Saccharomyces cerevisiae]
 gi|1723756|sp|P50090|KEL2_YEAST  Kelch repeats protein 2
 gi|1362338|pir||S57704  hypothetical protein YGR238c - yeast (Saccharomyces cerevisiae)
 gi|886933|emb|CAA61189.1|  (X87941) ORF 882 [Saccharomyces cerevisiae]
 gi|1323431|emb|CAA97266.1|  (Z73023) ORF YGR238c [Saccharomyces cerevisiae]
          Length = 882

 Score = 52.4 bits (124), Expect = 9e-07
 Identities = 29/94 (30%), Positives = 46/94 (48%), Gaps = 8/94 (8%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQE---MNVIIIFGGRDDSRSNPYFNDIFAYKIFE---- 228
           W IP+  G+ PL R  H  S +        + +FGG+ D     YFND+  + +
Sbjct: 184 WTIPQPIGRRPLGRYGHKISIIASNPMQTKLYLFGGQVDET---YFNDLVVFDLSSFRRP 240

Query: 229 -KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
               W  L+  G++P P  +H+ VAY +++  FGG
Sbjct: 241 NSHWEFLEPVGDLPPPLTNHTMVAYDNKLWVFGG 274



 Score = 46.6 bits (109), Expect = 5e-05
 Identities = 28/91 (30%), Positives = 40/91 (43%)
 Frame = +1
```

```
Query: 55  QPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKE 234
           +PN+ W   E  G  P P   HT   V   N + +FGG      +   ND + Y   + E
Sbjct: 239 RPNSHWEFLEPVGDLPPPLTNHTM--VAYDNKLWVFGGETPKTIS---NDTYRYDPAQSE 293

Query: 235 WVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
           W ++   G  P P   H++V YK  +   GG
Sbjct: 294 WSKVKTTGEKPPPIQEHASVVYKHLMCVLGG 324




 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 30/101 (29%), Positives = 50/101 (48%), Gaps = 1/101 (0%)
 Frame = +1

Query: 46  KLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIF 225
           + D   + W   +TTG+ P P IQ  AS V + +++ + GG+D    N Y ND++   +
Sbjct: 286 RYDPAQSEWSKVKTTGEKP-PPIQEHASVVYK-HLMCVLGGKDTH--NAYSNDVYFLNLL 341

Query: 226 EKEWVQLD-IYGNIPRPRASHSAVAYKSQVLYFGGVSMNGY 345
             +W +L  +   IP+ R+ HS    K++ L   G     Y
Sbjct: 342 SLKWYKLPRMKEGIPQERSGHSLTLMKNEKLLIMGGDKTDY 382




 Score = 32.0 bits (71), Expect = 1.3
 Identities = 26/95 (27%), Positives = 44/95 (45%), Gaps = 11/95 (11%)
 Frame = +1

Query: 100 PLPRIQHTASY-VQEMNVIIIFGGRDDSRSNPYFNDIFAYKI------FEKEWVQLDIYG 258
           P PR +H++S+ V   N I + GG  D      + D++          F  +  ++DI
Sbjct: 82  PFPRYRHSSSFIVTNDNRIFVTGGLHDQS---VYGDVWQIAANADGTSFTSK--RIDIDQ 136

Query: 259 NIPRPRASHSAVAYKSQVLYFGG----VSMNGYLE 351
           N P PR  H++    +   + FGG    ++ NG L+
Sbjct: 137 NTPPPRVGHASTICGNAYVVFGGDTHKLNKNGLLD 171
```

>gi|7488151|pir||T00913 probable phosphoprotein phosphatase (EC 3.1.3.16) F21B7.27 - Arabidopsis thaliana

```
          Length = 729

 Score = 52.0 bits (123), Expect = 1e-06
 Identities = 30/105 (28%), Positives = 53/105 (49%), Gaps = 13/105 (12%)
 Frame = +1


Query: 67  PWVIPETTGKPPLPRIQHT--ASYVQEMNVIIIFGGRDDSRSNP-----------YFNDI 207
           P+ IP+     P PR  HT  A +V   + +I+FGG   + +N              N +
Sbjct: 69  PFTIPDEDW--PGPRCGHTLTAVFVNNSHQLILFGGSTTAVANHNSSLPEISLDGVTNSV 126

Query: 208 FAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
            ++ +  ++W +L+  G++P PRA H+A  Y + +L  GG+  +G
Sbjct: 127 HSFDVLTRKWTRLNPIGDVPSPRACHAAALYGTLILIQGGIGPSG 171
```

```
 Score = 33.5 bits (75), Expect = 0.44
 Identities = 26/95 (27%), Positives = 41/95 (42%), Gaps = 1/95 (1%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W      G  P PR  H A+    +  I+I GG   S   P   D++   +   +W++
Sbjct: 136 WTRLNPIGDVPSPRACHAAALYGTL--ILIQGGIGPS--GPSDGDVYMLDMTNNKWIKFL 191

Query: 250 IYGNIPRPRASH-SAVAYKSQVLYFGGVSMNGYLE 351
           + G  P PR  H   +A +  ++ F G   NG L+
Sbjct: 192 VGGETPSPRYGHVMDIAAQRWLVIFSG--NNGMLQ 224




 Score = 33.1 bits (74), Expect = 0.58
 Identities = 25/98 (25%), Positives = 40/98 (40%), Gaps = 4/98 (4%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS----RSNPYFNDIFAY 216
           LD  NN W+     G+ P PR  H    +     ++IF G +             D +
Sbjct: 180 LDMTNNKWIKFLVGGETPSPRYGHVMD-IAAQRWLVIFSGNNGMLQVVLEKMTLGDTYGL 238

Query: 217 KIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGV 330
           K+ + + V   +    P PR  H+AV   S++   GG+
Sbjct: 239 KM-DSDNVWTPVPAVAPSPRYQHTAVFGGSKLHVIGGI 275



>gi|15218816|ref|NP_171844.1|  (NM_100227) putative serine/threonine protein phosphatase
[Arabidopsis thaliana]
          Length = 797

 Score = 52.0 bits (123), Expect = 1e-06
 Identities = 30/105 (28%), Positives = 53/105 (49%), Gaps = 13/105 (12%)
 Frame = +1

Query: 67  PWVIPETTGKPPLPRIQHT--ASYVQEMNVIIIFGGRDDSRSNP-----------YFNDI 207
           P+ IP+     P PR  HT  A +V   + +I+FGG   + +N              N +
Sbjct: 69  PFTIPDEDW--PGPRCGHTLTAVFVNNSHQLILFGGSTTAVANHNSSLPEISLDGVTNSV 126

Query: 208 FAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
            ++ +  ++W +L+  G++P PRA H+A  Y + +L  GG+  +G
Sbjct: 127 HSFDVLTRKWTRLNPIGDVPSPRACHAAALYGTLILIQGGIGPSG 171




 Score = 33.5 bits (75), Expect = 0.44
 Identities = 26/95 (27%), Positives = 41/95 (42%), Gaps = 1/95 (1%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W      G  P PR  H A+    +  I+I GG   S   P   D++   +   +W++
Sbjct: 136 WTRLNPIGDVPSPRACHAAALYGTL--ILIQGGIGPS--GPSDGDVYMLDMTNNKWIKFL 191
```

```
Query: 250 IYGNIPRPRASH-SAVAYKSQVLYFGGVSMNGYLE 351
           + G  P PR  H   +A +  ++ F G   NG L+
Sbjct: 192 VGGETPSPRYGHVMDIAAQRWLVIFSG--NNGMLQ 224




 Score = 33.1 bits (74), Expect = 0.58
 Identities = 25/98 (25%), Positives = 40/98 (40%), Gaps = 4/98 (4%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS----RSNPYFNDIFAY 216
           LD  NN W+     G+ P PR  H    +     ++IF G +             D +
Sbjct: 180 LDMTNNKWIKFLVGGETPSPRYGHVMD-IAAQRWLVIFSGNNGMLQVVLEKMTLGDTYGL 238

Query: 217 KIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGV 330
           K+ + + V   +    P PR  H+AV   S++   GG+
Sbjct: 239 KM-DSDNVWTPVPAVAPSPRYQHTAVFGGSKLHVIGGI 275



>gi|9758913|dbj|BAB09450.1|  (AB012248) contains similarity to unknown
protein~gb|AAF49898.1~gene_id:MXI22.1 [Arabidopsis thaliana]
          Length = 596

 Score = 51.2 bits (121), Expect = 2e-06
 Identities = 35/112 (31%), Positives = 52/112 (46%), Gaps = 8/112 (7%)
 Frame = +1

Query: 16  QFYNH-----FMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           +FYN      +  L + D     W +  +   PP PR  H A  V   N + IFGG   S
Sbjct: 90  EFYNGQKTYVYGDLYRYDVEKQEWKLVSSPNSPP-PRSSHQA--VAWKNYLYIFGGEFTS 146

Query: 181 RSNPYFN---DIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +   F+   D +   +   +W QL++ G  P PR+ H  V YK +++ FGG
Sbjct: 147 PNQERFHHYKDFWMLDVKTNQWEQLNLKG-CPSPRSGHRMVLYKHKIIIFGG 197




 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 22/60 (36%), Positives = 37/60 (61%), Gaps = 1/60 (1%)
 Frame = +1

Query: 151 IIIFGGRDDSRSNPY-FNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
           +I++GG   +    Y + D++ Y + ++EW +L    N P PR+SH AVA+K+ +  FGG
Sbjct: 84  LILYGGEFYNGQKTYVYGDLYRYDVEKQEW-KLVSSPNSPPPRSSHQAVAWKNYLYIFGG 142



>gi|15229986|ref|NP_187193.1|  (NM_111415) unknown protein [Arabidopsis thaliana]
 gi|7596769|gb|AAF64540.1|  (AC009606) unknown protein [Arabidopsis thaliana]
          Length = 668

 Score = 51.2 bits (121), Expect = 2e-06
 Identities = 30/89 (33%), Positives = 48/89 (53%), Gaps = 1/89 (1%)
```

```
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W + +T GKPP+ R   + + V +   ++IFGG+D  RS    ND+    +    W ++D
Sbjct: 282 WSMLKTYGKPPVSRGGQSVTMVGK--TLVIFGGQDAKRS--LLNDLHILDLDTMTWDEID 337

Query: 250 IYGNIPRPRASHSAVAYKSQ-VLYFGGVS 333
             G  P PR+ H+A  +  + +L FGG S
Sbjct: 338 AVGVSPSPRSDHAAAVHAERFLLIFGGGS 366




 Score = 42.0 bits (97), Expect = 0.001
 Identities = 25/98 (25%), Positives = 46/98 (46%), Gaps = 10/98 (10%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W  P+T+G+ P  R +H A+ +Q+   + I+GG  + R   Y  D+    +    W +
Sbjct: 169 NQWTAPQTSGQRPKARYEHGAAVIQDK--MYIYGGNHNGR---YLGDLHVLDLKSWTWSR 223

Query: 244 LDI----------YGNIPRPRASHSAVAYKSQVLYFGG 327
           ++              +  P A HS +A+ +++L  GG
Sbjct: 224 VETKVATESQETSTPTLLAPCAGHSLIAWDNKLLSIGG 261




 Score = 41.2 bits (95), Expect = 0.002
 Identities = 27/93 (29%), Positives = 38/93 (40%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           LD     W   +  G  P PR  H A+ V     ++IFGG   +     F+D+    +
Sbjct: 326 LDLDTMTWDEIDAVGVSPSPRSDHAAA-VHAERFLLIFGGGSHATC---FDDLHVLDLQT 381

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            EW +    G+ P PRA H+ V         GG
Sbjct: 382 MEWSRPAQQGDAPTPRAGHAGVTIGENWFIVGG 414




>gi|18423130|ref|NP_568723.1|  (NM_124410) putative protein [Arabidopsis thaliana]
 gi|15724228|gb|AAL06507.1|AF412054_1  (AF412054) AT5g50310/MXI22_1 [Arabidopsis thaliana]
          Length = 666

 Score = 51.2 bits (121), Expect = 2e-06
 Identities = 35/112 (31%), Positives = 52/112 (46%), Gaps = 8/112 (7%)
 Frame = +1

Query: 16  QFYNH-----FMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           +FYN      +  L + D     W +  +   PP PR  H A  V   N + IFGG   S
Sbjct: 90  EFYNGQKTYVYGDLYRYDVEKQEWKLVSSPNSPP-PRSSHQA--VAWKNYLYIFGGEFTS 146


Query: 181 RSNPYFN---DIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +   F+   D +   +   +W QL++ G  P PR+ H  V YK +++ FGG
Sbjct: 147 PNQERFHHYKDFWMLDVKTNQWEQLNLKG-CPSPRSGHRMVLYKHKIIIFGG 197
```

```
 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 22/60 (36%), Positives = 37/60 (61%), Gaps = 1/60 (1%)
 Frame = +1

Query: 151 IIIFGGRDDSRSNPY-FNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
           +I++GG   +    Y + D++ Y + ++EW +L    N P PR+SH AVA+K+ +  FGG
Sbjct: 84  LILYGGEFYNGQKTYVYGDLYRYDVEKQEW-KLVSSPNSPPPRSSHQAVAWKNYLYIFGG 142


>gi|9795595|gb|AAF98413.1|AC026238_5  (AC026238) Hypothetical protein [Arabidopsis thaliana]
          Length = 556

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 32/114 (28%), Positives = 52/114 (45%), Gaps = 3/114 (2%)
 Frame = +1

Query: 13  TQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS---R 183
           T   N    L  LD  ++ W  P   G+ P  R  H+A+ V +   + +FGG   S
Sbjct: 96  TDGVNPLKDLYILDTSSHTWKCPSVRGEGPEAREGHSATLVGKR--LFVFGGCGKSSGIN 153

Query: 184 SNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGY 345
              Y+ND++ +      W +    GN P  R SHS  ++K++++  GG   + Y
Sbjct: 154 EEIYYNDVYIFNTETFVWKRAVTIGNPPSARDSHSCSSWKNKLVVIGGEDGHDY 207



 Score = 40.8 bits (94), Expect = 0.003
 Identities = 25/86 (29%), Positives = 37/86 (42%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W    T G PP  R  H+ S  +  N +++ GG D    + Y +D+         W +L+
Sbjct: 171 WKRAVTIGNPPSARDSHSCSSWK--NKLVVIGGEDGH--DYYLSDVHILDTDTLIWKELN 226

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGG 327
             G +  PRA H  V+       FGG
Sbjct: 227 TSGQLLTPRAGHVTVSLGRNFFVFGG 252



 Score = 40.0 bits (92), Expect = 0.005
 Identities = 23/76 (30%), Positives = 34/76 (44%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
           P  R  HT + ++  + + +FGG    R N   N +  +   ++ W Q  I G  P PR
Sbjct: 22  PGKRWGHTCNAIKGGSFLYVFGGY--GRDNCQTNQVHVFDAAKQIWTQPMINGTPPPPRD 79

Query: 280 SHSAVAYKSQVLYFGG 327
           SHS       +  FGG
Sbjct: 80  SHSCTTVGDNLFVFGG 95
```

```
>gi|13173410|gb|AAK14396.1|  (AF339882) attractin [Caenorhabditis elegans]
          Length = 1329

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 26/88 (29%), Positives = 48/88 (54%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W +   T   P  +  H +  + ++ V  I G   D  S+P  N+I  Y +  + W +++
Sbjct: 278 WDLIHPTNNAPTGKASHASIAIDDV-VWSIGGEFFDGSSDP--NNIDVYNVTSRIWSKVE 334

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVS 333
           + G++P+PR  H+ V YK+++  FGGV+
Sbjct: 335 VSGDMPKPRFDHTVVKYKNKLYMFGGVT 362



>gi|15221823|ref|NP_173296.1|  (NM_101719) hypothetical protein [Arabidopsis thaliana]
          Length = 572

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 32/114 (28%), Positives = 52/114 (45%), Gaps = 3/114 (2%)
 Frame = +1

Query: 13  TQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS---R 183
           T   N    L  LD  ++ W  P   G+ P  R  H+A+ V +   + +FGG   S
Sbjct: 114 TDGVNPLKDLYILDTSSHTWKCPSVRGEGPEAREGHSATLVGKR--LFVFGGCGKSSGIN 171

Query: 184 SNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGY 345
              Y+ND++ +      W +    GN P  R SHS  ++K++++  GG   + Y
Sbjct: 172 EEIYYNDVYIFNTETFVWKRAVTIGNPPSARDSHSCSSWKNKLVVIGGEDGHDY 225




 Score = 40.8 bits (94), Expect = 0.003
 Identities = 25/86 (29%), Positives = 37/86 (42%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W    T G PP  R  H+ S  +  N +++ GG D    + Y +D+         W +L+
Sbjct: 189 WKRAVTIGNPPSARDSHSCSSWK--NKLVVIGGEDGH--DYYLSDVHILDTDTLIWKELN 244

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGG 327
             G +  PRA H  V+       FGG
Sbjct: 245 TSGQLLTPRAGHVTVSLGRNFFVFGG 270




 Score = 32.0 bits (71), Expect = 1.3
 Identities = 26/92 (28%), Positives = 39/92 (42%), Gaps = 16/92 (17%)
 Frame = +1
```

```
Query: 100 PLPRIQHTASYVQEMNVIIIFGG--RDDSRSNPY--FNDIFAYKIF------------EK 231
           P  R  HT + ++  + + +FGG  RD+ ++N    F+  F    F            ++
Sbjct: 22  PGKRWGHTCNAIKGGSFLYVFGGYGRDNCQTNQVHVFDAAFGPLWFTCSLTVVILVAAKQ 81

Query: 232 EWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            W Q  I G  P PR SHS      +   FGG
Sbjct: 82  IWTQPMINGTPPPPRDSHSCTTVGDNLFVFGG 113



>gi|17567429|ref|NP_510443.1|  (NM_078042) perlecan like glycoprotein [Caenorhabditis
elegans]
 gi|2497611|sp|Q19981|YC81_CAEEL  Hypothetical 143.1 kDa protein F33C8.1 in chromosome X
precursor
 gi|7500382|pir||T21694  hypothetical protein F33C8.1 - Caenorhabditis elegans
 gi|3876598|emb|CAA93653.1|  (Z69790) predicted using Genefinder~Weak similarity to Mouse
basement membrane-specific heparin sulphate proteoglycan (SW:PGBM_MOUSE), contains
similarity to Pfam domain: PF00053 (Laminin EGF-like (Domains III and V)), Score=69.2, E-
value=2.8e-17, N=2;>
          Length = 1291

 Score = 50.8 bits (120), Expect = 3e-06
 Identities = 26/88 (29%), Positives = 48/88 (54%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W +   T   P  +  H +  + ++ V  I G   D  S+P  N+I  Y +  + W +++
Sbjct: 278 WDLIHPTNNAPTGKASHASIAIDDV-VWSIGGEFFDGSSDP--NNIDVYNVTSRIWSKVE 334

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVS 333
           + G++P+PR  H+ V YK+++  FGGV+
Sbjct: 335 VSGDMPKPRFDHTVVKYKNKLYMFGGVT 362



>gi|18043893|gb|AAH19800.1|AAH19800  (BC019800) Similar to Rab9 effector p40 [Mus musculus]
          Length = 372

 Score = 50.4 bits (119), Expect = 4e-06
 Identities = 32/93 (34%), Positives = 44/93 (46%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W  PET G PP PR  H    V     + I GG    +   +F+D+    I +  W +L
Sbjct: 180 WSQPETHGSPPSPRHGHVM--VAAGTKLFIHGGLAGDK---FFDDLHCIDIGDMSWQKLG 234

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYL 348
             G +P   A+H+AVA    V  FGG++  G L
Sbjct: 235 PTGAVPVGCAAHAAVAVGHHVYMFGGMTATGAL 267




 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 27/94 (28%), Positives = 44/94 (46%)
 Frame = +1
```

```
Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W  PE TG PP PR  HT+S     N + +FGG +          +  +      W Q +
Sbjct: 126 WSTPEVTGSPPSPRTFHTSS-AAIGNQLYVFGGGERGAQPVEDVKLHVFDANTLTWSQPE 184

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYLE 351
            +G+ P PR  H  VA  +++   GG++ + + +
Sbjct: 185 THGSPPSPRHGHVMVAAGTKLFIHGGLAGDKFFD 218



 Score = 33.9 bits (76), Expect = 0.34
 Identities = 25/78 (32%), Positives = 40/78 (51%), Gaps = 3/78 (3%)
 Frame = +1

Query: 103 LPRIQHTASYVQEMNV--IIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPR 276
           LPR +H AS++   +   I +FGG D S +    N +      ++ W   ++ G+ P PR
Sbjct: 84  LPRYEH-ASFLPSCSPHSIWVFGGADQSGNR---NCLQVMSPEDRTWSTPEVTGSPPSPR 139

Query: 277 ASH-SAVAYKSQVLYFGG 327
             H S+ A  +Q+  FGG
Sbjct: 140 TFHTSSAAIGNQLYVFGG 157



>gi|7019405|ref|NP_037452.1|  (NM_013320) host cell factor 2 [Homo sapiens]
 gi|4689221|gb|AAD27814.1|AF117210_1  (AF117210) host cell factor 2 [Homo sapiens]
          Length = 792

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 38/113 (33%), Positives = 55/113 (48%), Gaps = 17/113 (15%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----RDDSRSNPY--- 195
           L +LD     W  PET G  PLPR  HTAS +   N + IFGG      +++ ++P+
Sbjct: 222 LWQLDLETMSWSKPETKGTVPLPRSLHTASVIG--NKMYIFGGWVPHKGENTETSPHDCE 279

Query: 196 --FNDIFAY-KIFEKEWVQL------DIYGNIPRPRASHSAVAYKSQVLYFGG 327
                 F+Y  +   EW  L      D   + PRPRA H AVA  +++ ++ G
Sbjct: 280 WRCTSSFSYLNLDTTEWTTLVSDSQEDKKNSRPRPRAGHCAVAIGTRLYFWSG 332



 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 26/90 (28%), Positives = 43/90 (46%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNV----IIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    + +     + +FGG   +R     +D++   +    W
Sbjct: 177 WSIPVTKGVVPSPRESHTAVIYCKKDSGSPKMYVFGGMCGAR----LDDLWQLDLETMSW 232

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            + +  G +P PR+ H+A    +++  FGG
Sbjct: 233 SKPETKGTVPLPRSLHTASVIGNKMYIFGG 262
```

```
 Score = 40.4 bits (93), Expect = 0.004
 Identities = 22/77 (28%), Positives = 40/77 (51%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
           P+PR +H    V    ++IIFGG ++  ++    ++  Y     +W    + G+IP   A
Sbjct: 18  PVPRARHGHRAVAIRELMIIFGGGNEGIAD----ELHVYNTATNQWFLPAVRGDIPPGCA 73

Query: 280 SHSAVAYKSQVLYFGGV 330
           +H  V   +++L FGG+
Sbjct: 74  AHGFVCDGTRILVFGGM 90




 Score = 39.3 bits (90), Expect = 0.008
 Identities = 31/100 (31%), Positives = 45/100 (45%), Gaps = 16/100 (16%)
 Frame = +1

Query: 79  PETTGKPPLPRIQHTASYVQEMNVIIIFGG----RDDSRSN--PYFNDIFAYKIFEKE-- 234
           P  +G PP PR+ H+ S     N   +FGG     +DS +N   Y ND +  ++
Sbjct: 117 PPPSGLPPCPRLGHSFSLYG--NKCYLFGGLANESEDSNNNVPRYLNDFYELELQHGSGV 174

Query: 235 --WVQLDIYGNIPRPRASHSAVAY------KSQVLYFGGV 330
             W      G +P PR SH+AV Y        ++  FGG+
Sbjct: 175 VGWSIPVTKGVVPSPRESHTAVIYCKKDSGSPKMYVFGGM 214




 Score = 32.3 bits (72), Expect = 0.99
 Identities = 24/94 (25%), Positives = 41/94 (43%), Gaps = 4/94 (4%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W +P   G  P     H   +V +   I++FGG  +     Y N+++  +     W +
Sbjct: 57  NQWFLPAVRGDIPPGCAAH--GFVCDGTRILVFGGMVEY--GRYSNELYELQASRWLWKK 112

Query: 244 LDIY----GNIPRPRASHSAVAYKSQVLYFGGVS 333
           +  +    G  P PR  HS   Y ++   FGG++
Sbjct: 113 VKPHPPPSGLPPCPRLGHSFSLYGNKCYLFGGLA 146




>gi|16306876|gb|AAH06558.1|AAH06558  (BC006558) Similar to host cell factor 2 [Homo sapiens]
          Length = 412

 Score = 50.1 bits (118), Expect = 5e-06
 Identities = 38/113 (33%), Positives = 55/113 (48%), Gaps = 17/113 (15%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----RDDSRSNPY--- 195
           L +LD     W  PET G  PLPR  HTAS +   N + IFGG      +++ ++P+
Sbjct: 222 LWQLDLETMSWSKPETKGTVPLPRSLHTASVIG--NKMYIFGGWVPHKGENTETSPHDCE 279

Query: 196 --FNDIFAY-KIFEKEWVQL------DIYGNIPRPRASHSAVAYKSQVLYFGG 327
                 F+Y  +   EW  L      D   + PRPRA H AVA  +++ ++ G
Sbjct: 280 WRCTSSFSYLNLDTTEWTTLVSDSQEDKKNSRPRPRAGHCAVAIGTRLYFWSG 332
```

```
 Score = 43.5 bits (101), Expect = 4e-04
 Identities = 26/90 (28%), Positives = 43/90 (46%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNV----IIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    + +     + +FGG   +R     +D++   +    W
Sbjct: 177 WSIPVTKGVVPSPRESHTAVIYCKKDSGSPKMYVFGGMCGAR----LDDLWQLDLETMSW 232

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            + +  G +P PR+ H+A    +++  FGG
Sbjct: 233 SKPETKGTVPLPRSLHTASVIGNKMYIFGG 262
```

```
 Score = 40.4 bits (93), Expect = 0.004
 Identities = 22/77 (28%), Positives = 40/77 (51%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
           P+PR +H    V    ++IIFGG ++  ++    ++  Y     +W    + G+IP   A
Sbjct: 18  PVPRARHGHRAVAIRELMIIFGGGNEGIAD----ELHVYNTATNQWFLPAVRGDIPPGCA 73

Query: 280 SHSAVAYKSQVLYFGGV 330
           +H  V   +++L FGG+
Sbjct: 74  AHGFVCDGTRILVFGGM 90
```

```
 Score = 39.3 bits (90), Expect = 0.008
 Identities = 31/100 (31%), Positives = 45/100 (45%), Gaps = 16/100 (16%)
 Frame = +1

Query: 79  PETTGKPPLPRIQHTASYVQEMNVIIIFGG----RDDSRSN--PYFNDIFAYKIFEKE-- 234
           P  +G PP PR+ H+ S     N   +FGG     +DS +N   Y ND +  ++
Sbjct: 117 PPPSGLPPCPRLGHSFSLYG--NKCYLFGGLANESEDSNNNVPRYLNDFYELELQHGSGV 174

Query: 235 --WVQLDIYGNIPRPRASHSAVAY------KSQVLYFGGV 330
             W      G +P PR SH+AV Y        ++  FGG+
Sbjct: 175 VGWSIPVTKGVVPSPRESHTAVIYCKKDSGSPKMYVFGGM 214
```

```
 Score = 32.3 bits (72), Expect = 0.99
 Identities = 24/94 (25%), Positives = 41/94 (43%), Gaps = 4/94 (4%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W +P   G  P     H   +V +   I++FGG  +     Y N+++  +     W +
Sbjct: 57  NQWFLPAVRGDIPPGCAAH--GFVCDGTRILVFGGMVEY--GRYSNELYELQASRWLWKK 112
```

```
Query: 244 LDIY----GNIPRPRASHSAVAYKSQVLYFGGVS 333
           +  +    G  P PR  HS   Y ++   FGG++
Sbjct: 113 VKPHPPPSGLPPCPRLGHSFSLYGNKCYLFGGLA 146



>gi|18277872|sp|Q39610|DYHA_CHLRE  Dynein alpha chain, flagellar outer arm (DHC alpha)
 gi|6007859|gb|AAA57316.2|  (L26049) dynein heavy chain alpha [Chlamydomonas reinhardtii]
          Length = 4499

 Score = 49.7 bits (117), Expect = 6e-06
 Identities = 30/113 (26%), Positives = 50/113 (43%)
 Frame = +1

Query: 4   DKNTQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSR 183
           D     +N    L   D     W        PP PR +H A  + +   +++FGG +
Sbjct: 40  DGKAAAFNDLYELDTSDPDEYKWKELVVANAPP-PRARHAAIALDDKR-LLVFGGLN--- 94

Query: 184 SNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
               +ND++ +   +K W  +++ G  P PRA  +A  + S+V  FGG   +G
Sbjct: 95  KRIRYNDVWLFNYDDKSWTCMEVEGAAPEPRAHFTATRFGSRVFIFGGYGGSG 147



 Score = 48.1 bits (113), Expect = 2e-05
 Identities = 30/92 (32%), Positives = 49/92 (52%), Gaps = 4/92 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIF---EKEWV 240
           W +P   G+ P PR  H+ + + E  V+    GR D ++   FND++        E +W
Sbjct: 5   WEVPNAQGEAPCPRSGHSFTVLGERFVLFGGCGRKDGKA-AAFNDLYELDTSDPDEYKWK 63

Query: 241 QLDIYGNIPRPRASHSAVAY-KSQVLYFGGVS 333
           +L +  N P PRA H+A+A    ++L FGG++
Sbjct: 64  EL-VVANAPPPRARHAAIALDDKRLLVFGGLN 94



 Score = 31.2 bits (69), Expect = 2.2
 Identities = 25/81 (30%), Positives = 32/81 (38%), Gaps = 4/81 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK--EWVQ 243
           W   E  G  P PR   TA+       + IFGG   S     +N+++     E    W
Sbjct: 112 WTCMEVEGAAPEPRAHFTATRFGSR--VFIFGGYGGS--GQVYNEMWVLHFGEDGFRWQN 167

Query: 244 L--DIYGNIPRPRASHSAVAY 300
           +   I G  P PR  HSA  Y
Sbjct: 168 ITESIEGTGPAPRFDHSAFIY 188



>gi|18398038|ref|NP_566316.1|  (NM_111651) expressed protein [Arabidopsis thaliana]
 gi|6466955|gb|AAF13090.1|AC009176_17  (AC009176) unknown protein [Arabidopsis thaliana]
 gi|6648184|gb|AAF21182.1|AC013483_6  (AC013483) unknown protein [Arabidopsis thaliana]
 gi|11692832|gb|AAG40019.1|AF324668_1  (AF324668) MLP3.17 [Arabidopsis thaliana]
```

```
 gi|11993873|gb|AAG42920.1|AF329503_1  (AF329503) unknown protein [Arabidopsis thaliana]
 gi|14517448|gb|AAK62614.1|  (AY039559) AT3g07720/F17A17_6 [Arabidopsis thaliana]
          Length = 329

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 30/103 (29%), Positives = 44/103 (42%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L   D     W I E +G  P PR+    + V    +I  FGGRD +      N+++ +
Sbjct: 52  LYVFDLETQTWSIQEASGDAPPPRVGVAMAAVGP--IIYFFGGRDSTHQE--LNELYCFN 107

Query: 220 IFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYL 348
               +W  L      P+ R+ HS  A    V  FGG  ++G L
Sbjct: 108 TLTNQWKLLSSGETGPQNRSYHSITADSQNVYVFGGCGVDGRL 150




 Score = 34.3 bits (77), Expect = 0.26
 Identities = 23/86 (26%), Positives = 36/86 (41%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           WV  +  G  P  R  H  + V   N +  FGG    R  P  N ++ + +  + W   +
Sbjct: 10  WVQLKQKGTGPGARSSHAIALVG--NKMYAFGGEFQPRV-PVDNQLYVFDLETQTWSIQE 66

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGG 327
             G+ P PR   +  A    + +FGG
Sbjct: 67  ASGDAPPPRVGVAMAAVGPIIYFFGG 92



>gi|15237715|ref|NP_196062.1|  (NM_120524) putative protein [Arabidopsis thaliana]
 gi|11358331|pir||T48438  hypothetical protein T32M21.20 - Arabidopsis thaliana
 gi|7406446|emb|CAB85548.1|  (AL162875) putative protein [Arabidopsis thaliana]
          Length = 514

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 34/118 (28%), Positives = 56/118 (46%), Gaps = 1/118 (0%)
 Frame = +1

Query: 1   DDKNTQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           +DKN +  N    L  L      W + ET    P+PR  HTA+   +   ++IFGG   S
Sbjct: 171 EDKNRRLLND---LHVLHLETMTWDVVETKQTRPVPRFDHTAATHSD-RYLLIFGGCSHS 226


Query: 181 RSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAY-KSQVLYFGGVSMNGYLE 351
               +++D+    +   EW Q  + G++  PRA H+ +   ++  +  GG +  G LE
Sbjct: 227 I---FYSDLHILDLQTMEWSQPHVQGDVVTPRAGHAGITIDENWYIVGGGDNSTGCLE 281




 Score = 35.0 bits (79), Expect = 0.15
 Identities = 14/32 (43%), Positives = 19/32 (58%)
 Frame = +1
```

```
Query: 244 LDIYGNIPRPRASHSAVAYKSQVLYFGGVSMN 339
           +D++GN+P  R  HS     S+VL FGG   N
Sbjct: 143 IDVFGNVPASRGGHSITLVGSRVLVFGGEDKN 174




>gi|15724206|gb|AAL06496.1|AF412043_1  (AF412043) AT5g04420/T32M21_20 [Arabidopsis thaliana]
          Length = 514

 Score = 48.9 bits (115), Expect = 1e-05
 Identities = 34/118 (28%), Positives = 56/118 (46%), Gaps = 1/118 (0%)
 Frame = +1

Query: 1   DDKNTQFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDS 180
           +DKN +  N    L  L      W + ET    P+PR  HTA+   +   ++IFGG   S
Sbjct: 171 EDKNRRLLND---LHVLHLETMTWDVVETKQTRPVPRFDHTAATHSD-RYLLIFGGCSHS 226

Query: 181 RSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAY-KSQVLYFGGVSMNGYLE 351
               +++D+    +   EW Q  + G++  PRA H+ +   ++  +  GG +  G LE
Sbjct: 227 I---FYSDLHILDLQTMEWSQPHVQGDVVTPRAGHAGITIDENWYIVGGGDNSTGCLE 281





 Score = 34.3 bits (77), Expect = 0.26
 Identities = 14/32 (43%), Positives = 19/32 (58%)
 Frame = +1

Query: 244 LDIYGNIPRPRASHSAVAYKSQVLYFGGVSMN 339
           +D++GN+P  R  HS     S+VL FGG   N
Sbjct: 143 IDVFGNVPASRDGHSITLVGSRVLVFGGEDKN 174




>gi|18029285|gb|AAL56463.1|  (AF374376) similar to host cell factor [Oikopleura dioica]
          Length = 683

 Score = 48.5 bits (114), Expect = 1e-05
 Identities = 29/91 (31%), Positives = 46/91 (49%), Gaps = 4/91 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQ---EMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWV 240
           W  PE TG  P PR  HT+   +      ++I++GG+D+ R      DIF+       W
Sbjct: 202 WSKPECTGTAPCPRESHTSVVYKPEGRAQMVIVYGGKDERRR---LGDIFSLDTGTNSWT 258


Query: 241 QLDIYGNIPRPRASHSAVAY-KSQVLYFGGV 330
           +L   G  P  R+ HSAV   +++++  GG+
Sbjct: 259 ELKPLGVPPLNRSLHSAVMISETKMVVLGGL 289





 Score = 43.1 bits (100), Expect = 6e-04
 Identities = 25/96 (26%), Positives = 46/96 (47%), Gaps = 16/96 (16%)
 Frame = +1
```

```
Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP-------YFNDIFAYKIFEK----E 234
           +G  P PR+ H+ + +     +++FGG  +   +P       Y+ND++   +
Sbjct: 142 SGMVPCPRMCHSFTNLGSSGTVVLFGGVANKSDSPERIFRPDYYNDVYYLALKGGTGIYN 201

Query: 235 WVQLDIYGNIPRPRASHSAVAYKSQ-----VLYFGG 327
           W + +  G  P PR SH++V YK +     V+ +GG
Sbjct: 202 WSKPECTGTAPCPRESHTSVVYKPEGRAQMVIVYGG 237



 Score = 38.1 bits (87), Expect = 0.018
 Identities = 22/81 (27%), Positives = 38/81 (46%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           TG+   PR +H    V    +I++FGG D    +    ++  Y      W +    G+IP
Sbjct: 36  TGQGLFPRARHGHRLVAIRGLILLFGGGDGEIRD----ELMVYSTENGNWFKPSCSGDIP 91

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A+H  +A  +++L  GG+
Sbjct: 92  PGVAAHGMIADGTRILIHGGM 112



 Score = 30.8 bits (68), Expect = 2.9
 Identities = 13/42 (30%), Positives = 21/42 (49%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFG 165
           +  LD   N W   +  G PPL R  H+A  + E  ++++ G
Sbjct: 247 IFSLDTGTNSWTELKPLGVPPLNRSLHSAVMISETKMVVLGG 288



>gi|9280670|gb|AAF86539.1|AC002560_32  (AC002560) F21B7.7 [Arabidopsis thaliana]
          Length = 846

 Score = 47.8 bits (112), Expect = 2e-05
 Identities = 23/85 (27%), Positives = 44/85 (51%), Gaps = 11/85 (12%)
 Frame = +1

Query: 121 TASYVQEMNVIIIFGGRDDSRSNP-----------YFNDIFAYKIFEKEWVQLDIYGNIP 267
           TA +V   + +I+FGG   + +N              N + ++ +  ++W +L+  G++P
Sbjct: 55  TAVFVNNSHQLILFGGSTTAVANHNSSLPEISLDGVTNSVHSFDVLTRKWTRLNPIGDVP 114

Query: 268 RPRASHSAVAYKSQVLYFGGVSMNG 342
            PRA H+A  Y + +L  GG+  +G
Sbjct: 115 SPRACHAAALYGTLILIQGGIGPSG 139



 Score = 34.3 bits (77), Expect = 0.26
 Identities = 25/95 (26%), Positives = 41/95 (42%), Gaps = 1/95 (1%)
 Frame = +1
```

```
Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W      G  P PR  H A+    +  I+I GG   S   P   D++   +   +W++
Sbjct: 104 WTRLNPIGDVPSPRACHAAALYGTL--ILIQGGIGPS--GPSDGDVYMLDMTNNKWIKFL 159

Query: 250 IYGNIPRPRASH-SAVAYKSQVLYFGGVSMNGYLE 351
           + G  P PR  H   +A +  ++ F G + N  L+
Sbjct: 160 VGGETPSPRYGHVMDIAAQRWLVIFSGNNGNEILD 194




 Score = 32.0 bits (71), Expect = 1.3
 Identities = 27/100 (27%), Positives = 42/100 (42%), Gaps = 2/100 (2%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           LD   NN W+     G+ P PR  H    +     ++IF G   +  N   +D +A
Sbjct: 148 LDMTNNKWIKFLVGGETPSPRYGHVMD-IAAQRWLVIFSG---NNGNEILDDTWALDTRG 203

Query: 229 K-EWVQLDIYGNIPRPRASHSAVAYKSQV-LYFGGVSMNG 342
              W +L+  GN P  R   S  + +  + L  GG+  +G
Sbjct: 204 PFSWDRLNPSGNQPSGRMYASGSSREDGIFLLCGGIDHSG 243




>gi|6321952|ref|NP_012028.1|  (NC_001140) protein containing kelch repeats, similar to
YGR238c; Kel1p [Saccharomyces cerevisiae]
 gi|731732|sp|P38853|KEL1_YEAST  Kelch repeats protein 1
 gi|626671|pir||S46769  hypothetical protein YHR158c - yeast (Saccharomyces cerevisiae)
 gi|500665|gb|AAB68991.1|  (U10397) Yhr158cp [Saccharomyces cerevisiae]
          Length = 1164

 Score = 47.8 bits (112), Expect = 2e-05
 Identities = 28/94 (29%), Positives = 52/94 (54%), Gaps = 2/94 (2%)
 Frame = +1

Query: 52  DQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK 231
           D   N W I +TTG+ P P +Q  A+ V   +++ + GG+D+   + Y N ++   +  +
Sbjct: 328 DPAINDWFIIDTTGEKP-PPVQEHATVVYN-DLMCVVGGKDE--HDAYLNSVYFLNLKSR 383

Query: 232 EWVQLDIY-GNIPRPRASHSAVAYKS-QVLYFGG 327
           +W +L ++   IP+ R+ HS    K+ ++L  GG
Sbjct: 384 KWFKLPVFTAGIPQGRSGHSLTLLKNDKILIMGG 417




 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 28/99 (28%), Positives = 45/99 (45%), Gaps = 8/99 (8%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYV---QEMNVIIIFGGRDDSRSNPYFNDIFAYKIF----- 225
           W +P   G  PL R  H  S +   Q    + +FGG+ D   + YFND+  Y +
Sbjct: 224 WTVPAPVGPRPLGRYGHKISIIATTQMKTKLYVFGGQFD---DTYFNDLAVYDLSSFRRP 280

Query: 226 EKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNG 342
           +  W  L      P P  + + ++Y S++  FGG ++ G
Sbjct: 281 DSHWEFLKPRTFTPPPITNFTMISYDSKLWVFGGDTLQG 319
```

```
 Score = 38.9 bits (89), Expect = 0.011
 Identities = 26/81 (32%), Positives = 36/81 (44%), Gaps = 5/81 (6%)
 Frame = +1

Query: 100 PLPRIQHTAS-YVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK----EWVQLDIYGNI 264
           P PR +H AS YV + N I + GG  D      + D +    F+         +DI
Sbjct: 122 PFPRYRHVASAYVTDKNQIYVIGGLHDQS---VYGDTWILTAFDNATRFSTTTIDISEAT 178

Query: 265 PRPRASHSAVAYKSQVLYFGG 327
           P PR  H+AV   +  + FGG
Sbjct: 179 PPPRVGHAAVLCGNAFVVFGG 199



 Score = 34.7 bits (78), Expect = 0.20
 Identities = 23/93 (24%), Positives = 37/93 (39%), Gaps = 2/93 (2%)
 Frame = +1

Query: 55  QPNNPWVI--PETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           +P++ W    P T   PP+       + +   + + +FGG          ND+F Y
Sbjct: 279 RPDSHWEFLKPRTFTPPPITNF----TMISYDSKLWVFGG---DTLQGLVNDVFMYDPAI 331

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +W  +D  G  P P   H+ V Y   +   GG
Sbjct: 332 NDWFIIDTTGEKPPPVQEHATVVYNDLMCVVGG 364



>gi|7291384|gb|AAF46812.1| (AE003456) CG6758 gene product [Drosophila melanogaster]
          Length = 667

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 25/73 (34%), Positives = 42/73 (57%)
 Frame = +1

Query: 109 RIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRASHS 288
           R  H+A  V++ N + +FGG   S S+  FND++ + +    W +    G  P P+ S S
Sbjct: 103 RFAHSA--VRQDNSMYVFGG--GSSSDTTFNDLWRFDLTHMRWARPVATGTYPSPKGSAS 158

Query: 289 AVAYKSQVLYFGG 327
            VA+++Q++ FGG
Sbjct: 159 MVAWRNQLILFGG 171



 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 33/105 (31%), Positives = 50/105 (47%), Gaps = 6/105 (5%)
 Frame = +1

Query: 31  FMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-RDDSRSNPY---- 195
           F  L + D  +  W  P  TG  P P+   +AS V   N +I+FGG R  S   PY
Sbjct: 128 FNDLWRFDLTHMRWARPVATGTYPSPK--GSASMVAWRNQLILFGGWRYPSLHPPYQPWC 185
```

```
Query: 196 -FNDIFAYKIFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            F+++  Y + +  W+      + P P A HSA  +  +++ FGG
Sbjct: 186 LFDELHYYDLGKNRWLLRSSLSS-PPPMAGHSATVHGDRMVVFGG 229



>gi|15126782|gb|AAH12312.1|AAH12312  (BC012312) Similar to hypothetical protein
DKFZp434G0522 [Mus musculus]
          Length = 584

 Score = 47.4 bits (111), Expect = 3e-05
 Identities = 29/88 (32%), Positives = 46/88 (51%), Gaps = 4/88 (4%)
 Frame = +1

Query: 76  IPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPY-FNDIFAYKIFEKEWVQLDI 252
           + ET   PP PR+  + S   E + +I+FGG   +    + +N+++ Y I +  W ++DI
Sbjct: 53  VTETPCPPPSPRLNASLSAHPEKDELILFGGEYFNGQKTFMYNELYIYSIRKDTWTKVDI 112

Query: 253 YGNIPRPRASHSAVAYK---SQVLYFGG 327
            G  PR R +H AV       Q+  FGG
Sbjct: 113 PGPPPR-RCAHQAVVVPQGGGQLWVFGG 139



 Score = 38.1 bits (87), Expect = 0.018
 Identities = 28/108 (25%), Positives = 50/108 (45%), Gaps = 2/108 (1%)
 Frame = +1

Query: 16  QFYNHFMALLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-RDDSRSNP 192
           QFY H+  L  L      W    +TG P         ++ +++   I+FGG  + +R
Sbjct: 148 QFY-HYKDLWVLHLATKTWEQIRSTGGPSGRSGHRMVAWKRQL---ILFGGFHESARDYI 203


Query: 193 YFNDIFAYKIFEKEWVQLDIYGNIPRPRAS-HSAVAYKSQVLYFGGVS 333
           Y++D++ + +   +W +L   G  P PR+    AV  +  +  +GG S
Sbjct: 204 YYSDVYTFSLDTFQWSKLSPSGAGPTPRSGCLMAVTPQGSIAIYGGYS 251



>gi|6680185|ref|NP_032250.1| ☐ (NM_008224) host cell factor C1; VP16-accessory protein
[Mus musculus]
 gi|4098678|gb|AAD09225.1| ☐ (U80821) C1 transcription factor [Mus musculus]
          Length = 2045

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225
```

```
 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIETLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272



 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 232 LWTLDIETLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339



 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100



>gi|1293686|gb|AAB01163.1| ☐ (U53925) transcription factor C1 (HCF) [Mus musculus]
          Length = 2045

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188
```

```
Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225




 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIETLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +   + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272




 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +    N + +FGG       P   D
Sbjct: 232 LWTLDIETLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA +++   + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100




>gi|7304317|gb|AAF59349.1|  (AE003844) CG1710 gene product [Drosophila melanogaster]
         Length = 1499

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 27/93 (29%), Positives = 44/93 (47%), Gaps = 4/93 (4%)
 Frame = +1
```

```
Query: 61  NNPWVIPETTGKPPLPRIQHTA-SYVQEMN---VIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           N  W++P+T G  P PR  HT  S+  + N    ++I+GG    R      D++  +
Sbjct: 226 NGKWIVPKTYGDSPPPRESHTGISFATKSNGNLNLLIYGGMSGCR----LGDLWLLETDS 281

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
             W +    G  P PR+ HS+    +++  FGG
Sbjct: 282 MTWSKPKTSGEAPLPRSLHSSTMIGNKMYVFGG 314




 Score = 42.7 bits (99), Expect = 7e-04
 Identities = 32/106 (30%), Positives = 53/106 (49%), Gaps = 13/106 (12%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----RDDSRSN-----PYF 198
           L+  +  W  P+T+G+ PLPR  H+++ +   N + +FGG      +DS+S
Sbjct: 277 LETDSMTWSKPKTSGEAPLPRSLHSSTMIG--NKMYVFGGWVPLVINDSKSTTEREWKCT 334

Query: 199 NDIFAYKIFEKEW--VQLD-IYGNIPRPRASHSAVAYKSQVLYFGG 327
           N +    +    W  V LD +  N+PR RA H AV  +S++  + G
Sbjct: 335 NTLAVLDLETMTWENVTLDTVEENVPRARAGHCAVGIQSRLYVWSG 380




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 20/81 (24%), Positives = 41/81 (49%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           TG  P PR  H A  ++E+  +++FGG ++       +++  Y     +W    + G++P
Sbjct: 67  TGPQPRPRHGHRAINIKEL--MVVFGGGNEG----IVDELHVYNTVTNQWYVPVLKGDVP 120

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++  FGG+
Sbjct: 121 NGCAAYGFVVEGTRMFVFGGM 141




 Score = 31.6 bits (70), Expect = 1.7
 Identities = 23/94 (24%), Positives = 40/94 (42%), Gaps = 4/94 (4%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W +P   G  P     +   +V E   + +FGG  +     Y N+++  +  + EW +
Sbjct: 108 NQWYVPVLKGDVPNGCAAY--GFVVEGTRMFVFGGMIEY--GKYSNELYELQATKWEWRK 163

Query: 244 L----DIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
           +       G  P PR  HS      ++  FGG++
Sbjct: 164 MYPESPDSGLSPCPRLGHSFTMVGEKIFLFGGLA 197




>gi|14970918|emb|CAC44472.1| ☐ (AJ320236) host cell factor [Drosophila melanogaster]
          Length = 1500
```

```
 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 27/93 (29%), Positives = 44/93 (47%), Gaps = 4/93 (4%)
 Frame = +1

Query: 61  NNPWVIPETTGKPPLPRIQHTA-SYVQEMN---VIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           N  W++P+T G  P PR  HT  S+  + N    ++I+GG    R      D++  +
Sbjct: 226 NGKWIVPKTYGDSPPPRESHTGISFATKSNGNLNLLIYGGMSGCR----LGDLWLLETDS 281

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
             W +    G  P PR+ HS+    +++  FGG
Sbjct: 282 MTWSKPKTSGEAPLPRSLHSSTMIGNKMYVFGG 314




 Score = 42.7 bits (99), Expect = 7e-04
 Identities = 32/106 (30%), Positives = 53/106 (49%), Gaps = 13/106 (12%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----RDDSRSN-----PYF 198
           L+  +  W  P+T+G+ PLPR  H+++ +   N + +FGG      +DS+S
Sbjct: 277 LETDSMTWSKPKTSGEAPLPRSLHSSTMIG--NKMYVFGGWVPLVINDSKSTTEREWKCT 334

Query: 199 NDIFAYKIFEKEW--VQLD-IYGNIPRPRASHSAVAYKSQVLYFGG 327
           N +    +    W  V LD +  N+PR RA H AV  +S++  + G
Sbjct: 335 NTLAVLDLETMTWENVTLDTVEENVPRARAGHCAVGIQSRLYVWSG 380




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 20/81 (24%), Positives = 41/81 (49%)
 Frame = +1


Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           TG  P PR  H A  ++E+  +++FGG ++       +++  Y     +W    + G++P
Sbjct: 67  TGPQPRPRHGHRAINIKEL--MVVFGGGNEG----IVDELHVYNTVTNQWYVPVLKGDVP 120

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++  FGG+
Sbjct: 121 NGCAAYGFVVEGTRMFVFGGM 141




 Score = 31.6 bits (70), Expect = 1.7
 Identities = 23/94 (24%), Positives = 40/94 (42%), Gaps = 4/94 (4%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W +P   G  P     +   +V E   + +FGG  +     Y N+++  +  + EW +
Sbjct: 108 NQWYVPVLKGDVPNGCAAY--GFVVEGTRMFVFGGMIEY--GKYSNELYELQATKWEWRK 163

Query: 244 L----DIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
           +       G  P PR  HS       ++  FGG++
Sbjct: 164 MYPESPDSGLSPCPRLGHSFTMVGEKIFLFGGLA 197
```

```
>gi|4885403|ref|NP_005325.1|  (NM_005334) host cell factor C1 (VP16-accessory protein)
[Homo sapiens]
 gi|2135400|pir||I37453  host cell factor C1 - human
 gi|558349|emb|CAA55790.1|  (X79198) host cell factor [Homo sapiens]
          Length = 1938

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 32  GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 89

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 90  IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 126



 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1


Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 88  WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIDTLTW 143

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 144 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 173



 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 133 LWTLDIDTLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 185

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 186 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 240



>gi|17864168|ref|NP_524621.1|  (NM_079882) Host cell factor [Drosophila melanogaster]
 gi|13507075|gb|AAK28427.1|AF251006_1  (AF251006) host cell factor HCF [Drosophila
melanogaster]
          Length = 1500
```

```
 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 27/93 (29%), Positives = 44/93 (47%), Gaps = 4/93 (4%)
 Frame = +1

Query: 61  NNPWVIPETTGKPPLPRIQHTA-SYVQEMN---VIIIFGGRDDSRSNPYFNDIFAYKIFE 228
           N  W++P+T G  P PR  HT  S+  + N    ++I+GG    R      D++  +
Sbjct: 226 NGKWIVPKTYGDSPPPRESHTGISFATKSNGNLNLLIYGGMSGCR----LGDLWLLETDS 281

Query: 229 KEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
             W +    G  P PR+ HS+    +++  FGG
Sbjct: 282 MTWSKPKTSGEAPLPRSLHSSTMIGNKMYVFGG 314




 Score = 42.7 bits (99), Expect = 7e-04
 Identities = 32/106 (30%), Positives = 53/106 (49%), Gaps = 13/106 (12%)
 Frame = +1

Query: 49  LDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGG-----RDDSRSN-----PYF 198
           L+   +  W  P+T+G+ PLPR  H+++ +   N + +FGG      +DS+S
Sbjct: 277 LETDSMTWSKPKTSGEAPLPRSLHSSTMIG--NKMYVFGGWVPLVINDSKSTTEREWKCT 334

Query: 199 NDIFAYKIFEKEW--VQLD-IYGNIPRPRASHSAVAYKSQVLYFGG 327
           N +    +    W  V LD +  N+PR RA H AV  +S++  + G
Sbjct: 335 NTLAVLDLETMTWENVTLDTVEENVPRARAGHCAVGIQSRLYVWSG 380




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 20/81 (24%), Positives = 41/81 (49%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           TG  P PR  H A  ++E+  +++FGG ++       +++  Y     +W    + G++P
Sbjct: 67  TGPQPRPRHGHRAINIKEL--MVVFGGGNEG----IVDELHVYNTVTNQWYVPVLKGDVP 120

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++  FGG+
Sbjct: 121 NGCAAYGFVVEGTRMFVFGGM 141




 Score = 31.6 bits (70), Expect = 1.7
 Identities = 23/94 (24%), Positives = 40/94 (42%), Gaps = 4/94 (4%)
 Frame = +1

Query: 64  NPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQ 243
           N W +P   G  P     +   +V E   + +FGG  +     Y N+++  +  + EW +
Sbjct: 108 NQWYVPVLKGDVPNGCAAY--GFVVEGTRMFVFGGMIEY--GKYSNELYELQATKWEWRK 163

Query: 244 L----DIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
           +       G  P PR  HS      ++  FGG++
Sbjct: 164 MYPESPDSGLSPCPRLGHSFTMVGEKIFLFGGLA 197
```

```
>gi|1362760|pir||A56088  host cell factor C1 precursor - human
          Length = 2035

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225




 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIDTLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272




 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 232 LWTLDIDTLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79
```

```
Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100




>gi|14768289|ref|XP_048390.1|  (XM_048390) hypothetical protein XP_048390 [Homo sapiens]
 gi|476805|pir||A40718  host cell factor C1 precursor - human
 gi|399752|gb|AAB27583.1|  HCF, C1, VCAF, CFF=VP16 accessory protein host cell factor
[human, HeLa cell, Peptide, 2035 aa]
 gi|1302663|gb|AAC51751.1|  (U52112) host cell factor 1 [Homo sapiens]
          Length = 2035

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225




 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIDTLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272




 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1


Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 232 LWTLDIDTLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA +++   + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339
```

```
 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
            +G  P PR  H A  ++E+  I++FGG ++      +++   Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
               A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100



>gi|1708193|sp|P51610|HFC1_HUMAN  Host cell factor C1 (HCF) (VP16 accessory protein) (HFC1)
(VCAF) (CFF)
          Length = 2035

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
            G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
                YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225



 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
            W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIDTLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
             +  + G   P PR+ HSA    +++   FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272



 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
            L  LD     W  P   +G  PLPR  H+A+ +    N + +FGG       P   D
Sbjct: 232 LWTLDIDTLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284
```

```
Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100



>gi|1708194|sp|P51611|HFC1_MESAU  Host cell factor C1 (HCF) (VP16 accessory protein) (HFC1)
(VCAF) (CFF)
 gi|644882|dbj|BAA08258.1|  (D45419) HCF [Mesocricetus auratus]
          Length = 2090

 Score = 47.0 bits (110), Expect = 4e-05
 Identities = 32/97 (32%), Positives = 48/97 (48%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SH+AV Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMS 225




 Score = 44.7 bits (104), Expect = 2e-04
 Identities = 29/90 (32%), Positives = 41/90 (45%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR  HTA    E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHTAVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIETLTW 242

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272
```

```
 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
            L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 232 LWTLDIETLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
              EKEW        + LD           +  NIPR RA H AVA  +++   + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
            +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
               A++   V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100



>gi|20348947|ref|XP_109547.1|  (XM_109547) similar to transcription factor C1 (HCF) [Mus
musculus]
          Length = 2047

 Score = 46.6 bits (109), Expect = 5e-05
 Identities = 32/97 (32%), Positives = 47/97 (47%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
            G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 131 GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 188

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
                YG +P PR SHS V Y      KS+++ +GG+S
Sbjct: 189 IPITYGVLPPPRESHSVVVYTEKDNKKSKLVIYGGMS 225





 Score = 41.6 bits (96), Expect = 0.002
 Identities = 27/90 (30%), Positives = 40/90 (44%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
            W IP T G  P PR  H+     E +     ++I+GG    R      D++   I    W
Sbjct: 187 WDIPITYGVLPPPRESHSVVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIETLTW 242
```

```
Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 243 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 272




 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 232 LWTLDIETLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 284

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 285 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 339




 Score = 37.4 bits (85), Expect = 0.031
 Identities = 22/81 (27%), Positives = 42/81 (51%)
 Frame = +1

Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIP 267
           +G  P PR  H A  ++E+  I++FGG ++       +++  Y     +W    + G+IP
Sbjct: 26  SGPVPRPRHGHRAVAIKEL--IVVFGGGNEG----IVDELHVYNTATNQWFIPAVRGDIP 79

Query: 268 RPRASHSAVAYKSQVLYFGGV 330
              A++  V   +++L FGG+
Sbjct: 80  PGCAAYGFVCDGTRLLVFGGM 100



>gi|20260248|gb|AAM13022.1|  (AY093023) unknown protein [Arabidopsis thaliana]
          Length = 496

 Score = 46.6 bits (109), Expect = 5e-05
 Identities = 23/86 (26%), Positives = 43/86 (49%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLD 249
           W    + G  P PR    A+ +    +++  GG D  +   + +D++       EW++L
Sbjct: 114 WSELTSFGDLPTPRDFAAAAAIGSQKIVLC-GGWDGKK---WLSDVYVMDTMSLEWLELS 169

Query: 250 IYGNIPRPRASHSAVAYKSQVLYFGG 327
           + G++P PR  H+A   + ++L FGG
Sbjct: 170 VSGSLPPPRCGHTATMVEKRLLVFGG 195




 Score = 45.8 bits (107), Expect = 9e-05
 Identities = 29/88 (32%), Positives = 44/88 (49%), Gaps = 6/88 (6%)
 Frame = +1
```

```
Query: 88  TGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNI- 264
           +G PP  R  HTA  V + +++++FGG  D +   + +DI  Y I  K W + +  G+
Sbjct: 13  SGTPPQARSGHTAVNVGK-SMVVVFGGLVDKK---FLSDIIVYDIENKLWFEPECTGSES 68

Query: 265 -----PRPRASHSAVAYKSQVLYFGGVS 333
                P PRA H A+     +  FGG S
Sbjct: 69  EGQVGPTPRAFHVAITIDCHMFIFGGRS 96



 Score = 38.9 bits (89), Expect = 0.011
 Identities = 28/99 (28%), Positives = 43/99 (43%), Gaps = 4/99 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDD----SRSNPYFNDIFAYKIFEKEW 237
           W   +  G+ P  R  HT +     + +++FGG       SR + Y+ND         +W
Sbjct: 222 WTQLKLPGQAPSSRCGHTVT--SGGHYLLLFGGHGTGGWLSRYDVYYNDTIILDRVTAQW 279

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYLEF 354
            +L I    P PRA H+     ++ L  GG   +G L F
Sbjct: 280 KRLPIGNEPPPPRAYHTMTCIGARHLLIGG--FDGKLTF 316



>gi|6651060|gb|AAF22156.1|AF133093_11  (AF133093) host cell factor C1 [Mus musculus]
          Length = 1979

 Score = 46.6 bits (109), Expect = 5e-05
 Identities = 32/97 (32%), Positives = 47/97 (47%), Gaps = 16/97 (16%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNP------YFNDIFAYKIFEKE----WV 240
           G PP PR+ H+ S V   N   +FGG  +   +P      Y ND++  ++        W
Sbjct: 65  GPPPCPRLGHSFSLVG--NKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWD 122

Query: 241 QLDIYGNIPRPRASHSAVAY------KSQVLYFGGVS 333
               YG +P PR SHS V Y      KS+++ +GG+S
Sbjct: 123 IPITYGVLPPPRESHSVVVYTEKDNKKSKLVIYGGMS 159



 Score = 41.6 bits (96), Expect = 0.002
 Identities = 27/90 (30%), Positives = 40/90 (44%), Gaps = 4/90 (4%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMN----VIIIFGGRDDSRSNPYFNDIFAYKIFEKEW 237
           W IP T G  P PR   H+     E +     ++I+GG    R      D++   I    W
Sbjct: 121 WDIPITYGVLPPPRESHSVVVYTEKDNKKSKLVIYGGMSGCR----LGDLWTLDIETLTW 176

Query: 238 VQLDIYGNIPRPRASHSAVAYKSQVLYFGG 327
            +  + G  P PR+ HSA    +++  FGG
Sbjct: 177 NKPSLSGVAPLPRSLHSATTIGNKMYVFGG 206
```

```
 Score = 39.7 bits (91), Expect = 0.006
 Identities = 35/115 (30%), Positives = 47/115 (40%), Gaps = 19/115 (16%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           L  LD     W  P  +G  PLPR  H+A+ +   N + +FGG       P   D
Sbjct: 166 LWTLDIETLTWNKPSLSGVAPLPRSLHSATTIG--NKMYVFGGWV-----PLVMDDVKVA 218

Query: 220 IFEKEW--------VQLD-----------IYGNIPRPRASHSAVAYKSQVLYFGG 327
             EKEW        + LD           +  NIPR RA H AVA  +++  + G
Sbjct: 219 THEKEWKCTNTLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSG 273



>gi|7290955|gb|AAF46395.1| ☐ (AE003444) CG12081 gene product [Drosophila melanogaster]
          Length = 403

 Score = 46.2 bits (108), Expect = 7e-05
 Identities = 34/100 (34%), Positives = 44/100 (44%), Gaps = 6/100 (6%)
 Frame = +1

Query: 70  WVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDD------SRSNPYFNDIFAYKIFEK 231
           W   +T G PP  R  H AS   E   + IFGGR D      S+   Y ++I    +  K
Sbjct: 167 WRYVQTFGVPPSYRDFH-ASVAYEQERMYIFGGRGDKHSPYHSQEETYCHEIVYLDMKTK 225

Query: 232 EWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVSMNGYLE 351
            W +    G +P  R SHS   Y   +  FGG   NG L+
Sbjct: 226 VWHRPFTAGKVPVGRRSHSMFVYNKLIYVFGG--YNGLLD 263




 Score = 45.1 bits (105), Expect = 1e-04
 Identities = 27/98 (27%), Positives = 45/98 (45%)
 Frame = +1

Query: 40  LLKLDQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYK 219
           ++ LD     W  P T GK P+ R  H+     ++  I +FGG +    + +FND++ +
Sbjct: 217 IVYLDMKTKVWHRPFTAGKVPVGRRSHSMFVYNKL--IYVFGGYN-GLLDQHFNDLYTFD 273

Query: 220 IFEKEWVQLDIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
              K W  +   G  P  R    A+   +++  FGG S
Sbjct: 274 PRTKLWNLIRANGKAPTARRRQCAIVMGTRMFLFGGTS 311




 Score = 45.1 bits (105), Expect = 1e-04
 Identities = 30/93 (32%), Positives = 46/93 (49%), Gaps = 1/93 (1%)
 Frame = +1

Query: 52  DQPNNPWVIPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEK 231
           D     W  P+ TG  P  R  H+A  +   N + IFGG  D   N + +D+ +  +
Sbjct: 109 DPKTAQWSRPQVTGCLPGARDGHSACVIG--NSMYIFGGFVDE-INEFSSDVHSLNLDTM 165
```

```
Query: 232 EWVQLDIYGNIPRPRASHSAVAYKSQVLY-FGG 327
           EW  +  +G  P  R  H++VAY+ + +Y FGG
Sbjct: 166 EWRYVQTFGVPPSYRDFHASVAYEQERMYIFGG 198



 Score = 34.3 bits (77), Expect = 0.26
 Identities = 22/76 (28%), Positives = 34/76 (43%)
 Frame = +1

Query: 100 PLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPRPRA 279
           P  R  HT   V   + I I+GGR+D       N ++ +     +W +  + G +P  R
Sbjct: 75  PFQRYGHTV--VAYKDRIYIWGGRNDEN---LCNTLYCFDPKTAQWSRPQVTGCLPGARD 129

Query: 280 SHSAVAYKSQVLYFGG 327
            HSA    + +  FGG
Sbjct: 130 GHSACVIGNSMYIFGG 145



>gi|3811109|gb|AAC69437.1|  (U88869) protein serine/threonine phosphatase alpha [Plasmodium
falciparum]
          Length = 875

 Score = 46.2 bits (108), Expect = 7e-05
 Identities = 28/87 (32%), Positives = 44/87 (50%), Gaps = 1/87 (1%)
 Frame = +1

Query: 91  GKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQLDIYGNIPR 270
           G  P PR  HTA+Y+    V I      D+      +DI+ Y + + +W +L I  N P
Sbjct: 19  GDIPAPRFGHTATYLGNNKVAIFGRAIGDAGKYNITDDIYLYDLTQNKWKKL-ITENTPS 77

Query: 271 PRASHSAVAY-KSQVLYFGGVSMNGYL 348
            RA+H+A    + Q++ +GG +  G L
Sbjct: 78  ARAAHAAACVDEQQLVIYGGATGGGSL 104



 Score = 35.0 bits (79), Expect = 0.15
 Identities = 28/89 (31%), Positives = 39/89 (43%), Gaps = 1/89 (1%)
 Frame = +1

Query: 70  WV-IPETTGKPPLPRIQHTASYVQEMNVIIIFGGRDDSRSNPYFNDIFAYKIFEKEWVQL 246
           WV  P   G PP  R QHT+ ++   + I I GGR+D+      +    Y     EWV L
Sbjct: 235 WVEAPIKKGSPPEARYQHTSVFIG--SKIFILGGRNDNGCAVPLSTAL-YNTETIEWVTL 291

Query: 247 DIYGNIPRPRASHSAVAYKSQVLYFGGVS 333
                    +  H++  YK  +  FGG S
Sbjct: 292 PSIS-----KFRHTSWVYKYTIYTFGGFS 315



  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
```

```
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 222,546,477
Number of Sequences: 919285
Number of extensions: 4876041
Number of successful extensions: 16952
Number of sequences better than 10.0: 337
Number of HSP's better than 10.0 without gapping: 15995
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 16781
length of database: 288,558,979
effective HSP length: 94
effective length of database: 202,146,189
effective search space used: 4851508536
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020956767-011056-28129
```

7.1.17 Query= **hy-17_lf1-1a**

```
         (735 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           919,285 sequences; 288,558,979 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 200 Blast Hits on the Query Sequence

```
                                                                   Score    E
Sequences producing significant alignments:                        (bits) Value

gi|1360628|emb|CAA66648.1|  (X98016) histone H4-2 [Trichomon...     57   6e-13
gi|3023903|sp|P90516|H42_BLEJA  HISTONE H4 >gi|1708779|emb|C...     59   3e-11
gi|2851582|sp|P80737|H41_BLEJA  HISTONE H4-1                        57   7e-11
gi|1708778|emb|CAA66634.1|  (X97995) Histone H4 [Blepharisma...     57   7e-11
gi|70777|pir||HSTE42  histone H4, minor - Tetrahymena pyrifo...     65   7e-10
gi|122093|sp|P02310|H41_TETPY  HISTONE H4, MAJOR >gi|70778|p...     65   7e-10
gi|122095|sp|P02311|H42_TETPY  HISTONE H4, MINOR >gi|84352|p...     65   7e-10
gi|15430617|dbj|BAB64430.1|  (AB070962) histone H4 [Parameci...     64   2e-09
gi|122105|sp|P08436|H4_VOLCA  Histone H4 >gi|81288|pir||S009...     63   2e-09
gi|122094|sp|P23751|H42_EMENI  HISTONE H4.2 >gi|101786|pir||...     62   4e-09
gi|729679|sp|P40287|H4_ENTHI  HISTONE H4 >gi|1078772|pir||S5...     62   6e-09
gi|102290|pir||S10076  histone H4.2 - slime mold (Physarum p...     62   6e-09
gi|122101|sp|P04915|H4_PHYPO  HISTONE H4 >gi|84161|pir||A278...     62   6e-09
gi|102411|pir||S14185  histone H4 (clone H4g) - Stylonychia ...     61   9e-09
gi|663070|emb|CAA54829.1|  (X77806) histone H4 [Pyrenomonas ...     61   9e-09
gi|122100|sp|P18836|H4_OXYNO  HISTONE H4 >gi|84369|pir||JS01...     61   9e-09
gi|12854993|dbj|BAB30189.1|  (AK016310) data source:MGD, sou...     61   1e-08
gi|10953803|gb|AAG25601.1|AF297469_1  (AF297469) histone H4 ...     61   1e-08
gi|223582|prf||0901261A  histone H4 [Homo sapiens]                  61   1e-08
gi|1708110|sp|P50566|H4_CHLRE  HISTONE H4 >gi|2119025|pir||S...     61   1e-08
gi|17352498|ref|NP_478143.1|  (NM_058236) histone H4; Bicoid...     61   1e-08
gi|3745763|pdb|1AOI|F  Chain F, X-Ray Structure Of The Nucle...     61   1e-08
gi|70762|pir||HSBO4  histone H4 - bovine >gi|70763|pir||HSPG...     61   1e-08
```

```
gi|12845285|dbj|BAB26692.1|  (AK010085) data source:MGD, sou...    61   1e-08
gi|16797797|dbj|BAB71814.1|  (AB050889) histone H4 [Citrus j...    61   1e-08
gi|70774|pir||HSWT41  histone H4 (TH091) - wheat >gi|170747|...    61   1e-08
gi|15226944|ref|NP_180441.1|  (NM_128434) histone H4 [Arabid...    61   1e-08
gi|12847763|dbj|BAB27698.1|  (AK011560) data source:MGD, sou...    61   1e-08
gi|4504301|ref|NP_003529.1|  (NM_003538) H4 histone family, ...    61   1e-08
gi|462242|sp|P35059|H4_ACRFO  HISTONE H4 >gi|166311|gb|AAC37...    61   1e-08
gi|1806283|emb|CAB01913.1|  (Z79637) Histone H4 homologue [S...    61   1e-08
gi|1199967|emb|CAA64985.1|  (X95689) histone H4 [Allium cepa]      61   1e-08
gi|4139870|pdb|1HIO|D  Chain D, Histone Octamer (Chicken), C...    61   1e-08
gi|224293|prf||1101277A  histone H4 [Triticum aestivum]            61   1e-08
gi|70772|pir||HSWT4  histone H4 - wheat                            61   1e-08
gi|1883028|emb|CAA62811.1|  (X91511) histone H4 [Diprion pini]     61   1e-08
gi|19569887|gb|AAL92205.1|AC115576_7  (AC115576) hypothetica...    61   1e-08
gi|122099|sp|P04914|H4_NEUCR  Histone H4 >gi|83768|pir||S079...    60   2e-08
gi|122092|sp|P23750|H41_EMENI  HISTONE H4.1 >gi|101785|pir||...    60   2e-08
gi|223793|prf||0912198A  histone H4 [Physarum polycephalum]        60   2e-08
gi|515377|emb|CAA56154.1|  (X79715) histone H4 [Lolium temul...    60   2e-08
gi|484442|pir||JN0688  histone H4 - sea squirt (Styela plica...    60   3e-08
gi|462243|sp|P35057|H4_LYCES  HISTONE H4 >gi|421921|pir||S32...    60   3e-08
gi|2137382|pir||I48404  histone H4 (55AA) (1 is 3rd base in ...    60   3e-08
gi|20141289|sp|P80739|H4_EUPCR  Histone H4 >gi|1763310|gb|AA...    60   3e-08
gi|6006733|gb|AAF00589.1|AF140034_1  (AF140034) histone H4 [...    60   3e-08
gi|2062367|gb|AAB53361.1|  (U65675) histone H4 [Plasmodium f...    59   4e-08
gi|13812130|ref|NP_113257.1|  (NC_002752) Histone H4 [Guilla...    59   4e-08
gi|90626|pir||S03427  histone H4 (clone 53) - mouse >gi|5131...    59   5e-08
gi|886740|emb|CAA59110.1|  (X84376) histone 4 [Zea mays]           59   5e-08
gi|11121217|emb|CAC14793.1|  (AJ249812) histone H4 [Mortiere...    59   5e-08
gi|32097|emb|CAA24918.1|  (X00038) histone H4 [Homo sapiens]       59   6e-08
gi|1883034|emb|CAA62810.1|  (X91510) histone H4 [Diadromus p...    59   6e-08
gi|17540668|ref|NP_501203.1|  (NM_068802) Histone [Caenorhab...    59   6e-08
gi|70768|pir||HSUR4  histone H4 - sea urchin (Psammechinus m...    59   6e-08
gi|15988133|pdb|1ID3|B  Chain B, Crystal Structure Of The Ye...    59   6e-08
gi|7510983|pir||T27741  hypothetical protein ZK131.4 - Caeno...    59   6e-08
gi|84451|pir||JS0314  histone H4 - Caenorhabditis elegans >g...    59   6e-08
gi|3287225|emb|CAA76306.1|  (Y16587) histone H4 [Paracentrot...    59   6e-08
gi|17509199|ref|NP_492641.1|  (NM_060240) histone H4 [Caenor...    59   6e-08
gi|6319481|ref|NP_009563.1|  (NC_001134) Histone H4 (HHF1 an...    59   6e-08
gi|19880139|gb|AAM00266.1|AF361948_1  (AF361948) histone 4 [...    59   6e-08
gi|462244|sp|P35058|H4_PHACH  Histone H4 >gi|101931|pir||S25...    58   8e-08
gi|1885348|emb|CAA62815.1|  (X91515) histone H4 [Trichogramm...    58   1e-07
gi|6980692|pdb|2HIO|D  Chain D, Histone Octamer (Chicken), C...    58   1e-07
gi|2495141|sp|Q27443|H4_ASCSU  HISTONE H4 >gi|1177238|emb|CA...    58   1e-07
gi|4504319|ref|NP_003538.1|  (NM_003547) H4 histone family, ...    45   1e-07
gi|122104|sp|P27996|H4_SOLST  HISTONE H4 >gi|103584|pir||S20...    57   1e-07
gi|1883001|emb|CAA62809.1|  (X91509) histone H4 [Apis mellif...    57   1e-07
gi|1883026|emb|CAA62813.1|  (X91513) histone H4 [Diprion pini]     57   1e-07
gi|19112350|ref|NP_595558.1|  (NC_003423) histone H4 [Schizo...    57   2e-07
gi|6006741|gb|AAF00593.1|AF139876_1  (AF139876) histone H4 [...    55   5e-07
gi|1731927|emb|CAA71084.1|  (Y09953) histone H4 [Anopheles g...    53   3e-06
gi|1883030|emb|CAA62812.1|  (X91512) histone H4 [Diprion pini]     52   7e-06
gi|4379376|emb|CAA06069.1|  (AJ004709) histone H4 [Protocruz...    51   1e-05
gi|2564209|emb|CAA75404.1|  (Y15132) histone H4 [Arbacia lix...    50   2e-05
gi|4376214|emb|CAA06065.1|  (AJ004705) histone H4 [Blepharis...    47   2e-04
gi|4376196|emb|CAA06063.1|  (AJ004703) histone H4 [Blepharis...    47   2e-04
gi|4377539|emb|CAA06068.1|  (AJ004708) histone H4 [Euplotes ...    45   5e-04
gi|4376213|emb|CAA06064.1|  (AJ004704) histone H4 [Blepharis...    45   5e-04
gi|4377541|emb|CAA06067.1|  (AJ004707) histone H4 [Euplotes ...    45   5e-04
gi|4377535|emb|CAA06071.1|  (AJ004711) histone H4 [Euplotes ...    44   0.001
gi|4377536|emb|CAA06072.1|  (AJ004712) histone H4 [Euplotes ...    44   0.001
gi|18677172|gb|AAL78218.1|AF344869_1  (AF344869) histone Hgg...    44   0.002
```

```
gi|84353|pir||A02650  histone H4 - Tetrahymena thermophila (...    35   0.032
gi|2222684|emb|CAA74210.1|  (Y13915) Histone H4 [Leishmania ...    39   0.050
gi|5738233|gb|AAD50306.1|AF175386_1  (AF175386) histone H4 [...    39   0.050
gi|2222686|emb|CAA74211.1|  (Y13916) Histone H4 [Leishmania ...    39   0.050
gi|11022579|emb|CAC14237.1|  (AL445678) histone H4 [Leishman...    39   0.050
gi|2995265|emb|CAA06074.1|  (AJ004713) histone H4 [Prorodon ...    38   0.085
gi|2995229|emb|CAA06054.1|  (AJ004696) histone H4 [Obertrumi...    38   0.11
gi|2995268|emb|CAA06076.1|  (AJ004714) histone H4 [Prorodon ...    37   0.15
gi|4379375|emb|CAA06061.1|  (AJ004701) histone H4 [Protocruz...    37   0.25
gi|19173165|ref|NP_596968.1|  (NC_003238) HISTONE H4 [Enceph...    31   0.25
gi|2995235|emb|CAA06058.1|  (AJ004698) histone H4 [Colpoda c...    36   0.32
gi|2995217|emb|CAA06046.1|  (AJ004692) histone H4 [Colpidium...    36   0.32
gi|2995226|emb|CAA06052.1|  (AJ004695) histone H4 [Obertrumi...    36   0.32
gi|2995232|emb|CAA06056.1|  (AJ004697) histone H4 [Obertrumi...    36   0.42
gi|2995208|emb|CAA06040.1|  (AJ004689) histone H4 [Blepharis...    36   0.42
gi|18581195|ref|XP_101250.1|  (XM_101250) hypothetical prote...    33   2.1
```

Alignments

```
>gi|1360628|emb|CAA66648.1|  (X98016) histone H4-2 [Trichomonas vaginalis]
 gi|1360630|emb|CAA66649.1|  (X98017) histone H4-3 [Trichomonas vaginalis]
          Length = 101

 Score = 57.0 bits (136), Expect(2) = 6e-13
 Identities = 24/48 (50%), Positives = 38/48 (79%)
 Frame = -2

Query: 347 LFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           +++  R V+ +FL ++++D++ YTEHA+R TV AMDVVYALK+ G+ L
Sbjct: 51  IYEEVRGVLRHFLENVIRDSVTYTEHARRKTVTAMDVVYALKRQGKTL 98




 Score = 38.5 bits (88), Expect(2) = 6e-13
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS D Y   R
Sbjct: 28  QGITKPAIRRLARRGGVKRISGDIYEEVR 56




>gi|3023903|sp|P90516|H42_BLEJA  HISTONE H4
 gi|1708779|emb|CAA66635.1|  (X97996) Histone H4 [Blepharisma japonicum]
          Length = 89

 Score = 58.5 bits (140), Expect(2) = 3e-11
 Identities = 26/46 (56%), Positives = 37/46 (79%)
 Frame = -2

Query: 350 VLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYG 213
           +++D  R+V+  FL S+V+DA+ YTEHA+R TV A+DVVYALK+ G
Sbjct: 44  LVYDETRNVLKVFLESVVRDAVTYTEHAKRETVTALDVVYALKRQG 89
```

```
 Score = 31.2 bits (69), Expect(2) = 3e-11
 Identities = 19/30 (63%), Positives = 20/30 (66%), Gaps = 1/30 (3%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKR-ISSDSYPTTR 355
           Q IT   IRRLARRGGVKR +S   Y  TR
Sbjct: 21  QGITKPAIRRLARRGGVKRTLSGLVYDETR 50



>gi|2851582|sp|P80737|H41_BLEJA  HISTONE H4-1
          Length = 97

 Score = 57.4 bits (137), Expect(2) = 7e-11
 Identities = 25/46 (54%), Positives = 37/46 (80%)
 Frame = -2

Query: 350 VLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYG 213
           +++D  R+V+  FL ++V+DA+ YTEHA+R TV A+DVVYALK+ G
Sbjct: 52  LVYDETRNVLKVFLENVVRDAVTYTEHARRKTVTALDVVYALKRQG 97




 Score = 31.2 bits (69), Expect(2) = 7e-11
 Identities = 19/30 (63%), Positives = 20/30 (66%), Gaps = 1/30 (3%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKR-ISSDSYPTTR 355
           Q IT   IRRLARRGGVKR +S   Y  TR
Sbjct: 29  QGITKPAIRRLARRGGVKRTLSGLVYDETR 58



>gi|1708778|emb|CAA66634.1|  (X97995) Histone H4 [Blepharisma japonicum]
          Length = 89

 Score = 57.4 bits (137), Expect(2) = 7e-11
 Identities = 25/46 (54%), Positives = 37/46 (80%)
 Frame = -2

Query: 350 VLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYG 213
           +++D  R+V+  FL ++V+DA+ YTEHA+R TV A+DVVYALK+ G
Sbjct: 44  LVYDETRNVLKVFLENVVRDAVTYTEHARRKTVTALDVVYALKRQG 89




 Score = 31.2 bits (69), Expect(2) = 7e-11
 Identities = 19/30 (63%), Positives = 20/30 (66%), Gaps = 1/30 (3%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKR-ISSDSYPTTR 355
           Q IT   IRRLARRGGVKR +S   Y  TR
Sbjct: 21  QGITKPAIRRLARRGGVKRTLSGLVYDETR 50
```

```
>gi|70777|pir||HSTE42  histone H4, minor - Tetrahymena pyriformis
 gi|223273|prf||0702236B  histone H4 [Tetrahymena sp.]
          Length = 102

 Score = 65.1 bits (157), Expect = 7e-10
 Identities = 29/53 (54%), Positives = 41/53 (76%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++D +R V+ +FL ++V+DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISSFIYDDSRQVLKSFLENVVRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 97



 Score = 36.2 bits (82), Expect = 0.32
 Identities = 23/51 (45%), Positives = 26/51 (50%)
 Frame = -1

Query: 507 GQGKLAGAXXXXXXXXXXXKGGQHITNGDIRRLARRGGVKRISSDSYPTTR 355
           G GK+              +G   IT   IRRLARRGGVKRISS  Y  +R
Sbjct: 8   GMGKVGAKRHSRKSNKASIEG---ITKPAIRRLARRGGVKRISSFIYDDSR 55



>gi|122093|sp|P02310|H41_TETPY  HISTONE H4, MAJOR
 gi|70778|pir||HSTE41  histone H4, major - Tetrahymena pyriformis
 gi|352679|prf||1011244A  histone H4 [Tetrahymena pyriformis]
          Length = 102

 Score = 65.1 bits (157), Expect = 7e-10
 Identities = 29/53 (54%), Positives = 41/53 (76%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++D +R V+ +FL ++V+DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISSFIYDDSRQVLKSFLENVVRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 97



 Score = 36.2 bits (82), Expect = 0.32
 Identities = 23/51 (45%), Positives = 26/51 (50%)
 Frame = -1

Query: 507 GQGKLAGAXXXXXXXXXXXKGGQHITNGDIRRLARRGGVKRISSDSYPTTR 355
           G GK+              +G   IT   IRRLARRGGVKRISS  Y  +R
Sbjct: 8   GMGKVGAKRHSKRSNKASIEG---ITKPAIRRLARRGGVKRISSFIYDDSR 55



>gi|122095|sp|P02311|H42_TETPY  HISTONE H4, MINOR
 gi|84352|pir||A25875  histone H4 - Tetrahymena thermophila
```

```
 gi|10815|emb|CAA25121.1|  (X00417) H4-I histone [Tetrahymena thermophila]
 gi|10817|emb|CAA28452.1|  (X04755) histone H4 II (AA 1-103) [Tetrahymena thermophila]
          Length = 103

 Score = 65.1 bits (157), Expect = 7e-10
 Identities = 29/53 (54%), Positives = 41/53 (76%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++D +R V+ +FL ++V+DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISSFIYDDSRQVLKSFLENVVRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98




 Score = 36.2 bits (82), Expect = 0.32
 Identities = 23/51 (45%), Positives = 26/51 (50%)
 Frame = -1

Query: 507 GQGKLAGAXXXXXXXXXXXKGGQHITNGDIRRLARRGGVKRISSDSYPTTR 355
           G GK+            +G   IT   IRRLARRGGVKRISS  Y  +R
Sbjct: 9   GMGKVGAKRHSRKSNKASIEG---ITKPAIRRLARRGGVKRISSFIYDDSR 56




>gi|15430617|dbj|BAB64430.1|  (AB070962) histone H4 [Paramecium caudatum]
          Length = 102

 Score = 63.5 bits (153), Expect = 2e-09
 Identities = 28/53 (52%), Positives = 41/53 (76%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   L+D +R+V+ +FL ++++DAI YTEHA+R TV  MDVVY+LK+ GR L
Sbjct: 45  RISSFLYDDSRNVLKSFLENVIRDAITYTEHARRKTVTVMDVVYSLKRQGRTL 97




 Score = 35.8 bits (81), Expect = 0.42
 Identities = 19/27 (70%), Positives = 20/27 (73%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           IT   IRRLARRGGVKRISS  Y  +R
Sbjct: 29  ITKPAIRRLARRGGVKRISSFLYDDSR 55




>gi|122105|sp|P08436|H4_VOLCA  Histone H4
 gi|81288|pir||S00939  histone H4 - Volvox carteri
 gi|21984|emb|CAA30034.1|  (X06963) put. histone H4 [Volvox carteri]
 gi|21987|emb|CAA30036.1|  (X06964) put. histone H4 [Volvox carteri]
          Length = 103

 Score = 63.2 bits (152), Expect = 2e-09
 Identities = 27/53 (50%), Positives = 41/53 (76%)
```

```
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+ NFL ++++D++ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRTVLKNFLENVIRDSVTYTEHARRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y   TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|122094|sp|P23751|H42_EMENI  HISTONE H4.2
 gi|101786|pir||S11940  histone H4.2 - Emericella nidulans
 gi|296339|emb|CAA39156.1|  (X55550) histone H4.2 [Emericella nidulans]
 gi|529957|gb|AAA20821.1|  (U12631) histone H4.2 [Emericella nidulans]
 gi|227598|prf||1707275D  histone H4.2 [Emericella nidulans]
          Length = 103

 Score = 62.4 bits (150), Expect = 4e-09
 Identities = 26/53 (49%), Positives = 41/53 (77%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+ +FL S+++DA+ YTEHA+R TV ++DVVYALK+ GR L
Sbjct: 46  RISAMIYEETRGVLKSFLESVIRDAVTYTEHAKRKTVTSLDVVYALKRQGRTL 98



 Score = 37.4 bits (85), Expect = 0.15
 Identities = 20/29 (68%), Positives = 21/29 (71%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS+   Y   TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISAMIYEETR 56



>gi|729679|sp|P40287|H4_ENTHI  HISTONE H4
 gi|1078772|pir||S52264  histone H4 - Entamoeba histolytica
 gi|7439817|pir||S52262  histone H4 - Entamoeba histolytica
 gi|532583|gb|AAB67323.1|  (L35898) histone H4 [Entamoeba histolytica]
 gi|642232|emb|CAA58833.1|  (X84010) histone H4 [Entamoeba histolytica]
          Length = 118

 Score = 62.0 bits (149), Expect = 6e-09
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2
```

```
Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++D  R+V+  FL  +++D++ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 62  RINGAVYDETRNVLKQFLEQVIRDSVTYTEHAKRRTVTAMDVVYALKRQGRTL 114



 Score = 35.0 bits (79), Expect = 0.72
 Identities = 18/27 (66%), Positives = 19/27 (69%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           IT   IRRLARRGGVKRI+   Y  TR
Sbjct: 46  ITKPAIRRLARRGGVKRINGAVYDETR 72



>gi|102290|pir||S10076  histone H4.2 - slime mold (Physarum polycephalum)
 gi|1183935|emb|CAA33239.1|  (X15141) histone H42 [Physarum polycephalum]
          Length = 103

 Score = 62.0 bits (149), Expect = 6e-09
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   +++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISKTIYEETRGVLKTFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98



 Score = 37.7 bits (86), Expect = 0.11
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISKTIYEETR 56



>gi|122101|sp|P04915|H4_PHYPO  HISTONE H4
 gi|84161|pir||A27859  histone H4.1 - slime mold (Physarum polycephalum)
 gi|3196|emb|CAA25140.1|  (X00449) histone H4 [Physarum polycephalum]
 gi|3208|emb|CAA68442.1|  (Y00366) histone H4 (H42) [Physarum polycephalum]
 gi|295931|emb|CAA33240.1|  (X15142) H41 [Physarum polycephalum]
          Length = 103

 Score = 62.0 bits (149), Expect = 6e-09
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   +++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISNTIYEETRGVLKTFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98
```

```
 Score = 38.1 bits (87), Expect = 0.085
 Identities = 20/29 (68%), Positives = 21/29 (71%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS+  Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISNTIYEETR 56



>gi|102411|pir||S14185  histone H4 (clone H4g) - Stylonychia lemnae
          Length = 145

 Score = 61.2 bits (147), Expect = 9e-09
 Identities = 25/53 (47%), Positives = 42/53 (79%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R+V+ +FL ++++D++ YTEHA+R TV A+DVVYALK+ GR L
Sbjct: 88  RISSLIYEETRNVLRSFLENVIRDSVTYTEHAKRKTVTALDVVYALKRQGRTL 140



 Score = 37.0 bits (84), Expect = 0.19
 Identities = 20/27 (74%), Positives = 20/27 (74%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           IT   IRRLARRGGVKRISS  Y  TR
Sbjct: 72  ITKPAIRRLARRGGVKRISSLIYEETR 98



>gi|663070|emb|CAA54829.1|  (X77806) histone H4 [Pyrenomonas salina]
          Length = 103

 Score = 61.2 bits (147), Expect = 9e-09
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRSVLKVFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1
```

```
Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56




>gi|122100|sp|P18836|H4_OXYNO  HISTONE H4
 gi|84369|pir||JS0154  histone H4 - Oxytricha nova
 gi|102410|pir||S14184  histone H4 (clone H4K) - Stylonychia lemnae
 gi|10142|emb|CAA34151.1|  (X16018) histone H4 [Stylonychia lemnae]
 gi|159806|gb|AAA29395.1|  (M24411) H4 histone [Oxytricha nova]
 gi|974208|emb|CAA34152.1|  (X16019) histone H4 [Stylonychia lemnae]
 gi|6851293|gb|AAF29507.1|  (AF192970) histone H4 [Oxytricha trifallax]
          Length = 104

 Score = 61.2 bits (147), Expect = 9e-09
 Identities = 25/53 (47%), Positives = 42/53 (79%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R+V+ +FL ++++D++ YTEHA+R TV A+DVVYALK+ GR L
Sbjct: 47  RISSLIYEETRNVLRSFLENVIRDSVTYTEHAKRKTVTALDVVYALKRQGRTL 99




 Score = 37.0 bits (84), Expect = 0.19
 Identities = 20/27 (74%), Positives = 20/27 (74%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           IT   IRRLARRGGVKRISS  Y  TR
Sbjct: 31  ITKPAIRRLARRGGVKRISSLIYEETR 57




>gi|12854993|dbj|BAB30189.1|    (AK016310) data source:MGD, source key:MGI:96099,
           evidence:ISS~histone 4 protein~putative [Mus musculus]
          Length = 127

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 70  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 122




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 52  QGITKPAIRRLARRGGVKRISGLIYEETR 80
```

```
>gi|10953803|gb|AAG25601.1|AF297469_1  (AF297469) histone H4 [Schistosoma mansoni]
          Length = 98

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 44  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 96



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 26  QGITKPAIRRLARRGGVKRISGLIYEETR 54



>gi|223582|prf||0901261A  histone H4 [Homo sapiens]
          Length = 102

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 97



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 27  QGITKPAIRRLARRGGVKRISGLIYEETR 55



>gi|1708110|sp|P50566|H4_CHLRE  HISTONE H4
 gi|2119025|pir||S59586  histone H4 (clones CH-I, CH-II, and CH-III) - Chlamydomonas
reinhardtii
```

```
 gi|2119027|pir||S59124  histone H4 - Chlamydomonas reinhardtii
 gi|571471|gb|AAA98445.1|  (U16724) histone H4 [Chlamydomonas reinhardtii]
 gi|571476|gb|AAA98449.1|  (U16725) histone H4 [Chlamydomonas reinhardtii]
 gi|576634|gb|AAA98456.1|  (U16825) histone H4 [Chlamydomonas reinhardtii]
 gi|790701|gb|AAA99966.1|  (L41841) histone H4 [Chlamydomonas reinhardtii]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++D++ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRTVLKTFLENVIRDSVTYTEHARRKTVTAMDVVYALKRQGRTL 98




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|17352498|ref|NP_478143.1|  ☐ (NM_058236) histone H4; Bicoid interacting protein 2;
histone [Drosophila melanogaster]
 gi|17975542|ref|NP_524352.1|  ☐ (NM_079628) Histone H4 replacement [Drosophila
melanogaster]
 gi|122097|sp|P02307|H4_DROME  Histone H4
 gi|103010|pir||S18003  histone H4 - midge (Chironomus thummi thummi)
 gi|103052|pir||S09656  histone H4 - fruit fly (Drosophila hydei)
 gi|103199|pir||S10098  histone H4 - fruit fly (Drosophila melanogaster)
 gi|542539|pir||S40437  histone H4 - midge (Chironomus thummi thummi)
 gi|1085859|pir||B56654  histone H4 - Tigriopus californicus
 gi|2119026|pir||B56580  histone H4 - midge (Chironomus thummi thummi)
 gi|7428643|pir||HSFF4  histone H4 - fruit fly (Drosophila melanogaster)
 gi|7084|emb|CAA39772.1|  (X56335) histone H4 [Chironomus thummi]
 gi|7434|emb|CAA36806.1|  (X52576) histone H4 [Drosophila hydei]
 gi|7440|emb|CAA34920.1|  (X17072) histone H4 [Drosophila hydei]
 gi|8071|emb|CAA32435.1|  ☐ (X14215) H4 histone [Drosophila melanogaster]
 gi|10616|emb|CAA36639.1|  (X52393) histone H4 [Tigriopus californicus]
 gi|161896|gb|AAC41553.1|  (M84797) histone H4 [Tigriopus californicus]
 gi|297562|emb|CAA51323.1|  (X72803) histone H4 [Chironomus thummi]
 gi|1419478|emb|CAA66066.1|  (X97436) histone H4 [Drosophila hydei]
 gi|1419480|emb|CAA66067.1|  ☐ (X97437) histone H4 [Drosophila melanogaster]
 gi|1419482|emb|CAA66068.1|  ☐ (X97438) histone H4 [Drosophila melanogaster]
 gi|1882999|emb|CAA62808.1|  (X91508) histone H4 [Acrolepiopsis assectella]
 gi|1885324|emb|CAA62814.1|  (X91514) histone H4 [Myrmica ruginodis]
 gi|6686556|emb|CAB64686.1|  (AJ238321) putative H4 histone [Asellus aquaticus]
 gi|7299905|gb|AAF55080.1|  ☐ (AE003705) His4r gene product [Drosophila melanogaster]
 gi|14269423|gb|AAK58065.1|AF378198_5  (AF378198) histone H4 [Rhynchosciara americana]
          Length = 103
```

```
 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 98




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT    IRRLARRGGVKRIS    Y   TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|3745763|pdb|1AOI|F  Chain F, X-Ray Structure Of The Nucleosome Core Particle At 2.8 A
Resolution
 gi|3745759|pdb|1AOI|B  Chain B, X-Ray Structure Of The Nucleosome Core Particle At 2.8 A
Resolution
          Length = 87

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 30  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 82




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT    IRRLARRGGVKRIS    Y   TR
Sbjct: 12  QGITKPAIRRLARRGGVKRISGLIYEETR 40



>gi|70762|pir||HSBO4  histone H4 - bovine
 gi|70763|pir||HSPG4  histone H4 - pig
 gi|70765|pir||HSCH4  histone H4 - chicken
 gi|70766|pir||HSTR4  histone H4 - rainbow trout
          Length = 102

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2
```

```
Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 97




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 27  QGITKPAIRRLARRGGVKRISGLIYEETR 55



>gi|12845285|dbj|BAB26692.1|  (AK010085) data source:MGD, source key:MGI:96099,
evidence:ISS~histone 4 protein~putative [Mus musculus]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 98




 Score = 36.2 bits (82), Expect = 0.32
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPVIRRLARRGGVKRISGLIYEETR 56



>gi|16797797|dbj|BAB71814.1|  (AB050889) histone H4 [Citrus jambhiri]
          Length = 102

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98
```

```
 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|70774|pir||HSWT41  histone H4 (TH091) - wheat
 gi|170747|gb|AAA34292.1|  (M12277) histone H4 [Triticum aestivum]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|15226944|ref|NP_180441.1|  (NM_128434) histone H4 [Arabidopsis thaliana]
 gi|15231283|ref|NP_190179.1|  (NM_114462) Histone H4 - like protein [Arabidopsis thaliana]
 gi|15232318|ref|NP_190941.1|  (NM_115233) histone H4-like protein [Arabidopsis thaliana]
 gi|18390794|ref|NP_563793.1|  (NM_100639) histone H4 [Arabidopsis thaliana]
 gi|18390815|ref|NP_563797.1|  (NM_100656) histone H4 [Arabidopsis thaliana]
 gi|18424269|ref|NP_568911.1|  (NM_125361) histone H4 - like protein [Arabidopsis thaliana]
 gi|18424305|ref|NP_568918.1|  (NM_125390) histone H4 - like protein [Arabidopsis thaliana]
 gi|122106|sp|P02308|H4_WHEAT  HISTONE H4
 gi|70771|pir||HSZM4  histone H4 - maize
 gi|81642|pir||S06904  histone H4 - Arabidopsis thaliana
 gi|2119028|pir||S60475  histone H4 - garden pea
 gi|7428644|pir||HSPM4  histone H4 - garden pea
 gi|21795|emb|CAA24924.1|  (X00043) histone H4 [Triticum aestivum]
 gi|166740|gb|AAA32810.1|  (M17132) histone H4 [Arabidopsis thaliana]
 gi|166742|gb|AAA32811.1|  (M17133) histone H4 [Arabidopsis thaliana]
 gi|168499|gb|AAA33474.1|  (M36659) histone H4 (H4C13) [Zea mays]
 gi|168501|gb|AAA33475.1|  (M13370) histone H4 [Zea mays]
 gi|168503|gb|AAA33476.1|  (M13377) histone H4 [Zea mays]
 gi|498898|gb|AAA86948.1|  (U10042) histone H4 homolog [Pisum sativum]
 gi|1806285|emb|CAB01914.1|  (Z79638) histone H4 homologue [Sesbania rostrata]
 gi|3927823|gb|AAC79580.1|  (AC005727) histone H4 [Arabidopsis thaliana]
 gi|6009915|dbj|BAA85120.1|  (AB018245) histone H4-like protein [Solanum melongena]
```

```
 gi|6522611|emb|CAB62023.1|  (AL133298) histone H4-like protein [Arabidopsis thaliana]
 gi|7339494|emb|CAB82817.1|  (AL162459) Histone H4-like protein [Arabidopsis thaliana]
 gi|7629993|emb|CAB88335.1|  (AL132960) histone H4-like protein [Arabidopsis thaliana]
 gi|8439886|gb|AAF75072.1|AC007583_8  (AC007583) Identical to histone H4 from Arabidopsis
thaliana gi|S06904
 gi|8439903|gb|AAF75089.1|AC007583_25  (AC007583) Identical to histone H4 from Arabidopsis
thaliana gi|S06904
 gi|9757918|dbj|BAB08365.1|  (AB015475) histone H4 [Arabidopsis thaliana]
 gi|9758835|dbj|BAB09507.1|  (AB006705) histone H4 [Arabidopsis thaliana]
 gi|11762277|gb|AAG40410.1|AF325058_1  (AF325058) AT5g59690 [Arabidopsis thaliana]
 gi|12039318|gb|AAG46106.1|AC073166_4  (AC073166) histone H4 [Oryza sativa]
 gi|12248031|gb|AAG50107.1|AF334729_1  (AF334729) putative histone H4 protein [Arabidopsis
thaliana]
 gi|13277212|emb|CAC34411.1|  (Y18575) histone H4 [Flaveria trinervia]
 gi|16209693|gb|AAL14404.1|  (AY057609) AT5g59690/mth12_90 [Arabidopsis thaliana]
 gi|17065282|gb|AAL32795.1|  (AY062717) histone H4-like protein [Arabidopsis thaliana]
 gi|17380766|gb|AAL36213.1|  (AY063857) putative histone H4 protein [Arabidopsis thaliana]
 gi|20160804|dbj|BAB89744.1|  (AP003271) putative histone H4 [Oryza sativa (japonica
cultivar-group)]
 gi|20198175|gb|AAM15445.1|  (AC007184) histone H4 [Arabidopsis thaliana]
 gi|20260010|gb|AAM13352.1|  (AY093353) histone H4-like protein [Arabidopsis thaliana]
 gi|20466418|gb|AAM20526.1|  (AY099675) histone H4-like protein [Arabidopsis thaliana]
 gi|225838|prf||1314298A  histone H4 [Arabidopsis thaliana]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|12847763|dbj|BAB27698.1|  (AK011560) data source:MGD, source key:MGI:96099,
evidence:ISS~histone 4 protein~putative [Mus musculus]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRETVTAMDVVYALKRQGRTL 98
```

```
 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|4504301|ref|NP_003529.1| ☐ (NM_003538) H4 histone family, member A [Homo sapiens]
 gi|4504303|ref|NP_003530.1| ☐ (NM_003539) H4 histone family, member B [Homo sapiens]
 gi|4504305|ref|NP_003531.1| ☐ (NM_003540) H4 histone family, member C [Homo sapiens]
 gi|4504307|ref|NP_003532.1| ☐ (NM_003541) H4 histone family, member D [Homo sapiens]
 gi|4504309|ref|NP_003533.1| ☐ (NM_003542) H4 histone family, member G [Homo sapiens]
 gi|4504311|ref|NP_003534.1| ☐ (NM_003543) H4 histone family, member H [Homo sapiens]
 gi|4504313|ref|NP_003535.1| ☐ (NM_003544) H4 histone family, member I [Homo sapiens]
 gi|4504315|ref|NP_003536.1| ☐ (NM_003545) H4 histone family, member J [Homo sapiens]
 gi|4504317|ref|NP_003537.1| ☐ (NM_003546) H4 histone family, member K [Homo sapiens]
 gi|4504321|ref|NP_003486.1| ☐ (NM_003495) H4 histone family, member M; Histone 4 family,
member M [Homo sapiens]
 gi|4504323|ref|NP_003539.1| ☐ (NM_003548) H4 histone, family 2 [Homo sapiens]
 gi|7305141|ref|NP_038578.1| ☐ (NM_013550) histone 4 protein [Mus musculus]
 gi|11415030|ref|NP_068803.1| ☐ (NM_021968) H4 histone family, member E [Homo sapiens]
 gi|11418817|ref|XP_004408.1| ☐ (XM_004408) H4 histone family, member A [Homo sapiens]
 gi|11418837|ref|XP_004418.1| ☐ (XM_004418) H4 histone family, member B [Homo sapiens]
 gi|12083635|ref|NP_073177.1| ☐ (NM_022686) germinal histone H4 gene [Rattus norvegicus]
 gi|12732493|ref|XP_004424.2| ☐ (XM_004424) H4 histone family, member C [Homo sapiens]
 gi|14752061|ref|XP_030144.1| ☐ (XM_030144) H4 histone family, member G [Homo sapiens]
 gi|15304970|ref|XP_053631.1| ☐ (XM_053631) similar to HISTONE H4 [Homo sapiens]
 gi|16174327|ref|XP_057577.1| ☐ (XM_057577) similar to HISTONE H4 [Homo sapiens]
 gi|18564714|ref|XP_087934.1| ☐ (XM_087934) H4 histone family, member M [Homo sapiens]
 gi|18565224|ref|XP_094887.1| ☐ (XM_094887) similar to H4 histone family, member D (H.
sapiens) [Homo sapiens]
 gi|20345269|ref|XP_110041.1|  (XM_110041) similar to germinal histone H4 gene [Mus
musculus]
 gi|20345632|ref|XP_111473.1|  (XM_111473) similar to germinal histone H4 gene [Mus
musculus]
 gi|20345636|ref|XP_111476.1|  (XM_111476) similar to germinal histone H4 gene [Mus
musculus]
 gi|20345939|ref|XP_110064.1|  (XM_110064) similar to germinal histone H4 gene [Mus
musculus]
 gi|20345977|ref|XP_111512.1|  (XM_111512) similar to germinal histone H4 gene [Mus
musculus]
 gi|20345985|ref|XP_111506.1|  (XM_111506) similar to germinal histone H4 gene [Mus
musculus]
 gi|122098|sp|P02304|H4_HUMAN  Histone H4
 gi|86307|pir||JH0507  histone H4.III and H4.IV - chicken
 gi|90625|pir||S03426  histone H4 - mouse
 gi|102612|pir||S11312  histone H4 - polychaete (Platynereis dumerilii)
 gi|279637|pir||HSXL4  histone H4 - African clawed frog
```

```
 gi|630950|pir||S21367  histone H4 - Nile tilapia
 gi|1070598|pir||HSHU4  histone H4 - human
 gi|1078926|pir||D56618  histone H4 - spoonworm (Urechis caupo)
 gi|2134035|pir||I51433  histone H4 - Kenyan clawed frog
 gi|2134256|pir||I50459  H4 histone - muscovy duck
 gi|2144745|pir||HSRT4  histone H4 - rat
 gi|11513398|pdb|1F66|B  Chain B, 2.6 A Crystal Structure Of A Nucleosome Core Particle
Containing The Variant Histone H2a.Z
 gi|11513402|pdb|1F66|F  Chain F, 2.6 A Crystal Structure Of A Nucleosome Core Particle
Containing The Variant Histone H2a.Z
 gi|13096388|pdb|1HQ3|D  Chain D, Crystal Structure Of The Histone-Core-Octamer In
KclPHOSPHATE
 gi|13096392|pdb|1HQ3|H  Chain H, Crystal Structure Of The Histone-Core-Octamer In
KclPHOSPHATE
 gi|7767079|pdb|1EQZ|H  Chain H, X-Ray Structure Of The Nucleosome Core Particle At 2.5 A
Resolution
 gi|7767075|pdb|1EQZ|D  Chain D, X-Ray Structure Of The Nucleosome Core Particle At 2.5 A
Resolution
 gi|9818|emb|CAA37414.1|  (X53330) H4 protein (AA 1 - 103) [Platynereis dumerilii]
 gi|10866|emb|CAA41699.1|  (X58895) H4 histone [Urechis caupo]
 gi|31995|emb|CAA43011.1|   (X60481) H4 histone [Homo sapiens]
 gi|31997|emb|CAA43012.1|   (X60482) H4 histone [Homo sapiens]
 gi|31999|emb|CAA43013.1|   (X60483) H4 histone [Homo sapiens]
 gi|32001|emb|CAA43014.1|   (X60484) H4 histone [Homo sapiens]
 gi|32004|emb|CAA43016.1|   (X60486) H4 histone [Homo sapiens]
 gi|32006|emb|CAA43017.1|   (X60487) H4 histone [Homo sapiens]
 gi|32008|emb|CAA47464.1|   (X67081) histone [Homo sapiens]
 gi|51269|emb|CAA24130.1|   (V00753) histone H4 [Mus musculus]
 gi|51309|emb|CAA31621.1|   (X13235) histone H4 (AA 1 -103) [Mus musculus]
 gi|56349|emb|CAA31906.1|  (X13554) histone H4 precursor (AA -1 to 102) [Rattus norvegicus]
 gi|62733|emb|CAA32854.1|  (X14732) H4 histone (AA 1-103) [Cairina moschata]
 gi|62736|emb|CAA32857.1|  (X14732) H4 histone (AA 1-103) [Cairina moschata]
 gi|63472|emb|CAA26137.1|  (X02218) histone H4 (aa 1-103) [Gallus gallus]
 gi|63476|emb|CAA26140.1|  (X02218) H4 (aa 1-103) [Gallus gallus]
 gi|64321|emb|CAA26672.1|  (X02916) histone H4 (aa 1-103) [Oncorhynchus mykiss]
 gi|64427|emb|CAA38015.1|  (X54078) histone H4 [Oreochromis niloticus]
 gi|64765|emb|CAA25042.1|  (X00224) H4 histone [Xenopus laevis]
 gi|64768|emb|CAA26809.1|  (X03017) H4 (aa 1-103) [Xenopus laevis]
 gi|64773|emb|CAA26814.1|  (X03017) H4 (aa 1-103) [Xenopus laevis]
 gi|64779|emb|CAA26819.1|  (X03018) histone H4 (aa 1-103) [Xenopus laevis]
 gi|184076|gb|AAA63188.1|   (M60749) histone H4 [Homo sapiens]
 gi|204552|gb|AAA41306.1|   (M28409) histone H4 [Rattus norvegicus]
 gi|204601|gb|AAA60735.1|   (M27433) histone H4 [Rattus norvegicus]
 gi|211906|gb|AAA73091.1|  (M74533) [Chicken histone H4 protein gene, complete cds.], gene
product [Gallus gallus]
 gi|211908|gb|AAA73092.1|  (M74534) [Chicken histone H4 protein gene, complete cds.], gene
product [Gallus gallus]
 gi|214224|gb|AAA72138.1|  (J00985) [Xenopus borealis h4 histone mRNA.], gene product
 gi|214275|gb|AAA49761.1|  (M21286) histone H4 [Xenopus laevis]
 gi|214280|gb|AAA49766.1|  (M21286) histone H4 [Xenopus laevis]
 gi|214286|gb|AAA49771.1|  (M21287) histone H4 [Xenopus laevis]
 gi|386773|gb|AAA52652.1|   (M16707) histone H4 [Homo sapiens]
 gi|603556|emb|CAA58538.1|   (X83548) histone H4 [Homo sapiens]
 gi|1458137|gb|AAB04766.1|   (U62672) histone H4-D [Mus musculus]
 gi|1493809|gb|AAC59999.1|  (U37575) histone H4-VI [Gallus gallus]
 gi|1493811|gb|AAC60001.1|  (U37575) histone H4-VII [Gallus gallus]
```

```
 gi|1568565|emb|CAB02549.1|  ☐ (Z80787) histone H4 [Homo sapiens]
 gi|1840407|dbj|BAA19208.1|  ☐ (AB000905) H4 histone [Homo sapiens]
 gi|2292936|emb|CAA72967.1|  ☐ (Y12290) Histone H4 [Mus musculus]
 gi|2564108|gb|AAC15917.1|  (AF007904) histone H4 [Chaetopterus variopedatus]
 gi|2981288|gb|AAC39176.1|  (AF001288) histone H4.1 [Bos taurus]
 gi|3036788|emb|CAA16946.1|  ☐ (AL021807) dJ86C11.3 (H4 histone family, member M) [Homo
sapiens]
 gi|9863664|emb|CAC04128.1|  (AL353759) dJ221C16.1 (novel histone 4 family member) [Homo
sapiens]
 gi|9863671|emb|CAC04136.1|  (AL353759) dJ221C16.9 (novel histone 4 family member) [Homo
sapiens]
 gi|10198613|emb|CAC03414.1|  (AL031777) dJ34B20.12 (H4 histone family, member J) [Homo
sapiens]
 gi|10198616|emb|CAC03418.1|  (AL031777) dJ34B20.6 (H4 histone family, member C) [Homo
sapiens]
 gi|10198638|emb|CAC03426.1|  ☐ (AL049822) dJ160A22.2 (H4 histone family, member E) [Homo
sapiens]
 gi|10198639|emb|CAC03427.1|  ☐ (AL049822) dJ160A22.1 (H4 histone family, member D) [Homo
sapiens]
 gi|12841313|dbj|BAB25157.1|  (AK007642) data source:MGD, source key:MGI:96099,
evidence:ISS~histone 4 protein~putative [Mus musculus]
 gi|15012055|gb|AAH10926.1|AAH10926  (BC010926) Similar to H4 histone family, member A [Homo
sapiens]
 gi|15214896|gb|AAH12587.1|AAH12587  (BC012587) Similar to H4 histone family, member A [Homo
sapiens]
 gi|15559186|emb|CAC69642.1|  (AL021917) dJ45P21.7 (H4 histone family, member A) [Homo
sapiens]
 gi|16740965|gb|AAH16336.1|AAH16336  ☐ (BC016336) H4 histone family, member A [Homo
sapiens]
 gi|16878339|gb|AAH17361.1|AAH17361  ☐ (BC017361) H4 histone family, member A [Homo
sapiens]
 gi|18043704|gb|AAH19846.1|AAH19846  ☐ (BC019846) H4 histone family, member A [Homo
sapiens]
 gi|18043918|gb|AAH19757.1|AAH19757  (BC019757) H4 histone family, member A [Mus musculus]
 gi|18088762|gb|AAH20884.1|AAH20884  ☐ (BC020884) H4 histone family, member A [Homo
sapiens]
 gi|18857858|emb|CAD24074.1|  (Z98744) dJ193B12.12 (H4 histone family, member K) [Homo
sapiens]
 gi|20306504|gb|AAH28550.1|  (BC028550) H4 histone family, member A [Mus musculus]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1
```

```
Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56




>gi|462242|sp|P35059|H4_ACRFO  HISTONE H4
 gi|166311|gb|AAC37355.1|  (L11067) histone H4 [Acropora formosa]
 gi|455652|gb|AAB28739.1|  (S67324) histone H4; H4 [Acropora formosa]
 gi|450047|prf||1920342D  histone H4 [Acropora formosa]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 98




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56




>gi|1806283|emb|CAB01913.1|  (Z79637) Histone H4 homologue [Sesbania rostrata]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98




 Score = 35.4 bits (80), Expect = 0.55
 Identities = 19/29 (65%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLA+RGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLAKRGGVKRISGLIYEETR 56
```

```
>gi|1199967|emb|CAA64985.1|  (X95689) histone H4 [Allium cepa]
         Length = 58

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 1   RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 53



>gi|4139870|pdb|1HIO|D  Chain D, Histone Octamer (Chicken), Chromosomal Protein, Alpha
          Carbons Only
         Length = 76

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 19  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 71



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 1   QGITKPAIRRLARRGGVKRISGLIYEETR 29



>gi|224293|prf||1101277A  histone H4 [Triticum aestivum]
         Length = 102

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 97



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1
```

```
Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 27  QGITKPAIRRLARRGGVKRISGLIYEETR 55




>gi|70772|pir||HSWT4  histone H4 - wheat
          Length = 102

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 97





 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 27  QGITKPAIRRLARRGGVKRISGLIYEETR 55




>gi|1883028|emb|CAA62811.1|  (X91511) histone H4 [Diprion pini]
          Length = 103

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 27/53 (50%), Positives = 40/53 (74%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++  R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 98





 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56
```

```
>gi|19569887|gb|AAL92205.1|AC115576_7  (AC115576) hypothetical protein [Dictyostelium
discoideum]
 gi|19703116|gb|AAL93580.1|AC116100_14  (AC116100) HISTONE 4 PROTEIN [Dictyostelium
discoideum]
          Length = 108

 Score = 60.8 bits (146), Expect = 1e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   +++  R+V+  FL+++++D++ YTEHA R TV AMDVVYALK+ GR L
Sbjct: 51  RISFPIYEETRNVLRTFLTNVIRDSVAYTEHAGRRTVTAMDVVYALKRQGRTL 103



 Score = 35.8 bits (81), Expect = 0.42
 Identities = 19/27 (70%), Positives = 19/27 (70%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 35  ITKPAIRRLARRGGVKRISFPIYEETR 61



>gi|122099|sp|P04914|H4_NEUCR  Histone H4
 gi|83768|pir||S07913  histone H4 - Neurospora crassa
 gi|3018|emb|CAA25760.1|  (X01611) histone H4 [Neurospora crassa]
 gi|17644132|gb|AAL38972.1|  (AY062172) histone H4 [Neurospora crassa]
 gi|17644135|gb|AAL38974.1|  (AY062173) histone H4 [Neurospora crassa]
 gi|18307449|emb|CAD21509.1|  (AL670543) histone H4 [Neurospora crassa]
 gi|20145254|emb|CAD29611.1|  (AL731691) probable histone h4 [Aspergillus fumigatus]
          Length = 103

 Score = 60.5 bits (145), Expect = 2e-08
 Identities = 25/53 (47%), Positives = 39/53 (73%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL  +++DA+ YTEHA+R TV ++DVVYALK+ GR L
Sbjct: 46  RISAMIYEETRGVLKTFLEGVIRDAVTYTEHAKRKTVTSLDVVYALKRQGRTL 98



 Score = 37.4 bits (85), Expect = 0.15
 Identities = 20/29 (68%), Positives = 21/29 (71%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS+  Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISAMIYEETR 56
```

```
>gi|122092|sp|P23750|H41_EMENI  HISTONE H4.1
 gi|101785|pir||S11939  histone H4.1 - Emericella nidulans
 gi|296341|emb|CAA39155.1|  (X55549) H4.1 [Emericella nidulans]
 gi|529955|gb|AAA20820.1|  (U12630) histone H4.1 [Emericella nidulans]
 gi|9955877|dbj|BAB12238.1|  (AB033943) histone H4 [Aspergillus oryzae]
 gi|227597|prf||1707275C  histone H4.1 [Emericella nidulans]
          Length = 103

 Score = 60.5 bits (145), Expect = 2e-08
 Identities = 25/53 (47%), Positives = 39/53 (73%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+   ++++   R V+   FL   +++DA+ YTEHA+R TV ++DVVYALK+ GR L
Sbjct: 46  RISAMIYEETRGVLKTFLEGVIRDAVTYTEHAKRKTVTSLDVVYALKRQGRTL 98



 Score = 37.4 bits (85), Expect = 0.15
 Identities = 20/29 (68%), Positives = 21/29 (71%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS+   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISAMIYEETR 56



>gi|223793|prf||0912198A  histone H4 [Physarum polycephalum]
          Length = 102

 Score = 60.5 bits (145), Expect = 2e-08
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+    +++   R V+   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 45  RISBTIYZZTRGVLKVFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 97



 Score = 37.0 bits (84), Expect = 0.19
 Identities = 19/29 (65%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           + IT   IRRLARRGGVKRIS    Y  TR
Sbjct: 27  ZGITKPAIRRLARRGGVKRISBTIYZZTR 55



>gi|515377|emb|CAA56154.1|  (X79715) histone H4 [Lolium temulentum]
          Length = 103
```

```
 Score = 60.1 bits (144), Expect = 2e-08
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDAVTYTEHAXRKTVTAMDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|484442|pir||JN0688  histone H4 - sea squirt (Styela plicata)
 gi|7522682|gb|AAB27670.2|  (S64499) H4 histone [Styela plicata]
          Length = 103

 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV A+DVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTALDVVYALKRQGRTL 98



 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|462243|sp|P35057|H4_LYCES  HISTONE H4
 gi|421921|pir||S32769  histone H4 - tomato
 gi|297150|emb|CAA48923.1|  (X69179) histone H4 [Lycopersicon esculentum]
 gi|297152|emb|CAA48924.1|  (X69180) histone H4 [Lycopersicon esculentum]
 gi|2746721|gb|AAB94924.1|  (AF038387) histone H4 [Capsicum annuum]
          Length = 103

 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2
```

```
Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++D++ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVLKIFLENVIRDSVTYTEHARRKTVTAMDVVYALKRQGRTL 98




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|2137382|pir||I48404  histone H4 (55AA) (1 is 3rd base in codon) - mouse  (fragment)
 gi|817960|emb|CAA28350.1|    (X04652) histone H4 (55AA) (1 is 3rd base in codon) [Mus
musculus]
          Length = 55

 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 26/49 (53%), Positives = 38/49 (77%)
 Frame = -2

Query: 350 VLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 2   LIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTL 50



>gi|20141289|sp|P80739|H4_EUPCR  Histone H4
 gi|1763310|gb|AAB39722.1|  (U75430) histone H4 [Euplotes crassus]
          Length = 107

 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 25/53 (47%), Positives = 40/53 (75%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL S+++D++ YTEHA+R TV A+DVVYALK+ G+ L
Sbjct: 50  RISSLVYEETRAVLKGFLESVIRDSVTYTEHAKRKTVTALDVVYALKRQGKTL 102




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 19/27 (70%), Positives = 20/27 (73%)
 Frame = -1

Query: 435 ITNGDIRRLARRGGVKRISSDSYPTTR 355
           +T   IRRLARRGGVKRISS  Y  TR
Sbjct: 34  VTKPAIRRLARRGGVKRISSLVYEETR 60
```

```
>gi|6006733|gb|AAF00589.1|AF140034_1  (AF140034) histone H4 [Mastigamoeba balamuthi]
          Length = 108

 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++D++ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 51  RISGLIYEETRGVLKVFLENVIRDSVTYTEHARRKTVTAMDVVYALKRQGRTL 103




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 33  QGITKPAIRRLARRGGVKRISGLIYEETR 61



>gi|2062367|gb|AAB53361.1|  (U65675) histone H4 [Plasmodium falciparum]
          Length = 79

 Score = 59.3 bits (142), Expect = 4e-08
 Identities = 26/53 (49%), Positives = 41/53 (77%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL +++KD+I+YTEHA+R TV AMD+VY+LK+ GR L
Sbjct: 27  RISGLIYEEIRGVLKVFLENVIKDSIMYTEHAKRKTVTAMDIVYSLKRQGRTL 79




 Score = 34.3 bits (77), Expect = 1.2
 Identities = 19/29 (65%), Positives = 19/29 (65%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y   R
Sbjct: 9   QGITKPAIRRLARRGGVKRISGLIYEEIR 37



>gi|13812130|ref|NP_113257.1|  (NC_002752) Histone H4 [Guillardia theta]
 gi|13794442|gb|AAK39817.1|AF165818_25  (AF165818) Histone H4 [Guillardia theta]
          Length = 104

 Score = 59.3 bits (142), Expect = 4e-08
 Identities = 26/53 (49%), Positives = 40/53 (75%)
 Frame = -2
```

```
Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  + V+  FL ++VKDA+ YTEHA+R TV AMDV+YALK+ GR +
Sbjct: 47  RISGLIYEETKIVLKLFLENVVKDAVTYTEHARRKTVTAMDVIYALKRQGRTI 99




 Score = 35.4 bits (80), Expect = 0.55
 Identities = 19/29 (65%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  T+
Sbjct: 29  QGITKPAIRRLARRGGVKRISGLIYEETK 57



>gi|90626|pir||S03427  histone H4 (clone 53) - mouse
 gi|51311|emb|CAA31622.1|  (X13236) histone H4 (AA 1 - 103) [Mus musculus]
          Length = 103

 Score = 58.9 bits (141), Expect = 5e-08
 Identities = 26/53 (49%), Positives = 39/53 (73%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V+  FL ++++DA+ YTEHA+R TV AMDVVY LK+ GR L
Sbjct: 46  RISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYRLKRQGRTL 98




 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56



>gi|886740|emb|CAA59110.1|  (X84376) histone 4 [Zea mays]
          Length = 103

 Score = 58.9 bits (141), Expect = 5e-08
 Identities = 27/53 (50%), Positives = 39/53 (72%)
 Frame = -2

Query: 362 RLGIVLFDPNRDVIVNFLSSLVKDAIIYTEHAQRNTVQAMDVVYALKKYGRNL 204
           R+  ++++  R V   FL ++++DA+ YTEHA+R TV AMDVVYALK+ GR L
Sbjct: 46  RISGLIYEETRGVRKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTL 98
```

```
 Score = 36.6 bits (83), Expect = 0.25
 Identities = 20/29 (68%), Positives = 20/29 (68%)
 Frame = -1

Query: 441 QHITNGDIRRLARRGGVKRISSDSYPTTR 355
           Q IT   IRRLARRGGVKRIS   Y  TR
Sbjct: 28  QGITKPAIRRLARRGGVKRISGLIYEETR 56




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 9, 2002  6:14 AM
  Number of letters in database: 288,558,979
  Number of sequences in database:  919,285

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 339,700,645
Number of Sequences: 919285
Number of extensions: 6560764
Number of successful extensions: 23773
Number of sequences better than 10.0: 224
Number of HSP's better than 10.0 without gapping: 22763
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 23755
length of database: 288,558,979
effective HSP length: 117
effective length of database: 181,002,634
effective search space used: 22987334518
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

BLASTX 2.2.3 [Apr-24-2002]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1020947669-028505-4569
```

7.1.18 Query= **04199901 (*upstream message*)**

```
         (237 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           933,889 sequences; 292,803,047 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 110 Blast Hits on the Query Sequence

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits)  Value

gi|7438334|pir||T05954  transcription factor myb4 - barley (...    38    0.022
gi|11358531|pir||T51640  myb-related transcription factor MY...    36    0.082
gi|15242947|ref|NP_200039.1|  (NM_124605) putative protein [...    36    0.082
gi|11358535|pir||T51649  myb-related transcription factor MY...    34    0.31
gi|15233006|ref|NP_189488.1|  (NM_113767) myb transcription ...    34    0.31
gi|15225822|ref|NP_180263.1|  (NM_128252) putative MYB famil...    34    0.41
gi|11358536|pir||T51651  myb-related transcription factor MY...    34    0.41
gi|15237193|ref|NP_197691.1|  (NM_122206) myb-related transc...    34    0.41
gi|7446163|pir||T02640  hypothetical protein F12C20.1 - Arab...    34    0.41
gi|7446182|pir||T03830  probable myb factor - rice >gi|19452...    34    0.41
gi|15219601|ref|NP_176797.1|  (NM_105294) myb-related transc...    34    0.41
gi|11358686|pir||T51650  probable transcription factor MYB36...    33    0.53
gi|15242793|ref|NP_200570.1|  (NM_125143) Myb-related transc...    33    0.53
gi|15241377|ref|NP_196938.1|  (NM_121438) MYB40 - putative t...    33    0.53
gi|11358538|pir||T51653  myb-related transcription factor MY...    33    0.53
gi|7446173|pir||T02912  probable transcription factor MYB41 ...    33    0.91
gi|19548467|gb|AAL90657.1|AF470089_1  (AF470089) P-type R2R3...    33    0.91
gi|19074969|ref|NP_586475.1|  (NC_003237) similarity to Myb-...    33    0.91
gi|15237380|ref|NP_197163.1|  (NM_121666) transcription fact...    33    0.91
gi|15234385|ref|NP_194540.1|  (NM_118951) putative transcrip...    33    0.91
gi|15229407|ref|NP_188966.1|  (NM_113226) myb-related transc...    32    1.2
gi|11358530|pir||T51639  myb-related transcription factor MY...    32    1.2
gi|15234199|ref|NP_195071.1|  (NM_119499) putative transcrip...    32    1.2
```

```
gi|15230123|ref|NP_189640.1|  (NM_113920) myb-like transcrip...    32   1.2
gi|15232334|ref|NP_191605.1|  (NM_115910) myb protein - like...    32   1.2
gi|2832404|emb|CAA74603.1|  (Y14207) R2R3-MYB transcription ...    32   1.2
gi|15235149|ref|NP_194286.1|  (NM_118688) myb-like protein [...    32   1.2
gi|7269406|emb|CAB81366.1|  (AL161563) myb-like protein [Ara...    32   1.2
gi|7446177|pir||T07393  myb-related transcription factor - t...    32   1.5
gi|15220910|ref|NP_173237.1|  (NM_101658) myb-like protein, ...    32   1.5
gi|7438328|pir||T03828  myb protein - rice >gi|1946267|emb|C...    32   1.5
gi|11358559|pir||T51679  myb-related transcription factor MY...    32   1.5
gi|9665072|gb|AAF97274.1|AC034106_17  (AC034106) Contains si...    32   1.5
gi|4886264|emb|CAB43399.1|  (AJ006292) Myb-related transcrip...    32   1.5
gi|19073324|gb|AAL84760.1|AF474127_1  (AF474127) typical P-t...    32   1.5
gi|15234593|ref|NP_192419.1|  (NM_116749) MYB - like protein...    32   1.5
gi|17065086|gb|AAL32697.1|  (AY062619) myb-related protein M...    32   2.0
gi|18415694|ref|NP_567626.1|  (NM_118264) myb-related protei...    32   2.0
gi|2129563|pir||S58280  cM4 protein - Arabidopsis thaliana         32   2.0
gi|11358541|pir||T51660  myb-related transcription factor MY...    32   2.0
gi|15238370|ref|NP_201326.1|  (NM_125921) transcription fact...    32   2.0
gi|6467225|gb|AAF13101.1|U26934_1  (U26934) DNA-binding prot...    32   2.0
gi|19548457|gb|AAL90652.1|AF470084_1  (AF470084) P-type R2R3...    32   2.0
gi|7428545|pir||S58293  myb-related protein M4 - Arabidopsis...    32   2.0
gi|15225824|ref|NP_180264.1|  (NM_128253) putative MYB famil...    31   2.6
gi|15290138|dbj|BAB63829.1|  (AP003627) putative Myb factor ...    31   2.6
gi|15222362|ref|NP_177115.1|  (NM_105625) MYB-family transcr...    31   2.6
gi|13346188|gb|AAK19616.1|AF336283_1  (AF336283) GHMYB25 [Go...    31   2.6
gi|7438352|pir||T02006  transcription factor MYB4 homolog T1...    31   2.6
gi|6979341|gb|AAF34434.1|AF172282_23  (AF172282) myb-like pr...    31   2.6
gi|9295726|gb|AAF87032.1|AC006535_10  (AC006535) T24P13.16 [...    31   2.6
gi|19073332|gb|AAL84764.1|AF474131_1  (AF474131) typical P-t...    31   2.6
gi|15234262|ref|NP_192077.1|  (NM_116398) putative transcrip...    31   2.6
gi|18396049|ref|NP_564261.1|  (NM_102441) myb-related protei...    31   2.6
gi|6492385|dbj|BAA81733.2|  (AB029162) GmMYB29A2 [Glycine max]     31   2.6
gi|11358563|pir||T51683  myb-related transcription factor MY...    31   2.6
gi|11994486|dbj|BAB02527.1|  (AP000371) ATMYB3 [Arabidopsis ...    31   3.5
gi|5139814|dbj|BAA81736.1|  (AB029165) GmMYB29B2 [Glycine max]     31   3.5
gi|13366181|dbj|BAB39404.1|  (AP002901) putative transcripti...    31   3.5
gi|11358543|pir||T51662  myb-related transcription factor MY...    31   3.5
gi|15219462|ref|NP_177484.1|  (NM_106001) myb-like transcrip...    31   3.5
gi|11358682|pir||T51642  probable transcription factor MYB21...    31   3.5
gi|730090|sp|Q08759|MYB_XENLA  Myb protein >gi|2118431|pir||...    31   3.5
gi|11967859|emb|CAC19439.1|  (AJ237661) Myb factor protein [...    31   3.5
gi|19073322|gb|AAL84759.1|AF474126_1  (AF474126) typical P-t...    31   3.5
gi|15229055|ref|NP_190461.1|  (NM_114751) myb-like protein [...    31   3.5
gi|18419456|gb|AAL69334.1|AF426174_1  (AF426174) blind [Lyco...    31   3.5
gi|15219259|ref|NP_173098.1|  (NM_101514) putative myb-relat...    30   4.5
gi|28686|emb|CAA31656.1|  (X13294) A-myb N-terminal region )...    30   4.5
gi|1171089|sp|P10243|MYBA_HUMAN  Myb-related protein A (A-My...    30   4.5
gi|19072768|gb|AAL84629.1|AF474139_1  (AF474139) typical P-t...    30   4.5
gi|13486684|dbj|BAB39921.1|  (AP002912) putative transcripti...    30   4.5
gi|2506883|sp|P51960|MYBA_MOUSE  Myb-related protein A (A-Myb)     30   4.5
gi|2118433|pir||I49497  transforming protein A-myb - mouse >...    30   4.5
gi|6678972|ref|NP_032677.1|  (NM_008651) myeloblastosis onco...    30   4.5
gi|11358532|pir||T51641  myb-related transcription factor MY...    30   4.5
gi|19072740|gb|AAL84615.1|AF474118_1  (AF474118) typical P-t...    30   4.5
gi|462669|sp|P80073|MYB2_PHYPA  Myb-related protein Pp2 >gi|...    30   5.9
gi|15238116|ref|NP_196590.1|  (NM_121066) putative transcrip...    30   5.9
gi|15239158|ref|NP_201380.1|  (NM_125976) transcription fact...    30   5.9
gi|6322952|ref|NP_013025.1|  (NC_001143) Transcription facto...    30   5.9
gi|7939552|dbj|BAA95755.1|  (AB025615) MYB transcription fac...    30   5.9
gi|15224628|ref|NP_180676.1|  (NM_128674) myb-related protei...    30   5.9
gi|18406019|ref|NP_566841.1|  (NM_113823) expressed protein ...    30   5.9
```

```
gi|11273952|pir||T51673  myb-related transcription factor MY...    30   5.9
gi|15375293|gb|AAK25747.2|AF334814_1  (AF334814) putative tr...    30   5.9
gi|18421977|ref|NP_568581.1|  (NM_123400) MYB -like protein ...    30   5.9
gi|7446172|pir||A71448  probable MYB transcription factor - ...    30   5.9
gi|127579|sp|P20026|MYB1_HORVU  Myb-related protein Hv1 >gi|...    30   5.9
gi|18414965|ref|NP_567540.1|  (NM_117888) MYB transcription ...    30   5.9
gi|10178146|dbj|BAB11591.1|  (AB006702) contains similarity ...    30   5.9
gi|15236297|ref|NP_193084.1|  (NM_117422) myb-like protein [...    30   5.9
gi|18461236|dbj|BAB84433.1|  (AP003292) Myb-related transcri...    30   5.9
gi|15225582|ref|NP_181517.1|  (NM_129546) putative MYB famil...    30   5.9
gi|15222565|ref|NP_176575.1|  (NM_105065) putative MYB famil...    30   5.9
gi|15239578|ref|NP_200234.1|  (NM_124803) Myb-related transc...    30   7.7
gi|15232137|ref|NP_186802.1|  (NM_111019) putative transcrip...    30   7.7
gi|7438337|pir||T03850  myb-related protein myb1, TMV-induci...    30   7.7
gi|19072736|gb|AAL84613.1|AF474116_1  (AF474116) typical P-t...    30   7.7
gi|11273960|pir||T51645  myb-related transcription factor MY...    30   7.7

                                      Alignments


>gi|7438334|pir||T05954  transcription factor myb4 - barley (fragment)
 gi|1617325|emb|CAA68235.1|  (X99973) myb4 transcription factor [Hordeum vulgare subsp.
vulgare]
          Length = 288

 Score = 38.1 bits (87), Expect = 0.022
 Identities = 20/65 (30%), Positives = 37/65 (56%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLNEKEKVRGV 211
           FS  E++ I++L   +G  +++I    PG+T N IKN ++  L+   +    E+++  G
Sbjct: 68  FSDEEEQSIIQLHQLLGNRWSAIAARLPGRTDNEIKNVWHTHLKKRLDPSAQEQQEEAGA 127

Query: 212 RKQRK 226
            K+RK
Sbjct: 128 AKKRK 132




>gi|11358531|pir||T51640  myb-related transcription factor MYB19 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941428|gb|AAC83590.1|  (AF062868) putative transcription factor [Arabidopsis thaliana]
          Length = 215

 Score = 36.2 bits (82), Expect = 0.082
 Identities = 21/49 (42%), Positives = 30/49 (60%), Gaps = 2/49 (4%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY--YKKLRYMK 172
           FS  E++ IL L   +G  ++ I KY PG+T N IKN +  Y K R++K
Sbjct: 17  FSEEEEETILTLHSSLGNKWSRIAKYLPGRTDNEIKNYWHSYLKKRWLK 65




>gi|15242947|ref|NP_200039.1|  (NM_124605) putative protein [Arabidopsis thaliana]
 gi|8885539|dbj|BAA97469.1|  (AB025603) contains similarity to myb-related transcription
factor~gene_id:F17P19.16 [Arabidopsis thaliana]
          Length = 268
```

```
 Score = 36.2 bits (82), Expect = 0.082
 Identities = 21/49 (42%), Positives = 30/49 (60%), Gaps = 2/49 (4%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY--YKKLRYMK 172
            FS  E++ IL L   +G  ++ I KY PG+T N IKN +  Y K R++K
Sbjct: 70  FSEEEEETILTLHSSLGNKWSRIAKYLPGRTDNEIKNYWHSYLKKRWLK 118




>gi|11358535|pir||T51649  myb-related transcription factor MYB35 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941446|gb|AAC83599.1|  (AF062877) putative transcription factor [Arabidopsis thaliana]
          Length = 237

 Score = 34.3 bits (77), Expect = 0.31
 Identities = 17/49 (34%), Positives = 28/49 (56%)
 Frame = +2

Query: 17  LNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
            L    FS  E++ I+E    +G  ++SI +  PG+T N +KN +  KL+
Sbjct: 12  LKHDSFSTQEEELIIECHRAIGSRWSSIARKLPGRTDNDVKNHWNTKLK 60




>gi|15233006|ref|NP_189488.1|  (NM_113767) myb transcription factor, putative [Arabidopsis
thaliana]
 gi|9294582|dbj|BAB02863.1|  (AB026644) transcription factor-like protein [Arabidopsis
thaliana]
          Length = 317

 Score = 34.3 bits (77), Expect = 0.31
 Identities = 17/49 (34%), Positives = 28/49 (56%)
 Frame = +2

Query: 17  LNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
            L    FS  E++ I+E    +G  ++SI +  PG+T N +KN +  KL+
Sbjct: 65  LKHDSFSTQEEELIIECHRAIGSRWSSIARKLPGRTDNDVKNHWNTKLK 113




>gi|15225822|ref|NP_180263.1|  (NM_128252) putative MYB family transcription factor
[Arabidopsis thaliana]
 gi|6598476|gb|AAF18614.1|  (AC005623) putative MYB family transcription factor [Arabidopsis
thaliana]
 gi|20197285|gb|AAM15011.1|  (AC005168) putative MYB family transcription factor
[Arabidopsis thaliana]
          Length = 375

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 17/45 (37%), Positives = 27/45 (59%)
 Frame = +2


Query: 8   KRKLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKN 142
            K  L    F+  E+KR+L+L   +G  ++ + + FPG+T N IKN
Sbjct: 55  KPSLKKGPFTDEEEKRVLQLHAVLGNKWSQMAREFPGRTDNEIKN 99
```

```
>gi|11358536|pir||T51651  myb-related transcription factor MYB37 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941450|gb|AAC83601.1|  (AF062879) putative transcription factor [Arabidopsis thaliana]
          Length = 275

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 18/44 (40%), Positives = 25/44 (55%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           FS  ED+ I  L   +G  ++ I  + PG+T N IKN +  KLR
Sbjct: 17  FSEEEDRIIFSLFAAIGSRWSIIAAHLPGRTDNDIKNYWNTKLR 60



>gi|15237193|ref|NP_197691.1|  (NM_122206) myb-related transcription factor-like
[Arabidopsis thaliana]
 gi|9758402|dbj|BAB08873.1|  (AB026660) myb-related transcription factor-like [Arabidopsis
thaliana]
 gi|18072740|emb|CAC80101.1|  (AJ131517) R2R3-MYB transcription factor [Arabidopsis
thaliana]
          Length = 329

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 18/44 (40%), Positives = 25/44 (55%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           FS  ED+ I  L   +G  ++ I  + PG+T N IKN +  KLR
Sbjct: 71  FSEEEDRIIFSLFAAIGSRWSIIAAHLPGRTDNDIKNYWNTKLR 114



>gi|7446163|pir||T02640  hypothetical protein F12C20.1 - Arabidopsis thaliana
          Length = 386

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 17/45 (37%), Positives = 27/45 (59%)
 Frame = +2

Query: 8   KRKLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKN 142
           K  L    F+  E+KR+L+L   +G  ++ + + FPG+T N IKN
Sbjct: 66  KPSLKKGPFTDEEEKRVLQLHAVLGNKWSQMAREFPGRTDNEIKN 110



>gi|7446182|pir||T03830  probable myb factor - rice
 gi|1945283|emb|CAA72186.1|  (Y11351) myb factor [Oryza sativa (japonica cultivar-group)]
          Length = 279

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 20/57 (35%), Positives = 32/57 (56%), Gaps = 4/57 (7%)
 Frame = +2
```

```
Query: 35  SYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEK 193
           S +E+K +++L   +G  ++ I  + PG+T N IKN +     KKLR M  D +  K
Sbjct: 71  SETEEKTVIDLHEQLGNRWSKIASHLPGRTDNEIKNHWNTHIKKKLRKMGIDPVTHK 127



>gi|15219601|ref|NP_176797.1|  (NM_105294) myb-related transcription factor, putative
[Arabidopsis thaliana]
 gi|12323572|gb|AAG51765.1|AC066691_5  (AC066691) myb-related transcription factor,
putative; 17635-18559 [Arabidopsis thaliana]
          Length = 282

 Score = 33.9 bits (76), Expect = 0.41
 Identities = 20/54 (37%), Positives = 30/54 (55%), Gaps = 4/54 (7%)
 Frame = +2

Query: 44  EDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEK 193
           E+K +++L   +G  ++ I  + PG+T N IKN +     KKLR M  D L  K
Sbjct: 74  EEKMVIDLHSQLGNRWSKIASHLPGRTDNEIKNHWNTHIKKKLRKMGIDPLTHK 127



>gi|11358686|pir||T51650  probable transcription factor MYB36 [imported] - Arabidopsis
thaliana
 gi|3941448|gb|AAC83600.1|  (AF062878) putative transcription factor [Arabidopsis thaliana]
          Length = 333

 Score = 33.5 bits (75), Expect = 0.53
 Identities = 23/63 (36%), Positives = 32/63 (50%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLNEKEKVRGV 211
           FS  ED+ IL L   +G  ++ I    PG+T N IKN +  KL          K+K+ G
Sbjct: 71  FSEEEDRIILSLYISIGSRWSIIAAQLPGRTDNDIKNYWNTKL----------KKKLLGR 120

Query: 212 RKQ 220
           +KQ
Sbjct: 121 QKQ 123



>gi|15242793|ref|NP_200570.1|  (NM_125143) Myb-related transcription factor-like protein
[Arabidopsis thaliana]
 gi|9758323|dbj|BAB08797.1|  (AB011482) Myb-related transcription factor-like protein
[Arabidopsis thaliana]
          Length = 333

 Score = 33.5 bits (75), Expect = 0.53
 Identities = 23/63 (36%), Positives = 32/63 (50%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLNEKEKVRGV 211
           FS  ED+ IL L   +G  ++ I    PG+T N IKN +  KL          K+K+ G
Sbjct: 71  FSEEEDRIILSLYISIGSRWSIIAAQLPGRTDNDIKNYWNTKL----------KKKLLGR 120

Query: 212 RKQ 220
           +KQ
Sbjct: 121 QKQ 123
```

```
>gi|15241377|ref|NP_196938.1|  (NM_121438) MYB40 - putative transcription factor
[Arabidopsis thaliana]
 gi|11274001|pir||T48607  probable transcription factor MYB40 - Arabidopsis thaliana
 gi|7573459|emb|CAB87773.1|  (AL163817) MYB40-putative transcription factor [Arabidopsis
thaliana]
          Length = 263

 Score = 33.5 bits (75), Expect = 0.53
 Identities = 16/44 (36%), Positives = 28/44 (63%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           F+ +E+ RI+EL   +G  ++ I  +F G+T N IKN +  K++
Sbjct: 70  FTDAEEDRIMELHSQLGNRWSKIASHFSGRTDNEIKNHWNTKIK 113



>gi|11358538|pir||T51653  myb-related transcription factor MYB40 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941454|gb|AAC83603.1|  (AF062881) putative transcription factor [Arabidopsis thaliana]
          Length = 203

 Score = 33.5 bits (75), Expect = 0.53
 Identities = 16/44 (36%), Positives = 28/44 (63%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           F+ +E+ RI+EL   +G  ++ I  +F G+T N IKN +  K++
Sbjct: 17  FTDAEEDRIMELHSQLGNRWSKIASHFSGRTDNEIKNHWNTKIK 60



>gi|7446173|pir||T02912  probable transcription factor MYB41 [similarity] - Arabidopsis
thaliana (fragment)
 gi|3941456|gb|AAC83604.1|  (AF062882) putative transcription factor [Arabidopsis thaliana]
 gi|4455370|emb|CAB36780.1|  (AL035524) putative transcription factor MYB41 [Arabidopsis
thaliana]
          Length = 218

 Score = 32.7 bits (73), Expect = 0.91
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I     PG+T N IKN +   +R
Sbjct: 5   RFSFEEEETIIQLHSVMGNKWSAIAARLPGRTDNEIKNHWNTHIR 49



>gi|19548467|gb|AAL90657.1|AF470089_1  (AF470089) P-type R2R3 Myb protein [Zea mays]
          Length = 157

 Score = 32.7 bits (73), Expect = 0.91
 Identities = 18/45 (40%), Positives = 26/45 (57%)
```

```
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           KFS  E++ IL L   +G  + +I  + PG+T N IKN +   LR
Sbjct: 69  KFSREEEQTILHLHSVLGNKWGAIATHLPGRTDNEIKNFWNTHLR 113




>gi|19074969|ref|NP_586475.1|  (NC_003237) similarity to Myb-related transcription factor
(fragment) [Encephalitozoon cuniculi]
 gi|19069694|emb|CAD26079.1|  (AL590450) similarity to Myb-related transcription factor
(fragment) [Encephalitozoon cuniculi]
          Length = 208

 Score = 32.7 bits (73), Expect = 0.91
 Identities = 17/48 (35%), Positives = 26/48 (53%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKE 175
           F+  E+  I+EL    G  ++ I KY PG+T N IKN +   ++   E
Sbjct: 75  FTMEEEGVIIELHSKFGNRWSEIAKYLPGRTDNAIKNYWNSSIQRRSE 122




>gi|15237380|ref|NP_197163.1|  (NM_121666) transcription factor (gb|AAD53095.1) [Arabidopsis
thaliana]
 gi|5823313|gb|AAD53095.1|AF175990_1  (AF175990) putative transcription factor [Arabidopsis
thaliana]
 gi|10176969|dbj|BAB10187.1|  (AB008270) transcription factor [Arabidopsis thaliana]
          Length = 327

 Score = 32.7 bits (73), Expect = 0.91
 Identities = 19/54 (35%), Positives = 30/54 (55%), Gaps = 4/54 (7%)
 Frame = +2

Query: 44  EDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEK 193
           E+++++ L   +G  ++ I  + PG+T N IKN +     KKLR M  D L  K
Sbjct: 74  EEQKVINLHAQLGNRWSKIASHLPGRTDNEIKNHWNTHIKKKLRKMGIDPLTHK 127




>gi|15234385|ref|NP_194540.1|  (NM_118951) putative transcription factor MYB41 [Arabidopsis
thaliana]
 gi|7269665|emb|CAB79613.1|  (AL161572) putative transcription factor MYB41 [Arabidopsis
thaliana]
          Length = 282

 Score = 32.7 bits (73), Expect = 0.91
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 69  RFSFEEEETIIQLHSVMGNKWSAIAARLPGRTDNEIKNHWNTHIR 113
```

```
>gi|15229407|ref|NP_188966.1|  (NM_113226) myb-related transcription factor, putative
[Arabidopsis thaliana]
 gi|7939535|dbj|BAA95738.1|  (AB025608) myb-related transcription factor [Arabidopsis
thaliana]
          Length = 285

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 20/65 (30%), Positives = 32/65 (48%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLNEKEKVRGV 211
           F+  E+  I+ L   +G  +++I    PG+T N IKN ++  L+   ED    K K
Sbjct: 70  FTKEEEDAIISLHQILGNRWSAIAAKLPGRTDNEIKNVWHTHLKKRLEDYQPAKPKTSNK 129

Query: 212 RKQRK 226
           +K  K
Sbjct: 130 KKGTK 134




>gi|11358530|pir||T51639  myb-related transcription factor MYB18 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941426|gb|AAC83589.1|  (AF062867) putative transcription factor [Arabidopsis thaliana]
          Length = 232

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 17/58 (29%), Positives = 32/58 (54%), Gaps = 2/58 (3%)
 Frame = +2

Query: 35  SYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYK--KLRYMKEDDLNEKEKV 202
           S  E++ IL     +G  ++ I K+ PG+T N IKN ++   K +++K   L + + +
Sbjct: 18  SAEEEETILTFHSPLGNKWSQIAKFLPGRTDNEIKNYWHSHLKKKWLKSQSLQDAKSI 75




>gi|15234199|ref|NP_195071.1|  (NM_119499) putative transcription factor [Arabidopsis
thaliana]
 gi|7446188|pir||T05996  hypothetical protein F17M5.210 - Arabidopsis thaliana
 gi|4490312|emb|CAB38803.1|  (AL035678) putative transcription factor [Arabidopsis thaliana]
 gi|7270293|emb|CAB80062.1|  (AL161583) putative transcription factor [Arabidopsis thaliana]
          Length = 250

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 15/52 (28%), Positives = 30/52 (56%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLN 187
           F+  E++R+L+     G  + SI + FPG+T N +KN ++  +   K ++ +
Sbjct: 74  FTEEEEERLLKAHRIQGNRWASIARLFPGRTDNAVKNHFHVIMARRKRENFS 125




>gi|15230123|ref|NP_189640.1|  (NM_113920) myb-like transcription factor, putative
[Arabidopsis thaliana]
 gi|9294205|dbj|BAB02107.1|  (AB028615) MYB-related transcription factor-like protein
[Arabidopsis thaliana]
 gi|14161421|gb|AAK54743.1|AF371980_1  (AF371980) putative transcription factor MYB121
[Arabidopsis thaliana]
```

```
          Length = 276

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 24/66 (36%), Positives = 35/66 (52%), Gaps = 3/66 (4%)
 Frame = +2

Query: 44  EDKRILELV**VGPNFNSIVKYFPGKTMNMIKN---RYYKKLRYMKEDDLNEKEKVRGVR 214
            E+  ILEL    G  +++I +Y PG+T N IKN    +YKK    K     +K K    R
Sbjct: 89  EEGIILELHSLWGNKWSTIARYLPGRTDNEIKNYWRTHYKK--NQKSSSKQDKVKKSLSR 146

Query: 215 KQRKAN 232
            KQ++ +
Sbjct: 147 KQQQVD 152




>gi|15232334|ref|NP_191605.1|  (NM_115910) myb protein - like [Arabidopsis thaliana]
 gi|11273995|pir||T47857  myb protein-like - Arabidopsis thaliana
 gi|7287994|emb|CAB81832.1|  (AL138646) myb protein-like [Arabidopsis thaliana]
 gi|18766549|gb|AAL79015.1|AF469468_1  (AF469468) putative transcription factor [Arabidopsis
thaliana]
          Length = 297

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 16/41 (39%), Positives = 23/41 (56%)
 Frame = +2

Query: 20  NDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKN 142
            N  KFS  E++ ++EL    G  +  I  Y PG+T N +KN
Sbjct: 63  NGCKFSADEERTVIELQSEFGNKWARIATYLPGRTDNDVKN 103




>gi|2832404|emb|CAA74603.1|  (Y14207) R2R3-MYB transcription factor [Arabidopsis thaliana]
          Length = 273

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 20/65 (30%), Positives = 32/65 (48%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMKEDDLNEKEKVRGV 211
            F+  E+  I+ L   +G  +++I    PG+T N IKN ++  L+   ED    K K
Sbjct: 70  FTKEEEDAIISLHQILGNRWSAIAAKLPGRTDNEIKNVWHTHLKKRLEDYQPAKPKTSNK 129

Query: 212 RKQRK 226
            +K  K
Sbjct: 130 KKGTK 134




>gi|15235149|ref|NP_194286.1|  (NM_118688) myb-like protein [Arabidopsis thaliana]
 gi|7446167|pir||T05791  myb-related protein homolog M7J2.70 - Arabidopsis thaliana
 gi|2980794|emb|CAA18170.1|  (AL022197) myb-like protein [Arabidopsis thaliana]
          Length = 283

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 17/58 (29%), Positives = 32/58 (54%), Gaps = 2/58 (3%)
 Frame = +2
```

```
Query: 35  SYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYK--KLRYMKEDDLNEKEKV 202
           S  E++ IL     +G  ++ I K+ PG+T N IKN ++    K +++K   L + + +
Sbjct: 69  SAEEEETILTFHSSLGNKWSQIAKFLPGRTDNEIKNYWHSHLKKKWLKSQSLQDAKSI 126




>gi|7269406|emb|CAB81366.1|  (AL161563) myb-like protein [Arabidopsis thaliana]
          Length = 282

 Score = 32.3 bits (72), Expect = 1.2
 Identities = 17/58 (29%), Positives = 32/58 (54%), Gaps = 2/58 (3%)
 Frame = +2

Query: 35  SYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYK--KLRYMKEDDLNEKEKV 202
           S  E++ IL     +G  ++ I K+ PG+T N IKN ++    K +++K   L + + +
Sbjct: 68  SAEEEETILTFHSSLGNKWSQIAKFLPGRTDNEIKNYWHSHLKKKWLKSQSLQDAKSI 125




>gi|7446177|pir||T07393  myb-related transcription factor - tomato
 gi|1430846|emb|CAA67600.1|  (X99210) myb-related transcription factor [Lycopersicon
esculentum]
          Length = 274

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 18/54 (33%), Positives = 30/54 (55%), Gaps = 4/54 (7%)
 Frame = +2

Query: 44  EDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEK 193
           E+K +++L    +G  ++ I   + PG+T N IKN +      KKL+ M   D +   K
Sbjct: 74  EEKMVIDLHAQLGNRWSKIASHLPGRTDNEIKNHWNTHIKKKLKKMGIDPITHK 127




>gi|15220910|ref|NP_173237.1|  (NM_101658) myb-like protein, putative [Arabidopsis thaliana]
          Length = 249

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 14/46 (30%), Positives = 28/46 (60%)
 Frame = +2

Query: 14  KLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYY 151
           ++N   F+  E++R+L      G  ++ I ++FPG+T N +KN ++
Sbjct: 54  RINRNPFTEEEEERLLASHRIHGNRWSVIARFFPGRTDNAVKNHWH 99




>gi|7438328|pir||T03828  myb protein - rice
 gi|1946267|emb|CAA72218.1|  (Y11415) myb [Oryza sativa (japonica cultivar-group)]
          Length = 368

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2
```

```
Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 69  RFSFEEEEAIIQLHSILGNKWSAIAARLPGRTDNEIKNYWNTHIR 113




>gi|11358559|pir||T51679  myb-related transcription factor MYB74 [imported] - Arabidopsis
thaliana (fragment)
 gi|3941506|gb|AAC83629.1|  (AF062907) putative transcription factor [Arabidopsis thaliana]
          Length = 259

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 5   RFSFEEEETIIQLHGIMGNKWSAIAARLPGRTDNEIKNYWNTHIR 49




>gi|9665072|gb|AAF97274.1|AC034106_17  (AC034106) Contains similarity to myb homologue from
Arabidopsis thaliana gb|D10936 and contains two Myb-like DNA-binding PF|00249 domains
          Length = 248

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 14/46 (30%), Positives = 28/46 (60%)
 Frame = +2

Query: 14  KLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYY 151
           ++N   F+  E++R+L      G  ++ I ++FPG+T N +KN ++
Sbjct: 53  RINRNPFTEEEEERLLASHRIHGNRWSVIARFFPGRTDNAVKNHWH 98




>gi|4886264|emb|CAB43399.1|  (AJ006292) Myb-related transcription factor mixta-like 1
[Antirrhinum majus]
          Length = 359

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 21/65 (32%), Positives = 34/65 (52%), Gaps = 4/65 (6%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEKE 196
           KFS  E++ I++L   +G  +++I  + P +T N IKN +     K+L  M  D +  K
Sbjct: 69  KFSLQEEQAIIQLHAFLGNRWSAIATHLPKRTDNEIKNYWNTHLKKRLTKMGIDPMTHKP 128

Query: 197 KVRGV 211
           K   V
Sbjct: 129 KSHDV 133




>gi|19073324|gb|AAL84760.1|AF474127_1  (AF474127) typical P-type R2R3 Myb protein [Sorghum
bicolor]
          Length = 157
```

```
 Score = 32.0 bits (71), Expect = 1.5
 Identities = 17/44 (38%), Positives = 25/44 (56%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           FS  ED+ IL L   +G  ++ I    PG+T N IKN +  +L+
Sbjct: 71  FSEEEDRIILSLYISIGSRWSIIAAQLPGRTDNDIKNYWNTRLK 114



>gi|15234593|ref|NP_192419.1|  (NM_116749) MYB - like protein [Arabidopsis thaliana]
 gi|7267269|emb|CAB81052.1|  (AL161502) MYB-like protein [Arabidopsis thaliana]
 gi|14423390|gb|AAK62377.1|AF386932_1  (AF386932) Unknown protein [Arabidopsis thaliana]
          Length = 324

 Score = 32.0 bits (71), Expect = 1.5
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 70  RFSFEEEETIIQLHSIMGNKWSAIAARLPGRTDNEIKNYWNTHIR 114



>gi|17065086|gb|AAL32697.1|  (AY062619) myb-related protein M4 [Arabidopsis thaliana]
          Length = 350

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 69  RFSFEEEETIIQLHSFLGNKWSAIAARLPGRTDNEIKNFWNTHIR 113



>gi|18415694|ref|NP_567626.1|  (NM_118264) myb-related protein M4 [Arabidopsis thaliana]
 gi|7242487|emb|CAB77384.1|  (X90381) putative transcription factor [Arabidopsis thaliana]
 gi|7263579|emb|CAB81661.1|  (X90382) putative transcription factor [Arabidopsis thaliana]
          Length = 350

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 69  RFSFEEEETIIQLHSFLGNKWSAIAARLPGRTDNEIKNFWNTHIR 113



>gi|2129563|pir||S58280  cM4 protein - Arabidopsis thaliana
          Length = 360
```

```
 Score = 31.6 bits (70), Expect = 2.0
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 69  RFSFEEEEAIIQLHSFLGNKWSAIAARLPGRTDNEIKNFWNTHIR 113



>gi|11358541|pir||T51660  myb-related transcription factor MYB52 [imported] - Arabidopsis
thaliana
 gi|3941468|gb|AAC83610.1|  (AF062888) putative transcription factor [Arabidopsis thaliana]
          Length = 249

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 14/46 (30%), Positives = 28/46 (60%)
 Frame = +2

Query: 14  KLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYY 151
           ++N   F+  E++R+L      G  ++ I ++FPG+T N +KN ++
Sbjct: 54  RINRNPFTEEEEERLLAPHRIHGNRWSVIARFFPGRTDNAVKNHWH 99



>gi|15238370|ref|NP_201326.1|  (NM_125921) transcription factor-like protein [Arabidopsis
thaliana]
 gi|10178185|dbj|BAB11659.1|  (AB013395) transcription factor-like protein [Arabidopsis
thaliana]
          Length = 310

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 17/45 (37%), Positives = 26/45 (57%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           KFS  E++ IL L   +G  ++ I  + PG+T N IKN +   L+
Sbjct: 69  KFSAEEEETILNLHAVLGNKWSMIASHLPGRTDNEIKNFWNTHLK 113



>gi|6467225|gb|AAF13101.1|U26934_1  (U26934) DNA-binding protein [Arabidopsis thaliana]
          Length = 386

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 16/45 (35%), Positives = 26/45 (57%)
 Frame = +2

Query: 8   KRKLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKN 142
           K  L    F+  E+KR+L+L   +G  ++ + +  PG+T N IKN
Sbjct: 66  KPSLKKGPFTDEEEKRVLQLHAVLGNKWSQMARELPGRTDNEIKN 110



>gi|19548457|gb|AAL90652.1|AF470084_1  (AF470084) P-type R2R3 Myb protein [Zea mays]
          Length = 210
```

```
 Score = 31.6 bits (70), Expect = 2.0
 Identities = 21/54 (38%), Positives = 30/54 (54%), Gaps = 4/54 (7%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDD 181
           F+  E K I++L   VG  ++ I   FPG+T N IKN +     K+LR M  D+
Sbjct: 71  FTSEEQKSIVQLHAIVGNKWSMIAATFPGRTDNEIKNYWNTHLKKQLRRMGLDE 124



>gi|7428545|pir||S58293  myb-related protein M4 - Arabidopsis thaliana
 gi|3080388|emb|CAA18708.1|  (AL022603) myb-related protein [Arabidopsis thaliana]
 gi|3402763|emb|CAA20209.1|  (AL031187) myb-related protein M4 [Arabidopsis thaliana]
 gi|7268941|emb|CAB81251.1|  (AL161555) myb-related protein M4 [Arabidopsis thaliana]
          Length = 352

 Score = 31.6 bits (70), Expect = 2.0
 Identities = 15/45 (33%), Positives = 28/45 (61%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLR 163
           +FS+ E++ I++L   +G  +++I    PG+T N IKN +   +R
Sbjct: 71  RFSFEEEETIIQLHSFLGNKWSAIAARLPGRTDNEIKNFWNTHIR 115



>gi|15225824|ref|NP_180264.1|  (NM_128253) putative MYB family transcription factor
[Arabidopsis thaliana]
 gi|6598477|gb|AAF18615.1|  (AC005623) MYB family transcription factor [Arabidopsis
thaliana]
          Length = 427

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 15/47 (31%), Positives = 29/47 (60%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMK 172
           F+  E+KR++EL   +G  +  + +  PG+T N IKN +  +L+ ++
Sbjct: 78  FTEKEEKRVIELHALLGNKWARMAEELPGRTDNEIKNFWNTRLKRLQ 124



>gi|15290138|dbj|BAB63829.1|  (AP003627) putative Myb factor protein [Oryza sativa (japonica
cultivar-group)]
          Length = 221

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 16/48 (33%), Positives = 28/48 (58%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYYKKLRYMK 172
           + S  E++ +++L   +G  ++ I K  PG+T N IKN +   LR +K
Sbjct: 58  RMSPEEERMVVQLHAKLGNRWSRIAKSIPGRTDNEIKNYWRTHLRKLK 105
```

```
>gi|15222362|ref|NP_177115.1|  (NM_105625) MYB-family transcription factor, putative
[Arabidopsis thaliana]
 gi|6692252|gb|AAF24603.1|AC021046_1  (AC021046) myb-related transcription factor, putative;
43081-41930 [Arabidopsis thaliana]
 gi|7644366|gb|AAF65558.1|AF249308_1  (AF249308) putative transcription factor [Arabidopsis
thaliana]
 gi|12597789|gb|AAG60101.1|AC073178_12  (AC073178) MYB-family transcription factor, putative
[Arabidopsis thaliana]
          Length = 330

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 13/46 (28%), Positives = 27/46 (58%)
 Frame = +2

Query: 14  KLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYY 151
            ++N   F+  E++R+++     G  +  I + FPG+T N +KN ++
Sbjct: 156 RINRRAFTEEEEERLMQAHRLYGNKWAMIARLFPGRTDNSVKNHWH 201




>gi|13346188|gb|AAK19616.1|AF336283_1  (AF336283) GHMYB25 [Gossypium hirsutum]
          Length = 309

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 20/61 (32%), Positives = 33/61 (53%), Gaps = 4/61 (6%)
 Frame = +2

Query: 29  KFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEKE 196
            KFS  E++ I++L   +G  +++I  + P +T N IKN +     K+L  M  D +  K
Sbjct: 69  KFSLQEEQTIIQLHALLGNRWSAIATHLPKRTDNEIKNYWNTHLKKRLTKMGIDPVTHKP 128

Query: 197 K 199
            K
Sbjct: 129 K 129




>gi|7438352|pir||T02006  transcription factor MYB4 homolog T15B16.7 - Arabidopsis thaliana
 gi|3859598|gb|AAC72864.1|  (AF104919) contains similarity to Myb DNA-binding domains (Pfam:
PF00249, E=3.7e-27 N=3) [Arabidopsis thaliana]
          Length = 348

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 21/58 (36%), Positives = 29/58 (49%), Gaps = 4/58 (6%)
 Frame = +2

Query: 32  FSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRY----YKKLRYMKEDDLNEK 193
            FS  E+  I+EL   +G  ++ I     PG+T N IKN +     KKLR    D +  K
Sbjct: 82  FSQDEENLIIELHAVLGNRWSQIAAQLPGRTDNEIKNLWNSCLKKKLRLRGIDPVTHK 139




>gi|6979341|gb|AAF34434.1|AF172282_23  (AF172282) myb-like protein [Oryza sativa]
          Length = 267

 Score = 31.2 bits (69), Expect = 2.6
 Identities = 14/46 (30%), Positives = 27/46 (58%)
 Frame = +2
```

```
Query: 14  KLNDPKFSYSEDKRILELV**VGPNFNSIVKYFPGKTMNMIKNRYY 151
           ++N   F+  E++R+L      G  +  I ++FPG+T N +KN ++
Sbjct: 75  RINKRPFTEEEEERLLAAHRHHGNKWALIARHFPGRTDNAVKNHWH 120




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  May 8, 2002  5:26 AM
  Number of letters in database: 292,803,047
  Number of sequences in database:  933,889

Lambda     K      H
   0.318    0.135    0.401

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 105,190,956
Number of Sequences: 933889
Number of extensions: 1627622
Number of successful extensions: 6011
Number of sequences better than 10.0: 220
Number of HSP's better than 10.0 without gapping: 5913
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 6010
length of database: 292,803,047
effective HSP length: 54
effective length of database: 242,373,041
effective search space used: 5816952984
frameshift window, decay const: 50,  0.1
T: 12
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
```

7.1.19 *blast*p-Datenbankvergleich mit pAC

BLASTP 2.2.1 [Apr-13-2001]

```
Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

RID: 1011023914-23618-5727
```

Query= pAC

```
         (865 letters)


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           846,869 sequences; 266,854,569 total letters


If you have any problems or questions with the results of this search
please refer to the BLAST FAQs

Taxonomy reports
```

Distribution of 39 Blast Hits on the Query Sequence

```
                                                                   Score     E
Sequences producing significant alignments:                        (bits)  Value

gi|17229758|ref|NP_486306.1|  (NC_003272) adenylate cyclase ...    84  6e-15
gi|1169150|sp|P43524|CYAA_ANACY  ADENYLATE CYCLASE (ATP PYRO...    83  1e-14
gi|15639476|ref|NP_218926.1|  (NC_000919) adenylate cyclase ...    73  1e-11
gi|16265216|ref|NP_438008.1|  (NC_003078) putative adenylate...    73  1e-11
gi|17229396|ref|NP_485944.1|  (NC_003272) adenylate cyclase ...    71  4e-11
gi|1754640|dbj|BAA13999.1|  (D89624) adenylate cyclase [Anab...    70  1e-10
gi|15890956|ref|NP_356628.1|  (NC_003063) AGR_L_1679p [Agrob...    67  9e-10
gi|15965714|ref|NP_386067.1|  (NC_003047) PUTATIVE ADENYLATE...    67  1e-09
gi|15889837|ref|NP_355518.1|  (NC_003062) AGR_C_4673p [Agrob...    65  3e-09
gi|14209675|gb|AAK56849.1|AF292554_1  (AF292554) putative ad...    60  1e-07
gi|5737842|gb|AAD50121.1|AF153362_1  (AF153362) adenylyl cyc...    59  2e-07
gi|15966612|ref|NP_386965.1|  (NC_003047) PUTATIVE ADENYLATE...    58  5e-07
gi|16265162|ref|NP_437954.1|  (NC_003078) putative adenylate...    58  5e-07
gi|2575805|dbj|BAA22996.1|  (D49530) adenylate cyclase [Spir...    55  2e-06
gi|6014766|sp|P40137|CYAA_STIAU  ADENYLATE CYCLASE 1 (ATP PY...    53  1e-05
gi|15966204|ref|NP_386557.1|  (NC_003047) PUTATIVE ADENYLATE...    52  2e-05
gi|95175|pir||B35266  adenylate cyclase (EC 4.6.1.1) - Rhizo...    52  2e-05
gi|15964012|ref|NP_384365.1|  (NC_003047) ADENYLATE CYCLASE ...    52  2e-05
gi|16263053|ref|NP_435846.1|  (NC_003037) Probable adenylate...    52  3e-05
gi|17228613|ref|NP_485161.1|  (NC_003272) adenylate cyclase ...    51  6e-05
gi|16263995|ref|NP_436787.1|  (NC_003078) putative adenylate...    50  1e-04
gi|16330472|ref|NP_441200.1|  (NC_000911) adenylate cyclase ...    49  2e-04
```

```
gi|6166050|sp|P40138|CYAB_STIAU  ADENYLATE CYCLASE 2 (ATP PY...    49  2e-04
gi|15598413|ref|NP_251907.1|  (NC_002516) probable adenylate...    49  2e-04
gi|12060475|dbj|BAB20624.1|  (AB052847) adenylate cyclase [T...    47  0.001
gi|7470942|pir||T17197  adenylate cyclase homolog - Spirulin...    43  0.014
gi|13473854|ref|NP_105422.1|  (NC_002678) adenylate cyclase ...    41  0.063
gi|15608460|ref|NP_215836.1|  (NC_000962) hypothetical prote...    40  0.081
gi|15840773|ref|NP_335810.1|  (NC_002755) adenylate cyclase,...    40  0.15
gi|17561794|ref|NP_505650.1|  (NM_073249) guanylate cyclase ...    39  0.18
gi|13472842|ref|NP_104409.1|  (NC_002678) contains partial s...    37  0.84
gi|1754642|dbj|BAA14000.1|  (D89625) adenylate cyclase [Anab...    37  0.86
gi|17232455|ref|NP_489003.1|  (NC_003272) adenylate cyclase ...    37  1.1
gi|15888487|ref|NP_354168.1|  (NC_003062) AGR_C_2127p [Agrob...    36  2.0
gi|9966799|ref|NP_065101.1|  (NM_020368) disrupter of silenc...    35  4.7
gi|6321948|ref|NP_012024.1|  (NC_001140) Establishes Silent ...    35  4.9
gi|13874614|dbj|BAB46914.1|  (AB060925) hypothetical protein...    34  5.9
gi|13272276|gb|AAK17065.1|  (AF244123) fibronectin binding a...    33  8.5
gi|17509149|ref|NP_492191.1|  (NM_059790) T22C1.5.p [Caenorh...    33  9.6

                                  Alignments


>gi|17229758|ref|NP_486306.1| (NC_003272) adenylate cyclase [Nostoc sp. PCC 7120]
 gi|15553050|dbj|BAA13998.2| (D89623) adenylate cyclase [Anabaena sp.]
 gi|17131357|dbj|BAB73965.1| (AP003588) adenylate cyclase [Nostoc sp. PCC 7120]
          Length = 859

 Score = 84.0 bits (206), Expect = 6e-15
 Identities = 61/215 (28%), Positives = 107/215 (49%), Gaps = 27/215 (12%)

Query: 525 DPMLPGKRKWAIYGFCDIRNFTDATEVL-QKDVMLFVNNIAEIVHSMVDRYQGSANKNIG 583
           D ++ G+RK     F DIR +T  TE L   +V+  +N   E +   V  Y+G+ +K IG
Sbjct: 590 DALMVGERKEVTVLFSDIRGYTTLTENLGAAEVVSLLNQYFETMVEAVFNYEGTLDKFIG 649

Query: 584 DAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDE 643
           DA + V                     F   L      A+  + + ++   ++ E+ N
Sbjct: 650 DALMAV---------------------FGAPLPLTENHAWQAVQSALDMRQRLKEF-NQR 687

Query: 644 RLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVL 703
           R+ Q  P  ++K+G G+  G  + G IGS  ++D + +   VN++SRLE  TK+YG  ++
Sbjct: 688 RIIQAQP--QIKIGIGISSGEVVSGNIGSHKRMDYTVIGDGVNLSSRLETVTKEYGCDII 745

Query: 704 ISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLF 738
           +S   +Q  S+ +  + RQ+DK+ VKG  + + ++
Sbjct: 746 LSEFTYQLCSDRI--WVRQLDKIRVKGKHQAVNIY 778




>gi|1169150|sp|P43524|CYAA_ANACY ADENYLATE CYCLASE (ATP PYROPHOSPHATE-LYASE) (ADENYLYL
CYCLASE)
 gi|2126532|pir||I39600 adenylate cyclase (EC 4.6.1.1) - Anabaena cylindrica
 gi|871829|dbj|BAA09511.1| (D55650) adenylate cyclase [Anabaena cylindrica]
          Length = 502

 Score = 83.2 bits (204), Expect = 1e-14
 Identities = 64/209 (30%), Positives = 103/209 (48%), Gaps = 36/209 (17%)

Query: 539 FCDIRNFTDATEVLQK-DVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           FCDIR +T  +E ++  ++  F+N+    +   +D    G  +K IGDA + +
Sbjct: 323 FCDIRGYTSMSEAMEPIEIFRFLNDYLACMGKAIDEAGGFIDKYIGDAIMAL-------- 374
```

```
Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQR----LPGYK 653
           ++DGN              DCAL A + +        + L+  NDER  Q     LP  +
Sbjct: 375 FDDGN-------------TDCALHAAILM-------QQALDKFNDERSMQTGKTGLP--R 412

Query: 654 VKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFS 713
           + +G G+H G  + G +G   +ID++ +   VN+ASR+E  TKQYG  +LI+  + +  S
Sbjct: 413 ISVGIGIHRGTVVMGTVGFTSRIDSTVIGDAVNVASRIEGLTKQYGCNILITESVVRNLS 472

Query: 714 NEVKKYTRQIDK-VTVKGSVKPIGLFTVE 741
                  R IDK V VKG  + I ++ V+
Sbjct: 473 CPESFSLRLIDKSVKVKGKDEAISIYEVK 501




>gi|15639476|ref|NP_218926.1| (NC_000919) adenylate cyclase [Treponema pallidum]
 gi|7520886|pir||C71320 probable adenylate cyclase - syphilis spirochete
 gi|3322767|gb|AAC65466.1| (AE001224) adenylate cyclase [Treponema pallidum]
          Length = 614

 Score = 72.8 bits (177), Expect = 1e-11
 Identities = 57/223 (25%), Positives = 107/223 (47%), Gaps = 26/223 (11%)

Query: 528 LPGKRKWAIYGFCDIRNFTDATEVL-QKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAF 586
           L G+RK A   F D+R+FT+ +E L  +DV+ F+N     +   +++  G  +K IGDA
Sbjct: 385 LGGERKTATIFFSDVRSFTEMSEKLPPEDVVEFLNEYMSCMVDCIEQTGGVVDKFIGDAI 444

Query: 587 LLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLS 646
           + +W    S                   L    L A   + A       +++  N++R++
Sbjct: 445 MAIWGAPVS-------------------LGSARLDALQSMKAVFLMRESLIQL-NEKRVA 484

Query: 647 QRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISS 706
              P  ++ +G G++ G  + G IGS  +++ + +   VN ASR+EA  K +G   LIS
Sbjct: 485 CSKP--RIGIGCGVNTGSCVAGQIGSSKRMEYTVIGDAVNTASRIEALNKPFGTDFLISE 542

Query: 707 ELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLF-TVEMEADDLP 748
             ++   + +     ++   +TVKG  +P+ ++  + ++  D P
Sbjct: 543 NTYELVKDML--IVEKMPPITVKGKREPLNVYAAINLKGHDGP 583




>gi|16265216|ref|NP_438008.1| (NC_003078) putative adenylate cyclase protein [Sinorhizobium
meliloti]
 gi|15141356|emb|CAC49868.1| (AL603647) putative adenylate cyclase protein [Sinorhizobium
meliloti]
          Length = 649

 Score = 72.8 bits (177), Expect = 1e-11
 Identities = 58/215 (26%), Positives = 98/215 (44%), Gaps = 30/215 (13%)

Query: 528 LPGKRKWAIYGFCDIRNFTDATEVLQKD---VMLFVNNIAEIVHSMVDRYQGSANKNIGD 584
           L G+R+     FCD+R FT  +E ++ D   +   +N +   +   V   +G+ +K IGD
Sbjct: 401 LGGERRTLTVLFCDVRGFTTISEDMKDDPEGLTTLINRLLTPLSEAVLNRRGTIDKYIGD 460

Query: 585 AFLLVWK--INDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRND 642
             +  W   ++D             +     Q A   L A   + A++  E K
Sbjct: 461 CLMAFWNAPLDDP-----------DHAVHAVQAARDMLTALGDLNAELQAEAK------- 502

Query: 643 ERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAV 702
              +   P   +++G G++ G   + G +GS  + D S L   VN+ASRLE A+K YGV +
Sbjct: 503 ---AAGRPPKALRIGIGINTGECVVGNMGSARRFDYSALGDAVNLASRLEGASKDYGVPL 559
```

```
Query: 703 LISSELHQYFSNEVKKYTRQIDKVTVKG--SVKPI 735
           L+        +   K    ++D++TVKG  +V P+
Sbjct: 560 LLGERTATLAAR--KFAVAELDRITVKGRSAVSPV 592



>gi|17229396|ref|NP_485944.1| (NC_003272) adenylate cyclase [Nostoc sp. PCC 7120]
 gi|17130994|dbj|BAB73603.1| (AP003587) adenylate cyclase [Nostoc sp. PCC 7120]
          Length = 860

 Score = 71.2 bits (173), Expect = 4e-11
 Identities = 64/254 (25%), Positives = 121/254 (47%), Gaps = 30/254 (11%)

Query: 525 DPMLPGKRKWAIYGFCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIG 583
           D  L G RK     F DIR +T  TE L+ ++V+  +N   E +   V +++G+ +K IG
Sbjct: 590 DAKLGGDRKEVSILFSDIRGYTTLTENLEAEEVVSMLNEYFESMVEAVFKHKGTLDKYIG 649

Query: 584 DAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDE 643
           DA + V+          G+ +                 A+M +   I    ++ E+ N +
Sbjct: 650 DAIMAVF----------GSPLPLEEH------------AWMAVKTSIEMRHRLQEF-NQK 686

Query: 644 RLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVL 703
           R +   P  ++ +G G++    I G IGS  +++ + +   VN+ SRLE+ +KQYG  ++
Sbjct: 687 RYAANKP--RINIGIGINSDTVISGNIGSSKRMEFTAIGDGVNLGSRLESVSKQYGCDII 744

Query: 704 ISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFT-VEMEADDLPPSK-QDYPQEEKQQ 761
           +S    +     +  + R++D + VKG  +P+ ++  + + +D +   K Q      K +
Sbjct: 745 LSDNTFKPCQENI--WARELDFIRVKGRNEPVSIYELIGLRSDPIASEKLQVIEHYHKGR 802

Query: 762 VMYEKKQIFLQQLE 775
             Y ++Q  L + E
Sbjct: 803 EYYLQRQFSLARAE 816



>gi|1754640|dbj|BAA13999.1| (D89624) adenylate cyclase [Anabaena sp.]
          Length = 860

 Score = 70.1 bits (170), Expect = 1e-10
 Identities = 56/215 (26%), Positives = 105/215 (48%), Gaps = 28/215 (13%)

Query: 525 DPMLPGKRKWAIYGFCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIG 583
           D  L G RK     F DIR +T  TE L+ ++V+  +N   E +   V +++G+ +K IG
Sbjct: 590 DAKLGGDRKEVSILFSDIRGYTTLTENLEAEEVVSMLNEYFESMVEAVFKHKGTLDKYIG 649

Query: 584 DAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDE 643
           DA + V+          G+ +                 A+M +   I    ++ E+ N +
Sbjct: 650 DAIMAVF----------GSPLPLEEH------------AWMAVKTSIEMRHRLQEF-NQK 686

Query: 644 RLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVL 703
           R +   P  ++ +G G++    I G IGS  +++ + +   VN+ SRLE+ +KQYG  ++
Sbjct: 687 RYAANKP--RINIGIGINSDTVISGNIGSSKRMEFTAIGDGVNLGSRLESVSKQYGCDII 744

Query: 704 ISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLF 738
           +S    +     +  + R++D + VKG  +P+ ++
Sbjct: 745 LSDNTFKPCQENI--WARELDFIRVKGRNEPVSIY 777
```

```
>gi|15890956|ref|NP_356628.1| (NC_003063) AGR_L_1679p [Agrobacterium tumefaciens]
[Agrobacterium tumefaciens str. C58 (Cereon)]
 gi|17937709|ref|NP_534498.1| (NC_003305) adenylate cyclase 1 [Agrobacterium tumefaciens
str. C58 (U. Washington)]
 gi|15159270|gb|AAK89413.1| (AE008282) AGR_L_1679p [Agrobacterium tumefaciens str. C58
(Cereon)]
 gi|17742456|gb|AAL44814.1| (AE009332) adenylate cyclase 1 [Agrobacterium tumefaciens str.
C58 (U. Washington)]
          Length = 414

 Score = 67.0 bits (162), Expect = 9e-10
 Identities = 54/209 (25%), Positives = 95/209 (44%), Gaps = 30/209 (14%)

Query: 540 CDIRNFTDATEVLQK-DVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQWY 598
           CD+RNFT  +++  + DV+  +N   + +   ++   G   K +GD  L ++ ++D +
Sbjct: 233 CDLRNFTHISDLWPRDDVIELLNGYFDAMCDPIEEAGGEILKFMGDGLLAIFPLSDPKAC 292

Query: 599 EDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMGF 658
           E        NL      A   L+A  +I  +   +P                GY    G
Sbjct: 293 E--------NLLKAIASAQKGLVALNEINCRKGHDPL---------------GY----GI 325

Query: 659 GLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNEVKK 718
           G+H+G  + G IGS+ ++D + + P VN+ASRLE  TK+ G  VL S E  +   N  +
Sbjct: 326 GVHVGDVMYGNIGSKTRLDFTVIGPAVNIASRLETLTKETGRNVLFSEEFVRMAGN--RD 383

Query: 719 YTRQIDKVTVKGSVKPIGLFTVEMEADDL 747
               +    ++G   P+ ++ +  E D +
Sbjct: 384 ALENLGPWLLRGLETPVEVYALPPETDTI 412



>gi|15965714|ref|NP_386067.1| (NC_003047) PUTATIVE ADENYLATE/GUANYLATE CYCLASE TRANSMEMBRANE
PROTEIN [Sinorhizobium meliloti]
 gi|15074983|emb|CAC46540.1| (AL591789) PUTATIVE ADENYLATE/GUANYLATE CYCLASE TRANSMEMBRANE
PROTEIN [Sinorhizobium meliloti]
          Length = 725

 Score = 66.6 bits (161), Expect = 1e-09
 Identities = 47/168 (27%), Positives = 80/168 (46%), Gaps = 24/168 (14%)

Query: 539 FCDIRNFTDATEVLQK-DVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           F D+RNFT+ +E L   +V+ F+N + + +   V   +G+ +K IGD+ +  W
Sbjct: 469 FVDVRNFTEISERLTPGEVVRFLNTLLDALSRHVIANEGTLDKFIGDSIMAFWNAPVDVA 528

Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
              G  ++             A +A  +  A++N E     + +D+         KV +G
Sbjct: 529 DHPGKAVR-------------AALAMRETLARLN-ESDAFGFGSDQ---------KVAIG 565

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLIS 705
            G+H G    G +G+E   + S +   VN+A+R+EAA K  G  +LIS
Sbjct: 566 VGIHTGLACVGNMGAEMHFNYSAVGDAVNIAARIEAACKDVGFDILIS 613



>gi|15889837|ref|NP_355518.1| (NC_003062) AGR_C_4673p [Agrobacterium tumefaciens]
[Agrobacterium tumefaciens str. C58 (Cereon)]
 gi|17936455|ref|NP_533245.1| (NC_003304) adenylate cyclase [Agrobacterium tumefaciens str.
C58 (U. Washington)]
```

```
 gi|15157775|gb|AAK88303.1| (AE008171) AGR_C_4673p [Agrobacterium tumefaciens str. C58
           (Cereon)]
 gi|17741075|gb|AAL43561.1| (AE009204) adenylate cyclase [Agrobacterium tumefaciens str. C58
(U. Washington)]
          Length = 564

 Score = 65.1 bits (157), Expect = 3e-09
 Identities = 51/184 (27%), Positives = 83/184 (44%), Gaps = 26/184 (14%)

Query: 523 DVDPMLPGKRKWAIYGFCDIRNFTDATEV-LQKDVMLFVNNIAEIVHSMVDRYQGSANKN 581
           D     PG+ +     FCD+RNFT  +E  L  DV+  +N    IV   V++  G  +K
Sbjct: 364 DAQSARPGREEEIAILFCDLRNFTTLSEAKLPYDVVFLLNRYFTIVGQAVEQADGHLDKF 423

Query: 582 IGDAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRN 641
           IGD  + ++ + +                      D A    MK  A I R+  IL
Sbjct: 424 IGDGAMALFGLGED--------------------TDRACRNAMKAAATILRDIAILNEGL 463

Query: 642 DERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVA 701
           +++ + RL     ++  G+H G  I G +G       + +   VN+ASRLE+  K++G A
Sbjct: 464 EKQFAVRL-----EVVVGIHSGPSIVGVMGYGAAKTLTAIGDTVNVASRLESKAKEFGAA 518

Query: 702 VLIS 705
           +++S
Sbjct: 519 IVVS 522
```

```
>gi|14209675|gb|AAK56849.1|AF292554_1 (AF292554) putative adenylate cyclase CyaA [Thiocapsa
roseopersicina]
          Length = 577

 Score = 60.1 bits (144), Expect = 1e-07
 Identities = 50/203 (24%), Positives = 86/203 (41%), Gaps = 32/203 (15%)

Query: 539 FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           F DIR FT   E +  +D + F+N     +   V    G  +  +GDA + V+
Sbjct: 327 FSDIRGFTSLVEGMAPEDAIDFINTYLARMEPAVLAEGGFIDNYVGDAIMAVF------- 379

Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
                          ++ AD ++ A + +   +       E+     L  + P   +++G
Sbjct: 380 ---------------DRGADASVRAGIAMLRSLG------EWTRPRELHGQSP---IRIG 415

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNEVK 717
            G+  G  I G IG+  ++    +   VN+A+R+E  TK YG+ +LI    +    +
Sbjct: 416 VGIATGEMIFGTIGAANRLKCGVVGDTVNLAARIEGLTKHYGLGLLIGGNTYASLDDPEC 475

Query: 718 KYTRQIDKVTVKGSVKPIGLFTV 740
              R+ID VTV G   P+ L+ V
Sbjct: 476 YLIREIDLVTVAGREAPVRLYEV 498
```

```
>gi|5737842|gb|AAD50121.1|AF153362_1 (AF153362) adenylyl cyclase [Dictyostelium discoideum]
          Length = 2123

 Score = 59.3 bits (142), Expect = 2e-07
 Identities = 54/207 (26%), Positives = 94/207 (45%), Gaps = 36/207 (17%)
```

```
Query: 539  FCDIRNFTDATE-VLQKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
            F DIR+FT  TE +L  DV+ F+N         +    G  +K IGDA + ++  +D
Sbjct: 1586 FSDIRDFTSTTEKMLVDDVIDFLNTYLAFALPSITDSGGFIDKFIGDAIMAIFPNSD--- 1642

Query: 598  YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYK---V 654
                                      MK+ A    +  I   R+ + +S  + G++   V
Sbjct: 1643 --------------------------MKLQAINAVKAAIRMMRSLDFMS--ISGFRFSSV 1674

Query: 655  KMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSN 714
            + G G++ G  I G +G+E +++ + L   VN+ASR E   K+Y   +LI+    +
Sbjct: 1675 ETGVGINTGKTIIGIVGTENRMEPTALGDAVNLASRTEQLCKEYQSRILITQFTMEAIGT 1734

Query: 715  EVKKYT-RQIDKVTVKGSVKPIGLFTV 740
             + ++  R +D VTVKG  + + ++ V
Sbjct: 1735 SIDEFVIRLVDSVTVKGKSEAVNIYEV 1761



>gi|15966612|ref|NP_386965.1| (NC_003047) PUTATIVE ADENYLATE/GUANYLATE CYCLASE PROTEIN
[Sinorhizobium meliloti]
 gi|15075884|emb|CAC47438.1| (AL591792) PUTATIVE ADENYLATE/GUANYLATE CYCLASE PROTEIN
[Sinorhizobium meliloti]
          Length = 423

 Score = 57.8 bits (138), Expect = 5e-07
 Identities = 55/210 (26%), Positives = 86/210 (40%), Gaps = 27/210 (12%)

Query: 535 AIYGFCDIRNFTD-ATEVLQKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKIN 593
           A+  F D+RN+T  A  V  +DV+ F+N+ + +V + V  + GS  K IGD  L ++
Sbjct: 236 AVMWFSDLRNYTRIADSVAPEDVIPFLNDYSGMVITAVHDHGGSVLKLIGDGVLAIFNGR 295

Query: 594 DSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYK 653
                          L                          +L   N  RL +  P
Sbjct: 296 KPAEACGAAIAAERQLRV------------------------MLSELNARRLDEGKPTTD 331

Query: 654 VKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFS 713
           V +  GLHIG    G IGS+ ++D + + P VN  SR+ A  +     V++SSE
Sbjct: 332 VYL--GLHIGEVFYGNIGSQDRLDFTVVGPAVNEVSRISAMCRSVERHVIMSSEFIAACP 389

Query: 714 NEVKKYTRQIDKVTVKGSVKPIGLFTVEME 743
            E +     + +  ++G  K   LFT + E
Sbjct: 390 VEHRANAVSLGRFALRGIAKAKELFTWDPE 419



>gi|16265162|ref|NP_437954.1| (NC_003078) putative adenylate cyclase protein [Sinorhizobium
meliloti]
 gi|15141302|emb|CAC49814.1| (AL603647) putative adenylate cyclase protein [Sinorhizobium
meliloti]
          Length = 420

 Score = 57.8 bits (138), Expect = 5e-07
 Identities = 44/168 (26%), Positives = 79/168 (46%), Gaps = 30/168 (17%)

Query: 540 CDIRNFTDATEVLQK-DVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQWY 598
           CD+R+FT  +++  + DV+  +N   + +   ++R+ G   K +GD  L ++ +++
Sbjct: 244 CDLRDFTKLSDLWPRDDVIELLNGYFDAMSEPIERHGGEILKFMGDGLLAIFPLSNP--- 300
```

```
Query: 599 EDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYK-VKMG 657
                              CA     ++   I      L   N+E L +   G+  +  G
Sbjct: 301 -------------------CAC---SQLLGAIGEAQAALAVLNEENLRK---GHDPLGYG 335

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLIS 705
            G+H+G  + G IGS  ++D + + P VN+ASRLE  TK+    VL+S
Sbjct: 336 IGVHVGDVMYGNIGSRRRLDFTVIGPAVNIASRLETLTKELKRPVLLS 383



>gi|2575805|dbj|BAA22996.1| (D49530) adenylate cyclase [Spirulina platensis]
          Length = 494

 Score = 55.5 bits (132), Expect = 2e-06
 Identities = 51/203 (25%), Positives = 91/203 (44%), Gaps = 32/203 (15%)

Query: 539 FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           F DIR+FT  +E +  +D    +N     +  +++ + G  +K IGD  + ++  +
Sbjct: 259 FSDIRDFTSISERISPQDNFNLINAYLSRMAPLINEHYGFIDKYIGDGIMALFGRSPDDA 318

Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
            +                A  ++++ +K Y K+  E            S  +P   +K+G
Sbjct: 319 VQ----------------AAISMLSSLKEYNKLRNE------------SGYIP---LKIG 347

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNEVK 717
            G++ G  I G +G    I+ + +S  VN+ASR+E  TK Y V +LIS        +
Sbjct: 348 IGINTGKLILGTVGYHNHIEGTVISDAVNLASRIEQLTKVYRVPLLISESTFVKLKHPHN 407

Query: 718 KYTRQIDKVTVKGSVKPIGLFTV 740
              R +DKV VKG  + + ++ V
Sbjct: 408 YGIRLVDKVQVKGRSEIVKVYEV 430



>gi|6014766|sp|P40137|CYAA_STIAU ADENYLATE CYCLASE 1 (ATP PYROPHOSPHATE-LYASE 1) (ADENYLYL
CYCLASE 1) (AC 1)
 gi|2828306|emb|CAA11549.1| (AJ223796) adenylate cyclase [Stigmatella aurantiaca]
          Length = 424

 Score = 53.1 bits (126), Expect = 1e-05
 Identities = 53/205 (25%), Positives = 92/205 (44%), Gaps = 33/205 (16%)

Query: 539 FCDIRNFTDATEVLQKD-VMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVW--KINDS 595
           F DIR+FT  +E L+ + V+  +N     +  +V R+ G+ +K IGDA ++ +   I D
Sbjct: 233 FADIRDFTSLSERLRPEQVVTLLNEYYGRMVEVVFRHGGTLDKFIGDALMVYFGAPIADP 292

Query: 596 QWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVK 655
                G                CAL    +           LE  N  R ++  P   ++
Sbjct: 293 AHARRG--------------VQCALDMVQE-----------LETVNALRSARGEPC--LR 325

Query: 656 MGFGLHIGWGIEGAIGSEF-KIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSN 714
           +G G+H G  + G IGS   +++ + +   VN+ASR+E+ TK   V +L S    +   +
Sbjct: 326 IGVGVHTGPAVLGNIGSATRRLEYTAIGDTVNLASRIESLTKTRDVPILASRATREQAGD 385

Query: 715 EVKKYTRQIDKVTVKGSVKPIGLFT 739
                  ++   +V G  +P+ +FT
Sbjct: 386 TF--LWNEMAPASVPGKSQPVAIFT 408
```

```
>gi|15966204|ref|NP_386557.1| (NC_003047) PUTATIVE ADENYLATE CYCLASE TRANSMEMBRANE PROTEIN
[Sinorhizobium meliloti]
 gi|15075474|emb|CAC47030.1| (AL591790) PUTATIVE ADENYLATE CYCLASE TRANSMEMBRANE PROTEIN
[Sinorhizobium meliloti]
          Length = 557

 Score = 52.4 bits (124), Expect = 2e-05
 Identities = 47/180 (26%), Positives = 78/180 (43%), Gaps = 26/180 (14%)

Query: 530 GKRKWAIYGFCDIRNFTDATE-VLQKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLL 588
           G+ +     FCD+R+FT   E  L  D +  +N   E+V   V+   G  +K IGD  L
Sbjct: 359 GRERRVAVLFCDLRDFTRIAEHRLPYDTVFLLNRYFEVVGEAVEGSGGVVDKFIGDGALA 418

Query: 589 VWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQR 648
           ++ +                 T + +    AL A     A++++  + L    +  L Q
Sbjct: 419 IFGLK----------------TPLPEACRQALSA----AARLSQGIRALNQTFEGELEQP 458

Query: 649 LPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSEL 708
           L     ++  GLH G  I G +G       + +   +N ASRLE   K++ V + +S+EL
Sbjct: 459 L-----RLAIGLHAGPAIIGEMGYGQATSLTVVGDTINTASRLEGLAKEHDVELAVSAEL 513



>gi|95175|pir||B35266 adenylate cyclase (EC 4.6.1.1) - Rhizobium meliloti  (fragment)
 gi|152153|gb|AAA26247.1| (M35096) adenylate cyclase (cyaA) [Sinorhizobium meliloti]
          Length = 193

 Score = 52.4 bits (124), Expect = 2e-05
 Identities = 47/205 (22%), Positives = 89/205 (42%), Gaps = 32/205 (15%)

Query: 539 FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWK--INDS 595
           F DI +FT  +E    ++V+  ++   ++   +V  + G+  +  GD+   +W   + D+
Sbjct: 2   FTDIYDFTTISEGRSPEEVVAMLSEYFDLFSEVVAAHDGTIIQFHGDSVFAMWNAPVADT 61

Query: 596 QWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVK 655
           +  E                  CAL    +           LE  N  + +  LP ++ +
Sbjct: 62  RHAEHA--------------CRCALAVEER-----------LEAFNSAQRASGLPEFRTR 96

Query: 656 MGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNE 715
             FG+H G  + G++G++ ++  + +   VN+ASRLE   K YG +VL S  +     +
Sbjct: 97  --FGIHTGTAVVGSVGAKERLQYTAMGDTVNVASRLEGMNKDYGTSVLASGAVVAQCKDM 154

Query: 716 VKKYTRQIDKVTVKGSVKPIGLFTV 740
           VK   R +     KG    + ++ V
Sbjct: 155 VK--FRPLGTAKAKGRSTALDIYEV 177



>gi|15964012|ref|NP_384365.1| (NC_003047) ADENYLATE CYCLASE 1 PROTEIN [Sinorhizobium
meliloti]
 gi|7404345|sp|P19485|CYA1_RHIME ADENYLATE CYCLASE 1 (ATP PYROPHOSPHATE-LYASE 1) (ADENYLYL
CYCLASE 1)
 gi|15073188|emb|CAC41696.1| (AL591783) ADENYLATE CYCLASE 1 PROTEIN [Sinorhizobium meliloti]
          Length = 665

 Score = 52.4 bits (124), Expect = 2e-05
 Identities = 47/205 (22%), Positives = 89/205 (42%), Gaps = 32/205 (15%)
```

```
Query: 539 FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWK--INDS 595
           F DI +FT  +E    ++V+  ++   ++   +V  + G+  +  GD+   +W   + D+
Sbjct: 474 FTDIYDFTTISEGRSPEEVVAMLSEYFDLFSEVVAAHDGTIIQFHGDSVFAMWNAPVADT 533

Query: 596 QWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVK 655
           +  E                  CAL    +           LE  N  + +  LP ++ +
Sbjct: 534 RHAEHA--------------CRCALAVEER-----------LEAFNSAQRASGLPEFRTR 568

Query: 656 MGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNE 715
             FG+H G  + G++G++ ++  + +   VN+ASRLE   K YG +VL S  +     +
Sbjct: 569 --FGIHTGTAVVGSVGAKERLQYTAMGDTVNVASRLEGMNKDYGTSVLASGAVVAQCKDM 626

Query: 716 VKKYTRQIDKVTVKGSVKPIGLFTV 740
           VK   R +     KG    + ++ V
Sbjct: 627 VK--FRPLGTAKAKGRSTALDIYEV 649



>gi|16263053|ref|NP_435846.1| (NC_003037) Probable adenylate cyclase [Sinorhizobium
meliloti]
 gi|14523709|gb|AAK65258.1| (AE007249) Probable adenylate cyclase [Sinorhizobium meliloti]
          Length = 584

 Score = 52.0 bits (123), Expect = 3e-05
 Identities = 47/191 (24%), Positives = 85/191 (43%), Gaps = 32/191 (16%)

Query: 518 MASSGDVDPMLPGKRKWAIYGFCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQG 576
           +    D+     G+R+     F DIR FT   E L    + +F+++    V    + + G
Sbjct: 383 LVGEDDLATWRQGRRQQGTILFIDIRGFTAYAEKLDPARLSIFISSFRRRVVRATEAFGG 442

Query: 577 SANKNIGDAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKI 636
             +K IGD  L+V+ I +                   Q  DC      +  A  +R   +
Sbjct: 443 VVDKFIGDGALVVFGIPEP------------------QSDDCT-----RAIACAHRLLAL 479

Query: 637 LEYRNDERLSQRLPGYK--VKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAA 694
           ++  N +R      G+   V++G G+H G    G +G + +++ + L   VN+A+++E A
Sbjct: 480 IDRWNIKR------GFDPPVRVGIGIHSGELYYGLVGDDHRLEFTVLGDTVNVAAKIEQA 533

Query: 695 TKQYGVAVLIS 705
           TK++  A+L S
Sbjct: 534 TKRFDTALLAS 544



>gi|17228613|ref|NP_485161.1| (NC_003272) adenylate cyclase [Nostoc sp. PCC 7120]
 gi|1754636|dbj|BAA13997.1| (D89622) adenylate cyclase [Anabaena sp.]
 gi|17130464|dbj|BAB73075.1| (AP003584) adenylate cyclase [Nostoc sp. PCC 7120]
          Length = 735

 Score = 50.8 bits (120), Expect = 6e-05
 Identities = 50/219 (22%), Positives = 103/219 (46%), Gaps = 40/219 (18%)

Query: 528 LPGKRKWAIYGFCDIRNFTDATEVLQKD-VMLFVNNIAEIVHSMVDRYQGSANKNIGDAF 586
           L G+R+     F D+R F+  +E L  + V+  +N    ++  ++++Y+G+ N+ IGD
Sbjct: 435 LGGERRKVTLLFSDLRGFSAMSEQLSPEQVVQILNLYLGVMTDVINQYKGTINEFIGDGI 494

Query: 587 LLVW-----KINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRN 641
            +++     + +DSQ                 +   CA IA  +   ++N +
Sbjct: 495 FIMFGAPICRPDDSQ-----------------RAIACA-IAMQRAMQQVNAKTH------ 530
```

```
Query: 642 DERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVA 701
                  LP  +++MG G++ G  + G IGS+ +   + +  +VN+A+R+E  T   G
Sbjct: 531 ----QMNLP--QLEMGIGINTGEVVAGNIGSQKRAQYTVIGSHVNLAARIETYT--VGGQ 582

Query: 702 VLISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFTV 740
           +LIS    Q  + +++   +   ++  KG  +PI ++ +
Sbjct: 583 ILISENTRQDANTDLQIAGQM--QIEPKGIKEPITIYEI 619



>gi|16263995|ref|NP_436787.1| (NC_003078) putative adenylate cyclase protein [Sinorhizobium
meliloti]
 gi|15140119|emb|CAC48647.1| (AL603642) putative adenylate cyclase protein [Sinorhizobium
meliloti]
          Length = 352

 Score = 49.7 bits (117), Expect = 1e-04
 Identities = 47/171 (27%), Positives = 77/171 (44%), Gaps = 26/171 (15%)

Query: 539 FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           + D+RNFT  +E  +   V+  +N+  + +   +  + G   K IGD  L ++ +ND
Sbjct: 176 YADLRNFTGLSEEHEPARVIAALNDWFDRLAGAIHAFGGEVLKFIGDGVLAIFPVND--- 232

Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
                           +  D AL A +   A        + +    RL Q LP   ++ G
Sbjct: 233 -------------LPAEACDRALRAVVAARAG-------MAHLEGLRLEQELP--PLRFG 270

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSEL 708
             LH+G  + G +GS  ++D + + P VN+ SRLE   +  G +VLIS  L
Sbjct: 271 AALHLGEILWGNVGSIDRLDFTAIGPAVNLVSRLEGLCRPLGRSVLISGAL 321



>gi|16330472|ref|NP_441200.1| (NC_000911) adenylate cyclase [Synechocystis sp. PCC 6803]
 gi|7469259|pir||S75018 adenylate cyclase - Synechocystis sp. (strain PCC 6803)
 gi|1652963|dbj|BAA17880.1| (D90910) adenylate cyclase [Synechocystis sp. PCC 6803]
          Length = 337

 Score = 49.3 bits (116), Expect = 2e-04
 Identities = 52/202 (25%), Positives = 92/202 (44%), Gaps = 31/202 (15%)

Query: 541 DIRNFTDATEVLQKDV--MLFVNNIAEIVHSMVDRYQGS-ANKNIGDAFLLVWKINDSQW 597
           D+RNFT   + +++++  ML  N   +  H +  R  GS  +K IGDA + +       W
Sbjct: 137 DMRNFTGMAQQVEEELLSMLIGNWFRQAGHIL--REAGSWVDKYIGDAVMAI-------W 187

Query: 598 YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
           +   NE   + +  I      A+     + AK+N++               LP + +++G
Sbjct: 188 FHGYNEATPAEIIQILH----AVNRLQAMTAKLNQK-------------YELP-FPLRIG 229

Query: 658 FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNEVK 717
            G++ G+ + G  GS    D + +   VN A RLE+ATKQ    + +S +   Y  +  +
Sbjct: 230 TGINTGYAMVGNTGSGDHPDYTAIGDTVNAAFRLESATKQAHFDLAMSEKTFSYLQDLPQ 289

Query: 718 -KYTRQIDKVTVKGSVKPIGLF 738
              T Q   + +KG   PI ++
Sbjct: 290 WSATVQQHTIELKGYTNPITIY 311
```

```
>gi|6166050|sp|P40138|CYAB_STIAU ADENYLATE CYCLASE 2 (ATP PYROPHOSPHATE-LYASE 2) (ADENYLYL
CYCLASE 2) (AC2)
 gi|7463015|pir||T10905 adenylate cyclase (EC 4.6.1.1) - Stigmatella aurantiaca
 gi|2828308|emb|CAA11548.1| (AJ223795) adenylate cyclase [Stigmatella aurantiaca]
          Length = 352

 Score = 49.3 bits (116), Expect = 2e-04
 Identities = 50/216 (23%), Positives = 96/216 (44%), Gaps = 23/216 (10%)

Query: 527 MLPGKRKWAIYGFCDIRNFTDATEVLQKD-VMLFVNNIAEIVHSMVDRYQGSANKNIGDA 585
           +L G++      F DIRNFT   E L  + V+  +N +   +   V    G+ +K +GD
Sbjct: 157 VLTGEKPEVTVLFADIRNFTGLAESLPPEQVVGVLNQVLGRLSDAVLTCGGTLDKFLGDG 216

Query: 586 FLLVWKINDSQWYEDGNEIKWSNLTFIN-QLADCALIAFMKIYAKINREPKILEYRNDER 644
            + VW          G  +  ++      Q A   + A +++     R+    E+  +ER
Sbjct: 217 LMAVW----------GAPVHRTDDALRALQAAKMMMTAMVEL-----RQAAQAEWAANER 261

Query: 645 LSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLI 704
           L + L    +++G G++ G  + G IG   + + + +   VN+A+RL A     G   ++
Sbjct: 262 LGRPL---VLELGIGINSGLAVAGNIGGSMRTEYTCIGDAVNVAARLCALA---GPGEIL 315

Query: 705 SSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFTV 740
           + E  +   +  +     +  V +KG  +P+ L+ V
Sbjct: 316 AGERTRELVSHREMPFEDLPPVRLKGKQQPVPLYRV 351
```

```
>gi|15598413|ref|NP_251907.1| (NC_002516) probable adenylate cyclase [Pseudomonas
aeruginosa]
 gi|11350916|pir||E83242 probable adenylate cyclase PA3217 [imported] - Pseudomonas
aeruginosa (strain PAO1)
 gi|9949338|gb|AAG06605.1|AE004745_2 (AE004745) probable adenylate cyclase [Pseudomonas
aeruginosa]
          Length = 463

 Score = 48.9 bits (115), Expect = 2e-04
 Identities = 54/231 (23%), Positives = 112/231 (48%), Gaps = 27/231 (11%)

Query: 518 MASSGDVDPMLPGKRKWAIYGFCDIRNFTDATEVLQKDVML-FVNNIAEIVHSMVDRYQG 576
           M  SG     L  +RK     F DIR FT+ +E L+ + +   +NN    +  +  +Y G
Sbjct: 211 MIFSGKKSVRLETQRKKLTVFFSDIRGFTELSEELEAEALTDLLNNYLNEMSKIALKYGG 270

Query: 577 SANKNIGDAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKI 636
           + +K +GD  ++ +    +Q  +             + +A  ++   M+ + K+ R+
Sbjct: 271 TIDKFVGDCVMVFFGDPSTQGAKK------------DAVAAVSMGIAMRKHMKVLRQ--- 315

Query: 637 LEYRNDERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATK 696
            ++R  + +++ L   +++M  G++ G+   G  G++ ++D + +   VN+ASRLE+A++
Sbjct: 316 -QWRA-QGITKPL---EIRM--GINTGYCTVGNFGADTRMDYTIIGREVNLASRLESASE 368

Query: 697 QYGVAVLISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFTVEMEADDL 747
                +LIS E +    + +    R   ++ VKG  +P+ ++ V     DL
Sbjct: 369 --AGEILISHETYSLIKDVI--MCRDKGQIAVKGFSRPVQIYQVVDSRRDL 415
```

```
>gi|12060475|dbj|BAB20624.1| (AB052847) adenylate cyclase [Thermosynechococcus elongatus]
          Length = 399

 Score = 46.6 bits (109), Expect = 0.001
 Identities = 47/176 (26%), Positives = 77/176 (43%), Gaps = 34/176 (19%)

Query: 541 DIRNFTDATEVLQKDVMLFVNNIAEIVH---SMVDRYQGSANKNIGDAFLLVWKINDSQW 597
           DIR+FT  T   Q +  L    I    H    ++ +Y    +K IGDA + VW I++S+
Sbjct: 203 DIRDFTQLTR--QLEEELLSELIGTWFHRAGEIIRQYGSWVDKYIGDAVMAVW-IHESE- 258

Query: 598 YEDGNEIKWSNLTF-INQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLP-GYKVK 655
                EI +  +   ++ L D  L+                       L Q+ P  + ++
Sbjct: 259 -----EIGYQEICHTLSALEDLRLMT--------------------GELHQQYPLPWPLR 293

Query: 656 MGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQY 711
           +G GL+ G+ + G  GS  + D + L   VN A RLE+ATK     + + +  + Y
Sbjct: 294 IGSGLNTGYAMVGNTGSGDRPDYTALGDTVNAAFRLESATKTIAADLAMGATTYHY 349
```

```
>gi|7470942|pir||T17197 adenylate cyclase homolog - Spirulina platensis
 gi|2575807|dbj|BAA22997.1| (D49692) adenylate cyclase [Spirulina platensis]
          Length = 1202

 Score = 42.7 bits (99), Expect = 0.014
 Identities = 49/198 (24%), Positives = 93/198 (46%), Gaps = 20/198 (10%)

Query: 519  ASSGDVDPMLPGKRKWAIYGFCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGS 577
            A++GD+   L  + +     F DI  FT  +  LQ + V   +N     +   V   QG+
Sbjct: 995  AAAGDLTLDLRPEPRLITILFSDIVGFTRMSNALQSQGVAELLNEYLGEMTRAVFENQGT 1054

Query: 578  ANKNIGDAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKIL 637
             +K +GDA + ++   +        +++ +  T     A   L+A  K+      + + L
Sbjct: 1055 VDKFVGDAIMALYGAPEEM--SPSEQVRRAIAT-----ARQMLVALEKLNQ--GWQERGL 1105

Query: 638  EYRNDERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQ 697
              RN+      +P   V+   G+H G  + G  GS+ + D + + P+VN+A+RL+ AT
Sbjct: 1106 VGRNE------VP--PVRFRCGIHQGMAVVGLFGSQERSDFTAIGPSVNIAARLQEATAP 1157

Query: 698  YGVAVLISSELHQYFSNE 715
               ++++S+ + QY  +E
Sbjct: 1158 N--SIMVSAMVAQYVPDE 1173
```

```
>gi|13473854|ref|NP_105422.1| (NC_002678) adenylate cyclase [Mesorhizobium loti]
 gi|14024605|dbj|BAB51208.1| (AP003004) adenylate cyclase [Mesorhizobium loti]
          Length = 458

 Score = 40.8 bits (94), Expect = 0.063
 Identities = 51/227 (22%), Positives = 96/227 (41%), Gaps = 30/227 (13%)

Query: 523 DVDPMLPGKRKWAIYGFCDIRNFTDATEVLQKD-VMLFVNNIAEIVHSMVDRYQGSANKN 581
           D +P  P +R+     F DI  FT  +E    + V   +      +  +V  + G+ +
Sbjct: 249 DDEPFGPVRRQDIAVLFADIVGFTHYSEDHPAEAVFELLRQFHRRMEQVVFDHHGTVDNY 308

Query: 582 IGDAFLLVWKINDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRN 641
           IGD  +  + +  +  ++D         T   Q A+  + A              LE  N
Sbjct: 309 IGDCIMATFGVPQAS-HDDA--------TRAIQCAEAMVAA--------------LEDWN 345
```

```
Query: 642 DERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVA 701
            +R+S+  P   V++  G   G  + GA+GSE  +  + +   VN+ASRL++  ++
Sbjct: 346 VQRVSRGYPPLDVRI--GAQYGAVVIGAVGSERNLSFAVVGDTVNVASRLQSLCRELQAN 403

Query: 702 VLISSELHQYFSNE---VKKYTRQIDKVTVKGSVKPIGLFTVEMEAD 745
           +   S L +    E    +   R    V+++G  +P+ ++ VE  A+
Sbjct: 404 ICFGSRLIEAAKAESPTTRLNARDHGPVSIRGRDEPVHVW-VERRAE 449



>gi|15608460|ref|NP_215836.1| (NC_000962) hypothetical protein Rv1320c [Mycobacterium
tuberculosis H37Rv]
 gi|1722970|sp|Q10633|YD20_MYCTU HYPOTHETICAL 61.9 KDA PROTEIN RV1320C
 gi|7476722|pir||D70769 hypothetical protein Rv1320c - Mycobacterium tuberculosis  (strain
H37RV)
 gi|1340086|emb|CAA98107.1| (Z73902) hypothetical protein Rv1320c [Mycobacterium
tuberculosis H37Rv]
          Length = 567

 Score = 40.4 bits (93), Expect = 0.081
 Identities = 54/244 (22%), Positives = 98/244 (40%), Gaps = 40/244 (16%)

Query: 514 IGTNMASSGDVD-PMLPGKRKWAIYGFCDIRNFTD-ATEVLQKDVMLFVNNIAEIVHSMV 571
           +G  +A++ + + P L G+ +     F DI   T   T     +V++ +N    ++   V
Sbjct: 338 VGREVAAAAERERPKLGGEERHVAVVFVDIVGSTQLVTSRPAAEVVMLLNRFFTVIVDEV 397

Query: 572 DRYQGSANKNIGDAFLLVWKI-NDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKI 630
           + ++G  NK  GDA L V+   N     ED
Sbjct: 398 NHHRGLVNKFQGDASLAVFGAPNRLSHPED------------------------------ 427

Query: 631 NREPKILEYRNDERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASR 690
                       +RL+  +P  + + G G+  G  + G +G+  + + + +   VN A+R
Sbjct: 428 --AALATARAIADRLASEMP--ECQAGIGVAAGQVVAGNVGAHERFEYTVIGEPVNEAAR 483

Query: 691 LEAATKQYGVAVLISSE-LHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFTVEMEADDLPP 749
           L    K Y   +L SS+ L     NE  +++   + VT++G  +PI L T  ++   +P
Sbjct: 484 LCELAKSYPSRLLASSQTLRGASENECARWSLG-ETVTLRGHDQPIRL-TSPVQQLQMPA 541

Query: 750 SKQD 753
              D
Sbjct: 542 QSAD 545



>gi|15840773|ref|NP_335810.1| (NC_002755) adenylate cyclase, putative [Mycobacterium
tuberculosis CDC1551]
 gi|13880966|gb|AAK45624.1| (AE007009) adenylate cyclase, putative [Mycobacterium
tuberculosis CDC1551]
          Length = 567

 Score = 39.7 bits (91), Expect = 0.15
 Identities = 51/228 (22%), Positives = 92/228 (39%), Gaps = 39/228 (17%)

Query: 514 IGTNMASSGDVD-PMLPGKRKWAIYGFCDIRNFTD-ATEVLQKDVMLFVNNIAEIVHSMV 571
           +G  +A++ + + P L G+ +     F DI   T   T     +V++ +N    ++   V
Sbjct: 338 VGREVAAAAERERPKLGGEERHVAVVFVDIVGSTQLVTSRPAAEVVMLLNRFFTVIVDEV 397

Query: 572 DRYQGSANKNIGDAFLLVWKI-NDSQWYEDGNEIKWSNLTFINQLADCALIAFMKIYAKI 630
           + ++G  NK  GDA L V+   N     ED
Sbjct: 398 NHHRGLVNKFQGDASLAVFGAPNRLSHPED------------------------------ 427
```

```
Query: 631 NREPKILEYRNDERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASR 690
                       +RL+  +P  + + G G+  G  + G +G+  + + + +   VN A+R
Sbjct: 428 --AALATARAIADRLASEMP--ECQAGIGVAAGQVVAGNVGAHERFEYTVIGEPVNEAAR 483

Query: 691 LEAATKQYGVAVLISSE-LHQYFSNEVKKYTRQIDKVTVKGSVKPIGL 737
           L    K Y   +L SS+ L     NE  +++   + VT++G  +PI L
Sbjct: 484 LCELAKSYPSRLLASSQTLRGASENECARWSLG-ETVTLRGHDQPIRL 530



>gi|17561794|ref|NP_505650.1| (NM_073249) guanylate cyclase  (2 domains) [Caenorhabditis
elegans]
 gi|7494714|pir||T18642 hypothetical protein B0024.6 - Caenorhabditis elegans
 gi|3873672|emb|CAA94879.1| (Z71178) contains similarity to Pfam domain: PF00069 (Eukaryotic
protein kinase domain), Score=41.9, E-value=5.1e-11, N=1; PF00211 (Adenylate and Guanylate
cyclase catalytic domain), Score=248.7, E-value=2.7e-71, N=2; PF01094 (Receptor family
ligand bi>
          Length = 1161

 Score = 39.3 bits (90), Expect = 0.18
 Identities = 38/181 (20%), Positives = 72/181 (38%), Gaps = 30/181 (16%)

Query: 539  FCDIRNFTD-ATEVLQKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
            F D+ +FT  A +     V+  +N +  I   +++++     + IGD + +   +
Sbjct: 954  FSDVVSFTTLAGKCTPLQVVNLLNGLYTIFDGIIEQHDVYKVETIGDGYFVASGVPRRNG 1013

Query: 598  YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
             E    I   ++ F+  LAD ++                            LPG K+K+
Sbjct: 1014 NEHTRNIASMSINFVKSLADFSI--------------------------PHLPGEKIKIR 1047

Query: 658  FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQ--YGVAVLISSELHQYFSNE 715
             G H G  + G +G             VN ASR+E+ +K+  +  +  ISS  H + + +
Sbjct: 1048 VGFHCGSVVAGVVGLTMP-RYCLFGDAVNTASRMESNSKRKIFQKSTNISSSFHLFSAGQ 1106

Query: 716  V 716
            +
Sbjct: 1107 I 1107



>gi|13472842|ref|NP_104409.1| (NC_002678) contains partial similarity to adenylate Cyclase
[Mesorhizobium loti]
 gi|14023589|dbj|BAB50195.1| (AP003001) contains partial similarity to adenylate cyclase
[Mesorhizobium loti]
          Length = 503

 Score = 37.0 bits (84), Expect = 0.84
 Identities = 24/78 (30%), Positives = 36/78 (45%), Gaps = 1/78 (1%)

Query: 635 KILEYRNDERLSQRLPGY-KVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEA 693
           K  E   D  L ++  G   +  G GLH+G  + G +G   ++  S     VN   RL+
Sbjct: 355 KATERMTDLNLRRKEQGLGDIGFGIGLHVGNVMFGNVGLTDRLTFSVFGSAVNEVQRLQT 414

Query: 694 ATKQYGVAVLISSELHQY 711
            TK+Y  +VL S +   Y
Sbjct: 415 LTKKYPHSVLASKDFASY 432
```

```
>gi|1754642|dbj|BAA14000.1| (D89625) adenylate cyclase [Anabaena sp.]
          Length = 1155

 Score = 37.0 bits (84), Expect = 0.86
 Identities = 43/178 (24%), Positives = 77/178 (43%), Gaps = 20/178 (11%)

Query: 539  FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
            F DI  FT     L+ + V   +N   E +   V    G+ +K +GDA L ++
Sbjct: 967  FSDIVGFTQLANTLRSRRVAELLNEYLEFMTKAVFDNGGTVDKFMGDAILALY------- 1019

Query: 598  YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
               G   + +    + +  + A  A     A++N+      +RN         G  V+
Sbjct: 1020 ---GAPEELTPNEQVRRAVNTAR-AMHSSLAQLNQ-----RWRNQGIFDTHDYG-GVQFR 1069

Query: 658  FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNE 715
             G+H G  + G  GS  + D + + P+VN+A+RL++A       +L+S+ +  Y  +E
Sbjct: 1070 CGIHQGTAVVGMFGSAERADYTAIGPSVNIAARLQSAA--VPGTILVSAAVADYLQDE 1125
```

```
>gi|17232455|ref|NP_489003.1| (NC_003272) adenylate cyclase carring two-component hybrid
sensor and regulator domains [Nostoc sp. PCC 7120]
 gi|17134101|dbj|BAB76662.1| (AP003598) adenylate cyclase carring two-component hybrid
sensor and regulator domains [Nostoc sp. PCC 7120]
          Length = 1155

 Score = 36.6 bits (83), Expect = 1.1
 Identities = 43/178 (24%), Positives = 77/178 (43%), Gaps = 20/178 (11%)

Query: 539  FCDIRNFTDATEVLQ-KDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQW 597
            F DI  FT     L+ + V   +N   E +   V    G+ +K +GDA L ++
Sbjct: 967  FSDIVGFTQLANTLRSRRVAELLNEYLEFMTKAVFDNGGTVDKFMGDAILALY------- 1019

Query: 598  YEDGNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMG 657
               G   + +    + +  + A  A     A++N+      +RN         G  V+
Sbjct: 1020 ---GAPEELTPNEQVRRAVNTAR-AMHSSLAQLNQ-----RWRNQGIFDTHDYG-GVQFR 1069

Query: 658  FGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNE 715
             G+H G  + G  GS  + D + + P+VN+A+RL++A       +L+S+ +  Y  +E
Sbjct: 1070 CGIHQGTAVVGMFGSAERADYTAIGPSVNIAARLQSAA--VPGTILVSAAVADYLQDE 1125
```

```
>gi|15888487|ref|NP_354168.1| (NC_003062) AGR_C_2127p [Agrobacterium tumefaciens]
[Agrobacterium tumefaciens str. C58 (Cereon)]
 gi|17935056|ref|NP_531846.1| (NC_003304) adenylate cyclase [Agrobacterium tumefaciens str.
C58 (U. Washington)]
 gi|13195123|gb|AAK13019.1| (AY026066) adenylate cyclase-like protein [Agrobacterium
tumefaciens]
 gi|15156187|gb|AAK86953.1| (AE008044) AGR_C_2127p [Agrobacterium tumefaciens str. C58
(Cereon)]
 gi|17739550|gb|AAL42162.1| (AE009078) adenylate cyclase [Agrobacterium tumefaciens str. C58
(U. Washington)]
          Length = 342

 Score = 35.8 bits (81), Expect = 2.0
 Identities = 30/104 (28%), Positives = 49/104 (46%), Gaps = 3/104 (2%)
```

```
Query: 637 LEYRNDERLSQRLPGYKVKMGFGLHIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATK 696
           +E + D  LS+   G   ++ F +H G  I   IG +     +Y    VN  SRLE+ +K
Sbjct: 234 MEEKRDMWLSRF--GEVPRLRFAIHGGPVITAEIGVDHH-KITYFGDTVNTTSRLESLSK 290

Query: 697 QYGVAVLISSELHQYFSNEVKKYTRQIDKVTVKGSVKPIGLFTV 740
             G  VLIS++L           +  + +  VKG  + +G+ T+
Sbjct: 291 MLGHPVLISADLAHQLVLPKGVRSEYLGEHAVKGRGQTLGVCTL 334



>gi|9966799|ref|NP_065101.1| (NM_020368) disrupter of silencing 10 [Homo sapiens]
 gi|9652121|gb|AAF91408.1|AF271212_1 (AF271212) disrupter of silencing SAS10 [Homo sapiens]
 gi|10439879|dbj|BAB15588.1| (AK026909) unnamed protein product [Homo sapiens]
 gi|13276683|emb|CAB66525.1| (AL136590) hypothetical protein [Homo sapiens]
 gi|13528699|gb|AAH04546.1|AAH04546 (BC004546) disrupter of silencing 10 [Homo sapiens]
          Length = 479

 Score = 34.7 bits (78), Expect = 4.7
 Identities = 23/76 (30%), Positives = 44/76 (57%), Gaps = 8/76 (10%)

Query: 222 KKENNTEQREGEDN----ANSQENRENQNMVAKQQQIDNSIPKV---QQIRRQSMKEPSQ 274
           K+EN+TE++  ED     A + +  +N+ +  ++++ID + P+V   ++ RR  ++   Q
Sbjct: 392 KEENSTEEQALEDQNAKRAITYQIAKNRGLTPRRKKIDRN-PRVKHREKFRRAKIRRRGQ 450

Query: 275 VKAKEKKESRVSKRLS 290
           V+   K+E R S  LS
Sbjct: 451 VREVRKEEQRYSGELS 466



>gi|6321948|ref|NP_012024.1| (NC_001140) Establishes Silent Chromatin; Esc4p [Saccharomyces
cerevisiae]
 gi|731729|sp|P38850|YHV4_YEAST HYPOTHETICAL 123.0 KD PROTEIN IN SPO16-REC104 INTERGENIC
REGION
 gi|626668|pir||S46755 hypothetical protein YHR154w - yeast (Saccharomyces cerevisiae)
 gi|500651|gb|AAB68978.1| (U10397) Yhr154wp [Saccharomyces cerevisiae]
          Length = 1070

 Score = 34.7 bits (78), Expect = 4.9
 Identities = 24/82 (29%), Positives = 41/82 (49%), Gaps = 3/82 (3%)

Query: 216 EQENLNKKENNTE-QREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQ 274
           ++ NL KK++N++   E E      E RE +       + D   PK Q+IR QS K+
Sbjct: 582 KKSNLQKKDSNSDISMETEVFCEGHEKREEKEFTKPITEYD--APKKQEIREQSRKKNDI 639

Query: 275 VKAKEKKESRVSKRLSDSTTKK 296
              KE++E+ +  +L   T ++
Sbjct: 640 DYKKEEEETELQVQLGQRTKRE 661



>gi|13874614|dbj|BAB46914.1| (AB060925) hypothetical protein [Macaca fascicularis]
          Length = 262

 Score = 34.3 bits (77), Expect = 5.9
 Identities = 23/76 (30%), Positives = 44/76 (57%), Gaps = 8/76 (10%)
```

```
Query: 222 KKENNTEQREGEDN----ANSQENRENQNMVAKQQQIDNSIPKV---QQIRRQSMKEPSQ 274
           K+EN+TE++  ED      A + +  +N+ +  ++++ID + P+V   ++ RR  ++    Q
Sbjct: 175 KEENSTEEQAVEDQNAKRAITYQIAKNRGLTPRRKKIDRN-PRVKHREKFRRAKIRRRGQ 233

Query: 275 VKAKEKKESRVSKRLS 290
           V+   K+E R S  LS
Sbjct: 234 VREVRKEEQRYSGELS 249



>gi|13272276|gb|AAK17065.1| (AF244123) fibronectin binding autolysin [Staphylococcus caprae]
          Length = 1395

 Score = 33.5 bits (75), Expect = 8.5
 Identities = 34/133 (25%), Positives = 56/133 (41%), Gaps = 24/133 (18%)

Query: 177 ASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQ-------------ENLNKK 223
           AS  ++     QP  ++ T  S    Y    ++I N  + +             +N NK+
Sbjct: 103 ASQQEQSNNSNQPTDSKETNTSSVNTYQDSTKDIKNNEISKDAVTSEDNETSKVDNTNKE 162

Query: 224 ENNTEQREGE-DNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKE 282
            + TEQ+E   +N N  EN  N N+ A   Q +N       +  Q  +E +  +A   K
Sbjct: 163 VSETEQQESTVENVNKDEN-SNNNIQASSNQTNN-------LNDQEKQEATSNQASNNKT 214

Query: 283 SRVSKRLSDSTTK 295
           S+  +  +DST K
Sbjct: 215 SQ--QATTDSTEK 225



>gi|17509149|ref|NP_492191.1| (NM_059790) T22C1.5.p [Caenorhabditis elegans]
 gi|7508201|pir||T25108 hypothetical protein T22C1.5 - Caenorhabditis elegans
 gi|3880035|emb|CAA99929.1| (Z75550) cDNA EST yk122c5.5 comes from this gene [Caenorhabditis
elegans]
          Length = 167

 Score = 33.5 bits (75), Expect = 9.6
 Identities = 18/73 (24%), Positives = 38/73 (51%), Gaps = 2/73 (2%)

Query: 218 ENLNKKENNTEQREGEDNANSQENRENQ--NMVAKQQQIDNSIPKVQQIRRQSMKEPSQV 275
           E L + +  T +++ ED   + EN+E +  N   ++ ++D   PK  Q++R+ +   S
Sbjct: 87  EKLRRMDEETAEQDAEDEKENDENKEPESPNRTLRKTKVDTPRPKTAQLKRKRIIMSSDD 146

Query: 276 KAKEKKESRVSKR 288
            ++++   R S R
Sbjct: 147 DSEDEFSMRSSPR 159




  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  Jan 14, 2002  1:01 AM
  Number of letters in database: 266,854,569
  Number of sequences in database:  846,869

Lambda     K      H
   0.314    0.132    0.366
```

```
Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 388,215,476
Number of Sequences: 846869
Number of extensions: 14913564
Number of successful extensions: 52987
Number of sequences better than 10.0: 385
Number of HSP's better than 10.0 without gapping: 106
Number of HSP's successfully gapped in prelim test: 294
Number of HSP's that attempted gapping in prelim test: 49944
Number of HSP's gapped (non-prelim): 1996
length of query: 865
length of database: 266,854,569
effective HSP length: 129
effective length of query: 736
effective length of database: 157,608,468
effective search space: 115999832448
effective search space used: 115999832448
T: 11
A: 40
X1: 16 ( 7.2 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (22.0 bits)
S2: 75 (33.5 bits)
```

Pfam 6.6 (St. Louis) : Home | Analyze a sequence | Browse alignments | Text search | Swisspfam | Help |

7.1.20 *Pfam*-Analyse pAC

Pfam HMM search results

[Go here for an explanation of the format of the results]

Model	Seq-from	Seq-to	HMM-from	HMM-to	Score	E-value	Description
guanylate_cyc	642	702	113	182	10.2	**0.0086**	Adenylate and Guanylate cyclase catalytic domain
TPR	802	835	1	34	5.8	0.57	TPR Domain
sodfe_C	817	830	100	113	2.7	0.75	Iron/manganese superoxide dismutases, C-terminal domain

```
Alignments of top-scoring domains:
guanylate_cyc: domain 1 of 1, from 642 to 702: score 10.2, E = 0.0086
                   *->qevnvghteggpPtvLyrglqlRiGIHtGpVvaGviGgaskrprYdv
                      +e+ +++++g         +++++G+H G  + G iG+   ++   +
       query   642    DERLSQRLPGYK-------VKMGFGLHIGWGIEGAIGS-EFKIDASY 680

                   wGdTVNvASRmEStGvpgekIhV<-*
                    +  VN+ASR+E+ ++++  ++V
       query   681 LSPNVNMASRLEAATKQY-GVAV    702

TPR: domain 1 of 1, from 802 to 835: score 5.8, E = 0.57
                   *->aeayynlGnaylklgkydeAiedyekALeldPnn<-*
                          +  G   + lg+++eA   +ekA +++Pn+
       query   802    FLHQFGQGFQGYLLGNWKEAHMLFEKAKQIRPND    835

sodfe_C: domain 1 of 1, from 817 to 830: score 2.7, E = 0.75
                   *->NWdevskRfeaAkk<-*
                      NW+e+   fe+Ak+
       query   817    NWKEAHMLFEKAKQ    830

//
```

Pfam 6.6 (St. Louis) : Home | Analyze a sequence | Browse alignments | Text search | Swisspfam | Help |

Comments, questions, flames? Email <pfam@genetics.wustl.edu>.

7.1.21 Smith/Waterman-Search mit pAC

Results of Search:

Program: bic2

Database: swall

Title: pAC (AS 1-532)

SeqLen: 533

```
OM of: EMBOSS_001 to: swall:*   out_format : pfs

Date: Fri Jan 18 15:03:47 2002

About: Results were produced by the GenCore software, version 5.0,
       Copyright (c) 1993-1999 Compugen Ltd.

Command line parameters:
-MODEL=sw.model -DEV=bic -Q=/net/nfs0/vol1/production/w3nobody/tmp/150498.bic2
-DB=swall:* -SUFFIX=pfs
-OUT=/net/nfs0/vol1/production/w3nobody/tmp/150498.bic2.res -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=150 -DOCALIGN=10 -THR_SCORE=quality -ALIGN=150
-MODE=LOCAL -OUTFMT=pfs -NONORM -HEAPSIZE=1000 -NCPU=6 -NEG_SCORES=0
-THREADS=1 -GAPOP=15 -GAPEXT=1 -BAND_WIDTH=100

Search information:
Query: EMBOSS_001
Query length: 521
Database: swall:*
Database sequences: 695814
Database length: 219930608
Search time (sec): 320

Score list:
Sequence        Strd Orig   Zscore Escore Len  !  Documentation  ..
swall:O77336      +  102.00 0.00   0.0   4550  !  O77336 PFC0425W PROTEIN. 6/2001
swall:Q9XSP9      +  100.00 0.00   0.0   273   !  Q9xsp9 VOLTAGE-GATED POTASSIUM CHANNEL
ISOFORM 4 (FRAGMENT). 6/2001
swall:Q9NGM0      +  98.00  0.00   0.0   1267  !  Q9ngm0 MYOSIN VI. 9/2001
swall:CIK1_MOUSE  +  97.00  0.00   0.0   495   !  P16388 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.1 (MK1) (MBK1). 7/1998
swall:CIK1_RAT    +  97.00  0.00   0.0   495   !  P10499 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.1 (IA) (RBK1) (RCK1). 7/1998
swall:Q9DAE0      +  97.00  0.00   0.0   315   !  Q9dae0 1700012K17RIK PROTEIN. 6/2001
swall:Q9D2A5      +  97.00  0.00   0.0   370   !  Q9d2a5 5330432F22RIK PROTEIN. 6/2001
swall:AAK73568    +  97.00  0.00   0.0   315   !  Aak73568 ATTACHES TO CRE. 9/2001
swall:CIK1_HUMAN  +  96.00  0.00   0.0   495   !  Q09470 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.1 (HUKI) (HBK1). 8/2001
swall:Q9DEY6      +  96.00  0.00   0.0   283   !  Q9dey6 PUTATIVE POTASSIUM CHANNEL KV1
(FRAGMENT). 9/2001
swall:Q9GPD5      +  95.00  0.00   0.0   771   !  Q9gpd5 ACP36DE (FRAGMENT). 6/2001
swall:Q9TUK5      +  95.00  0.00   0.0   274   !  Q9tuk5 VOLTAGE-GATED POTASSIUM CHANNEL
ISOFORM 5 (FRAGMENT). 6/2001
```

```
swall:Q09078      +  95.00  0.00  0.0  274  ! Q09078 POTASSIUM CHANNEL PROTEIN (CLONE KC4)
(FRAGMENT). 6/2001
swall:Q9N583      +  94.00  0.00  0.0  1724 ! Q9n583 Y110A7A.1 PROTEIN. 10/2000
swall:Q9GPD8      +  94.00  0.00  0.0  773  ! Q9gpd8 ACP36DE (FRAGMENT). 6/2001
swall:Q9GPD7      +  94.00  0.00  0.0  770  ! Q9gpd7 ACP36DE (FRAGMENT). 6/2001
swall:Q9GPD4      +  94.00  0.00  0.0  772  ! Q9gpd4 ACP36DE (FRAGMENT). 6/2001
swall:Q9GPD3      +  94.00  0.00  0.0  770  ! Q9gpd3 ACP36DE (FRAGMENT). 6/2001
swall:AAK68883    +  94.00  0.00  0.0  1758 ! Aak68883 HYPOTHETICAL PROTEIN Y110A7A.1B.
10/2001
swall:Q9VWS5      +  93.00  0.00  0.0  1895 ! Q9vws5 CG15040 PROTEIN. 6/2001
swall:Q9VGN4      +  93.00  0.00  0.0   399 ! Q9vgn4 CG14703 PROTEIN. 5/2000
swall:Q9XSQ0      +  92.00  0.00  0.0  260  ! Q9xsq0 VOLTAGE-GATED POTASSIUM CHANNEL
ISOFORM 1 (FRAGMENT). 6/2001
swall:CEC1_CAEEL  +  91.00  0.00  0.0  304  ! P34618 PROTEIN CEC-1. 8/2001
swall:MLH_TETTH   +  90.00  0.00  0.0  633  ! P40631 MICRONUCLEAR LINKER HISTONE
POLYPROTEIN (MIC LH) [CONTAINS: MICRONUCLEAR LINKER HISTONE-ALPHA; MICRONUCLEAR LINKER
HISTONE-BETA; MICRONUCLEAR LINKER HISTONE-DELTA; MICRONUCLEAR LINKER HISTONE-GAMMA]. 8/2001
swall:O60983      +  90.00  0.00  0.0  1038 ! O60983 REP-LIKE. 5/2000
swall:Q9GPD6      +  90.00  0.00  0.0  770  ! Q9gpd6 ACP36DE (FRAGMENT). 6/2001
swall:Q98QA1      +  90.00  0.00  0.0  622  ! Q98qa1 LIPOPROTEIN. 9/2001
swall:Q16322      +  89.00  0.00  0.0  511  ! Q16322 CYCLIC GMP GATED POTASSIUM CHANNEL.
6/2001
swall:P91642      +  88.00  0.00  0.0  454  ! P91642 CDK8. 9/2001
swall:O96923      +  88.00  0.00  0.0  1087 ! O96923 GELSOLIN-RELATED PROTEIN. 6/2001
swall:Q9N4C7      +  88.00  0.00  0.0  252  ! Q9n4c7 HYPOTHETICAL PROTEIN Y76B12C.4.
6/2001
swall:O35788      +  87.00  0.00  0.0  1339 ! O35788 CYCLIC NUCLEOTIDE-GATED CHANNEL BETA
SUBUNIT. 9/2001
swall:O44948      +  86.00  0.00  0.0  253  ! O44948 C17F3.3 PROTEIN. 11/1998
swall:Q9U0J9      +  86.00  0.00  0.0  842  ! Q9u0j9 HYPOTHETICAL 94.5 KDA PROTEIN. 6/2001
swall:Q91781      +  86.00  0.00  0.0  489  ! Q91781 POTASSIUM CHANNEL. 6/2001
swall:CIK2_RABIT  +  85.00  0.00  0.0  273  ! Q09081 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.2 (KC22) (FRAGMENT). 7/1998
swall:Q9NQS7      +  85.00  0.00  0.0  919  ! Q9nqs7 INNER CENTROMERE PROTEIN INCENP.
10/2000
swall:O76853      +  85.00  0.00  0.0  420  ! O76853 SRF RELATED PROTEIN. 6/2001
swall:Q9U0K5      +  85.00  0.00  0.0  1111 ! Q9u0k5 HYPOTHETICAL 132.5 KDA PROTEIN.
6/2001
swall:BAB58163    +  85.00  0.00  0.0  154  ! Bab58163 HYPOTHETICAL 17.8 KDA PROTEIN.
6/2001
swall:AAK85503    +  85.00  0.00  0.0  679  ! Aak85503 HYPOTHETICAL PROTEIN Y55B1BR.3.
10/2001
swall:NPL3_HUMAN  +  84.00  0.00  0.0  506  ! Q99457 NUCLEOSOME ASSEMBLY PROTEIN 1-LIKE 3.
8/2001
swall:P137_HUMAN  +  84.00  0.00  0.0  649  ! Q14444 GPI-ANCHORED PROTEIN P137 (P137GPI).
8/2001
swall:Q9BV09      +  84.00  0.00  0.0  656  ! Q9bv09 MEMBRANE COMPONENT, CHROMOSOME 11,
SURFACE MARKER 1. 6/2001
swall:O77306      +  84.00  0.00  0.0  1338 ! O77306 PFC0105W PROTEIN. 9/2001
swall:Q9PP74      +  84.00  0.00  0.0  719  ! Q9pp74 HYPOTHETICAL PROTEIN CJ0849C. 6/2001
swall:P137_MOUSE  +  83.00  0.00  0.0  656  ! Q60865 GPI-ANCHORED PROTEIN P137 (P137GPI).
8/2001
swall:O60788      +  83.00  0.00  0.0  506  ! O60788 DJ32F7.1 (FRAGMENT). 6/2001
swall:Q9NW13      +  83.00  0.00  0.0  759  ! Q9nw13 CDNA FLJ10377 FIS, CLONE
NT2RM2001989, WEAKLY SIMILAR TO NUCLEOLAR PROTEIN NOP4. 6/2001
swall:Q26597      +  83.00  0.00  0.0  512  ! Q26597 POTASSIUM CHANNEL PROTEIN. 6/2001
swall:Q9T069      +  83.00  0.00  0.0  532  ! Q9t069 HYPOTHETICAL 59.1 KDA PROTEIN.
10/2000
swall:Q54718      +  83.00  0.00  0.0  558  ! Q54718 M PROTEIN. 6/2001
swall:Q54447      +  83.00  0.00  0.0  563  ! Q54447 GLUCAN-BINDING PROTEIN PRECURSOR
(GBP). 6/2001
```

```
swall:AAH13889   +  83.00  0.00  0.0  759  ! Aah13889 UNKNOWN (PROTEIN FOR MGC:11192).
10/2001
swall:Q9XYL2     +  82.00  0.00  0.0  757  ! Q9xyl2 DEVELOPMENT PROTEIN DG1124
(FRAGMENT). 6/2001
swall:Q9V3R1     +  82.00  0.00  0.0  912  ! Q9v3r1 ACCESSORY GLAND PROTEIN ACP36DE
PRECURSOR. 3/2001
swall:Q9TUK7     +  82.00  0.00  0.0  279  ! Q9tuk7 VOLTAGE-GATED POTASSIUM CHANNEL
ISOFORM 2 (FRAGMENT). 6/2001
swall:Q9L6B1     +  82.00  0.00  0.0  2110 ! Q9l6b1 PUTATIVE FILAMENTOUS HEMAGGLUTININ.
10/2000
swall:Q9L6B0     +  82.00  0.00  0.0  2095 ! Q9l6b0 PUTATIVE FILAMENTOUS HEMAGGLUTININ
(FRAGMENT). 10/2000
swall:Q9CPH9     +  82.00  0.00  0.0  3919 ! Q9cph9 PFHB2. 6/2001
swall:AAK77711   +  82.00  0.00  0.0  1279 ! Aak77711 ORF42. 9/2001
swall:RBP2_PLAVB +  81.00  0.00  0.0  1251 ! Q00799 RETICULOCYTE BINDING PROTEIN 2
(FRAGMENT). 10/1996
swall:Q9W4J3     +  81.00  0.00  0.0  630  ! Q9w4j3 CG3546 PROTEIN. 5/2000
swall:Q9N2M3     +  81.00  0.00  0.0  2867 ! Q9n2m3 RETICULOCYTE-BINDING PROTEIN 2.
3/2001
swall:Q9VT57     +  81.00  0.00  0.0  454  ! Q9vt57 CDK8 PROTEIN. 9/2001
swall:Q9LH98     +  81.00  0.00  0.0  2081 ! Q9lh98 GENOMIC DNA, CHROMOSOME 3, BAC CLONE:
T19N8. 10/2000
swall:Q99QZ4     +  81.00  0.00  0.0  619  ! Q99qz4 SA2437 PROTEIN. 9/2001
swall:BAB58806   +  81.00  0.00  0.0  619  ! Bab58806 HYPOTHETICAL 69.2 KDA PROTEIN.
6/2001
swall:AAK77753   +  81.00  0.00  0.0  2314 ! Aak77753 ORF84. 9/2001
swall:AAL06398   +  81.00  0.00  0.0  231  ! Aal06398 HYPOTHETICAL 27.1 KDA PROTEIN.
10/2001
swall:INVO_CANFA +  80.00  0.00  0.0  285  ! P18174 INVOLUCRIN. 2/1996
swall:PHPA_PLACH +  80.00  0.00  0.0  441  ! Q02752 ACIDIC PHOSPHOPROTEIN PRECURSOR (50
KDA ANTIGEN). 6/1994
swall:Q23847     +  80.00  0.00  0.0  720  ! Q23847 GLUTAMINE-ASPARAGINE RICH PROTEIN
(FRAGMENT). 11/1996
swall:Q26223     +  80.00  0.00  0.0  2269 ! Q26223 RHOPTRY PROTEIN. 11/1998
swall:Q9Y0C9     +  80.00  0.00  0.0  838  ! Q9y0c9 RAS INTERACTING PROTEIN RIPA. 11/1999
swall:Q9W3U2     +  80.00  0.00  0.0  1086 ! Q9w3u2 CG14431 PROTEIN. 6/2001
swall:Q9U9S7     +  80.00  0.00  0.0  2123 ! Q9u9s7 ADENYLYL CYCLASE. 9/2001
swall:O97271     +  80.00  0.00  0.0  466  ! O97271 HYPOTHETICAL 54.8 KDA PROTEIN. 6/2001
swall:Q9NB35     +  80.00  0.00  0.0  1323 ! Q9nb35 MATURE PARASITE-INFECTED ERYTHROCYTE
SURFACE ANTIGEN. 6/2001
swall:Q9GSG6     +  80.00  0.00  0.0  757  ! Q9gsg6 AARDVARK. 6/2001
swall:Q28649     +  80.00  0.00  0.0  725  ! Q28649 CGMP-GATED POTASSIUM CHANNEL. 6/2001
swall:Q44782     +  80.00  0.00  0.0  218  ! Q44782 ERPB. 6/2001
swall:Q44791     +  80.00  0.00  0.0  338  ! Q44791 ERPD. 6/2001
swall:O31328     +  80.00  0.00  0.0  378  ! O31328 ERPJ. 6/2001
swall:O08264     +  80.00  0.00  0.0  378  ! O08264 ERPB2 PROTEIN. 6/2001
swall:Q9QR71     +  80.00  0.00  0.0  1129 ! Q9qr71 LATENT NUCLEAR ANTIGEN. 6/2001
swall:PTP3_DICDI +  79.00  0.00  0.0  989  ! P54637 PROTEIN-TYROSINE PHOSPHATASE 3 (EC
3.1.3.48) (PROTEIN-TYROSINE- PHOSPHATE PHOSPHOHYDROLASE 3). 11/1997
swall:Q9HI57     +  79.00  0.00  0.0  295  ! Q9hi57 HYPOTHETICAL PROTEIN TA1488. 3/2001
swall:P91570     +  79.00  0.00  0.0  312  ! P91570 COSMID ZK354. 11/1998
swall:O76719     +  79.00  0.00  0.0  335  ! O76719 F36H12.3 PROTEIN. 6/2001
swall:Q17909     +  79.00  0.00  0.0  385  ! Q17909 C11G6.3 PROTEIN. 6/2001
swall:Q9NG85     +  79.00  0.00  0.0  362  ! Q9ng85 CUT PROTEIN (FRAGMENT). 6/2001
swall:Q9GTW3     +  79.00  0.00  0.0  682  ! Q9gtw3 GLUTAMIC ACID-RICH PROTEIN. 3/2001
swall:Q97FK1     +  79.00  0.00  0.0  1163 ! Q97fk1 ATPASE INVOLVED IN DNA REPAIR. 9/2001
swall:YHV4_YEAST +  78.00  0.00  0.0  1070 ! P38850 HYPOTHETICAL 123.0 KDA PROTEIN IN
SPO16-REC104 INTERGENIC REGION. 2/1995
swall:MNN4_YEAST +  78.00  0.00  0.0  1178 ! P36044 MNN4 PROTEIN. 8/2001
swall:KAPC_DICDI +  78.00  0.00  0.0  648  ! P34099 CAMP-DEPENDENT PROTEIN KINASE
CATALYTIC SUBUNIT (EC 2.7.1.37). 5/2000
```

```
swall:O43988      +  78.00  0.00  0.0  800  ! O43988 HOMEOBOX-CONTAINING PROTEIN WARIAI
(FRAGMENT). 9/2001
swall:O00910      +  78.00  0.00  0.0  707  ! O00910 STAT PROTEIN. 6/2001
swall:O96133      +  78.00  0.00  0.0  1979 ! O96133 HYPOTHETICAL 237.7 KDA PROTEIN.
5/1999
swall:O96229      +  78.00  0.00  0.0  665  ! O96229 HYPOTHETICAL 78.6 KDA PROTEIN.
10/2000
swall:Q9VA29      +  78.00  0.00  0.0  578  ! Q9va29 CG12071 PROTEIN. 6/2001
swall:Q9TXB8      +  78.00  0.00  0.0  648  ! Q9txb8 SERINE/THREONINE PROTEIN KINASE.
9/2001
swall:O97242      +  78.00  0.00  0.0  1114 ! O97242 HYPOTHETICAL 132.2 KDA PROTEIN.
6/2001
swall:Q9GZ76      +  78.00  0.00  0.0  1387 ! Q9gz76 RETICULOCYTE-BINDING PROTEIN 2
HOMOLOG A (FRAGMENT). 3/2001
swall:Q06166      +  78.00  0.00  0.0  1661 ! Q06166 MATURE PARASITE-INFECTED ERYTHROCYTE
SURFACE ANTIGEN (ANTIGENIC PROTEIN PFEMP2). 6/2001
swall:Q9BK45      +  78.00  0.00  0.0  3254 ! Q9bk45 RETICULOCYTE BINDING PROTEIN 2
HOMOLOG B. 6/2001
swall:Q04111      +  78.00  0.00  0.0  891  ! Q04111 SURFACE EXCLUSION PROTEIN. 6/2001
swall:Q9S035      +  78.00  0.00  0.0  343  ! Q9s035 ERPB2 PROTEIN. 6/2001
swall:AAK54091    +  78.00  0.00  0.0  2062 ! Aak54091 HISTIDINE KINASE DHKJ. 6/2001
swall:AAK54092    +  78.00  0.00  0.0  1686 ! Aak54092 HISTIDINE KINASE DHKL (FRAGMENT).
6/2001
swall:AAK50002    +  78.00  0.00  0.0  1003 ! Aak50002 ORF73. 9/2001
swall:AAG02259    +  78.00  0.00  0.0  3254 ! Aag02259 RETICULOCYTE BINDING PROTEIN
ANALOG. 9/2001
swall:AAK84527    +  78.00  0.00  0.0  447  ! Aak84527 HYPOTHETICAL PROTEIN F40F4.8.
9/2001
swall:CIKS_DROME  +  77.00  0.00  0.0  656  ! P08510 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN SHAKER. 8/2001
swall:HMCU_DROME  +  77.00  0.00  0.0  2175 ! P10180 HOMEOBOX PROTEIN CUT. 8/2001
swall:CENC_HUMAN  +  77.00  0.00  0.0  943  ! Q03188 CENTROMERE PROTEIN C (CENP-C)
(CENTROMERE AUTOANTIGEN C). 8/2001
swall:Q25709      +  77.00  0.00  0.0  374  ! Q25709 10B ANTIGEN (FRAGMENT). 11/1998
swall:O61335      +  77.00  0.00  0.0  522  ! O61335 POTASSIUM CHANNEL. 6/2001
swall:Q22665      +  77.00  0.00  0.0  167  ! Q22665 HYPOTHETICAL 19.6 KDA PROTEIN T22C1.5
IN CHROMOSOME I. 1/1999
swall:O96201      +  77.00  0.00  0.0  1844 ! O96201 HYPOTHETICAL 214.7 KDA PROTEIN.
5/1999
swall:Q9NDS4      +  77.00  0.00  0.0  2678 ! Q9nds4 AMIB. 6/2001
swall:Q9ZU69      +  77.00  0.00  0.0  699  ! Q9zu69 PUTATIVE VICILIN STORAGE PROTEIN
(GLOBULIN-LIKE). 9/2001
swall:Q9FPT6      +  77.00  0.00  0.0  533  ! Q9fpt6 NOP58-LIKE PROTEIN F108. 6/2001
swall:Q9ZIU2      +  77.00  0.00  0.0  460  ! Q9ziu2 VIRULENT STRAIN ASSOCIATED
LIPOPROTEIN. 5/1999
swall:O50870      +  77.00  0.00  0.0  451  ! O50870 HYPOTHETICAL 54.3 KDA PROTEIN.
11/1998
swall:Q9DGL1      +  77.00  0.00  0.0  791  ! Q9dgl1 RETINITIS PIGMENTOSA GTPASE
REGULATOR-LIKE PROTEIN (FRAGMENT). 6/2001
swall:CIK2_HUMAN  +  76.00  0.00  0.0  499  ! P16389 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.2 (RBK2) (HBK5) (NGK1) (MK2) (HUKIV). 8/2001
swall:CIK2_MOUSE  +  76.00  0.00  0.0  499  ! P15386 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.2 (RBK2) (RCK5) (NGK1) (MK2) (HUKIV). 7/1998
swall:CIK4_MOUSE  +  76.00  0.00  0.0  654  ! Q61423 VOLTAGE-GATED POTASSIUM CHANNEL
PROTEIN KV1.4. 7/1998
swall:Q25920      +  76.00  0.00  0.0  1510 ! Q25920 MATURE-PARASITE-INFECTED ERYTHROCYTE
SURFACE ANTIGEN. 6/2001
swall:Q19974      +  76.00  0.00  0.0  754  ! Q19974 CODED FOR BY C. ELEGANS CDNA
YK69D11.5. 6/2001
swall:O01900      +  76.00  0.00  0.0  1621 ! O01900 CODED FOR BY C. ELEGANS CDNA
YK144H11.5. 6/2001
```

```
swall:O15763      +  76.00  0.00  0.0  1969 !  O15763 HYBRID HISTIDINE KINASE DHKB. 6/2001
swall:O96146      +  76.00  0.00  0.0  508  !  O96146 HYPOTHETICAL 61.3 KDA PROTEIN. 5/1999
swall:O96242      +  76.00  0.00  0.0  1351 !  O96242 HYPOTHETICAL 161.8 KDA PROTEIN.
5/1999
swall:O97140      +  76.00  0.00  0.0  1231 !  O97140 PRESPORE-SPECIFIC PROTEIN (FRAGMENT).
5/1999
swall:Q9Y114      +  76.00  0.00  0.0  656  !  Q9y114 BCDNA:GH10229 PROTEIN. 6/2001
swall:Q9VB01      +  76.00  0.00  0.0  1010 !  Q9vb01 CG12877 PROTEIN. 6/2001
swall:Q9U0K4      +  76.00  0.00  0.0  622  !  Q9u0k4 HYPOTHETICAL 73.3 KDA PROTEIN. 3/2001
swall:Q9GV71      +  76.00  0.00  0.0  1543 !  Q9gv71 MKPA PROTEIN. 6/2001
swall:Q28293      +  76.00  0.00  0.0  499  !  Q28293 DELAYED RECTIFIER K+ CHANNEL. 9/2001
swall:Q9MYX3      +  76.00  0.00  0.0  499  !  Q9myx3 POTASSIUM CHANNEL SUBUNIT KV 1.2.
9/2001
swall:Q09080      +  76.00  0.00  0.0  279  !  Q09080 POTASSIUM CHANNEL PROTEIN (CLONE
KC19) (FRAGMENT). 6/2001
swall:O31329      +  76.00  0.00  0.0  363  !  O31329 ERPM. 6/2001
swall:O40947      +  76.00  0.00  0.0  1089 !  O40947 ORF 73. 6/2001
swall:Q9DDN7      +  76.00  0.00  0.0  651  !  Q9ddn7 KV3.3 POTASSIUM CHANNEL SUBUNIT
PRECURSOR. 6/2001
swall:AAK54090    +  76.00  0.00  0.0  1736 !  Aak54090 HISTIDINE KINASE DHKI. 6/2001
swall:AAK72247    +  76.00  0.00  0.0  491  !  Aak72247 MB2 (FRAGMENT). 9/2001
swall:AAK72251    +  76.00  0.00  0.0  414  !  Aak72251 MB2 (FRAGMENT). 9/2001




Sequence name: swall:O77336

Sequence documentation:
ID   O77336      PRELIMINARY;      PRT;  4550 AA.
AC   O77336;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PFC0425W PROTEIN.
GN   PFC0425W.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O77336   ..

Alignment segment 1/1:

                     Quality:  102.00                     Escore:       0
             Matching length:     108               Total length:     108
 Matching Percent Similarity:   69.44  Matching Percent Identity:   23.15
    Total Percent Similarity:   69.44     Total Percent Identity:   23.15
                        Gaps:       0

Alignment:

                  .         .         .         .         .
     186 KRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDN 235
         ::: |: :  |::|     :  :|:||:  :::| |: :||:::::::::
    1902 RKEDIEKEKIKRGKYCHQEDRDDNMNNNDNNDNNNNNDNNNNNDNNNNND 1951
                  .         .         .         .         .
     236 ANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRV 285
          |:::|::|:|  : ::|: |:   | ::|:: |  | ::  : ::|| :
    1952 NNNNDNNNNDNNNSGEDQLINEDNKKRKINNLSTHEHTHNMSDHNKEPKK 2001
```

```
     286 SKRLSDST                                           293
         :|: |::|
    2002 KKNKSNET                                           2009



Sequence name: swall:Q9XSP9

Sequence documentation:
ID   Q9XSP9      PRELIMINARY;      PRT;   273 AA.
AC   Q9XSP9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL ISOFORM 4 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q9XSP9   ..

Alignment segment 1/1:

                     Quality:  100.00                      Escore:       0
             Matching length:      88                Total length:     108
 Matching Percent Similarity:   71.59   Matching Percent Identity:   28.41
    Total Percent Similarity:   58.33      Total Percent Identity:   23.15
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | ::  :|:| ||:|:::   |   :||:|::::  :|||| :  :
     135 DPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMNIIDIISIIPYL.. 182
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                        |   ::| | : |:: :: :: ::|::||:|: ::
     183 ...............ATLITELVQEDGAQRQQNMSLAILRIIRLVRVFRI 217

     174 YKAASYSQ                                           181
         :| : :|:
     218 FKLSRHSK                                           225



Sequence name: swall:Q9NGM0

Sequence documentation:
ID   Q9NGM0      PRELIMINARY;      PRT;  1267 AA.
AC   Q9NGM0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   MYOSIN VI.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.

Alignment of: EMBOSS_001 x Q9NGM0   ..
```

```
Alignment segment 1/1:

                     Quality:   98.00                      Escore:       0
             Matching length:      83                Total length:      83
 Matching Percent Similarity:   78.31   Matching Percent Identity:   24.10
    Total Percent Similarity:   78.31      Total Percent Identity:   24.10
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |:|:  | | :::: : |:: : | ::|:::::|:::||: :  |:::::
     959 EEERKLKAEMEAKRLREEEELRIQLEEEQKKLAAERKQIEADRQKLEEMQ 1008
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVI                  298
         ||: :|  : : :|::: ::::|||::::::||
    1009 RQAEEEQRRREEQERRDYELAQRLSEEANSQVI                  1041



Sequence name: swall:CIK1_MOUSE

Sequence documentation:
ID   CIK1_MOUSE     STANDARD;      PRT;   495 AA.
AC   P16388;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.1 (MK1) (MBK1).
GN   KCNA1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x CIK1_MOUSE   ..

Alignment segment 1/1:

                     Quality:   97.00                      Escore:       0
             Matching length:      97                Total length:     113
 Matching Percent Similarity:   70.10   Matching Percent Identity:   29.90
    Total Percent Similarity:   60.18      Total Percent Identity:   25.66
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      64 DDIRIIAFDKRADDGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFF 110
         |:  :|  ::   | | ::  :|:| ||:|:::   |   :||:|::::
     206 DNTTVIYTSNIFTDPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMN 255



                  .         .         .         .         .
     111 WLDIISTVSQILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIR 160
         ::||:: :  ::             ::::: |:| :: :  : ||   :|
     256 FIDIVAIIPYFI.............TLGTEIAEQEGNQKGEQATSLAILR 292
                  .
     161 VVRLVRLIRIVKL                                      173
         |:||||::|| ||
     293 VIRLVRVFRIFKL                                      305
```

```
Sequence name: swall:CIK1_RAT

Sequence documentation:
ID   CIK1_RAT       STANDARD;      PRT;   495 AA.
AC   P10499;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.1 (IA) (RBK1) (RCK1).
GN   KCNA1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Alignment of: EMBOSS_001 x CIK1_RAT   ..

Alignment segment 1/1:

                     Quality:   97.00                      Escore:       0
             Matching length:      97                Total length:     113
 Matching Percent Similarity:   70.10   Matching Percent Identity:   29.90
    Total Percent Similarity:   60.18      Total Percent Identity:   25.66
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      64 DDIRIIAFDKRADDGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFF 110
         |:  :|  ::   | | ::  :|:| ||:|:::   |   :||:|::::
     206 DNTTVIYTSNIFTDPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMN 255
                  .         .         .         .         .
     111 WLDIISTVSQILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIR 160
         ::||:: :  ::             ::::: |:| :: :  : ||   :|
     256 FIDIVAIIPYFI.............TLGTEIAEQEGNQKGEQATSLAILR 292
                  .
     161 VVRLVRLIRIVKL                                      173
         |:||||::|| ||
     293 VIRLVRVFRIFKL                                      305



Sequence name: swall:Q9DAE0

Sequence documentation:
ID   Q9DAE0      PRELIMINARY;      PRT;   315 AA.
AC   Q9DAE0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1700012K17RIK PROTEIN.
GN   1700012K17RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x Q9DAE0   ..

Alignment segment 1/1:

                     Quality:   97.00                      Escore:       0
```

```
             Matching length:     103               Total length:     116
 Matching Percent Similarity:   68.93   Matching Percent Identity:   29.13
    Total Percent Similarity:   61.21      Total Percent Identity:   25.86
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQN 246
         |::  |  :: ::|:|    |:: ::|::  :  |::  | |::|:: |:
     133 KAEERILKKIRRKIRNKQSAQDSRRRKKEYLDGLESRVAACSEQNQKLQR 182
                  .         .         .         .         .
     247 MVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK 296
                      |||:::||::   :||:  :|  :::|:| ::::|
     183 .............KVQELERQNIFLMEQVRQLQKLTAQTSSRAAQTSTCV 219
                  .
     297 VIILVILLLLIMPLFS                                   312
         :|:|  | |:|:| ||
     220 LILLFSLALIILPSFS                                   235



Sequence name: swall:Q9D2A5

Sequence documentation:
ID   Q9D2A5      PRELIMINARY;      PRT;   370 AA.
AC   Q9D2A5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   5330432F22RIK PROTEIN.
GN   5330432F22RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x Q9D2A5   ..

Alignment segment 1/1:

                     Quality:   97.00                     Escore:       0
             Matching length:     103               Total length:     116
 Matching Percent Similarity:   68.93   Matching Percent Identity:   29.13
    Total Percent Similarity:   61.21      Total Percent Identity:   25.86
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQN 246
         |::  |  :: ::|:|    |:: ::|::  :  |::  | |::|:: |:
     188 KAEERILKKIRRKIRNKQSAQDSRRRKKEYLDGLESRVAACSEQNQKLQR 237


                  .         .         .         .         .
     247 MVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK 296
                      |||:::||::   :||:  :|  :::|:| ::::|
     238 .............KVQELERQNIFLMEQVRQLQKLTAQTSSRAAQTSTCV 274
                  .
     297 VIILVILLLLIMPLFS                                   312
         :|:|  | |:|:| ||
     275 LILLFSLALIILPSFS                                   290
```

```
Sequence name: swall:AAK73568

Sequence documentation:
ID   AAK73568    PRELIMINARY;      PRT;   315 AA.
AC   AAK73568;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   ATTACHES TO CRE.
GN   ATCE1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x AAK73568   ..

Alignment segment 1/1:

                     Quality:   97.00                      Escore:       0
             Matching length:     103                Total length:     116
 Matching Percent Similarity:   68.93   Matching Percent Identity:   29.13
    Total Percent Similarity:   61.21      Total Percent Identity:   25.86
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQN 246
         |::  |  :: ::|:|    |:: ::|::  :  |::  | |::|:: |:
     133 KAEERILKKIRRKIRNKQSAQDSRRRKKEYLDGLESRVAACSEQNQKLQR 182
                  .         .         .         .         .
     247 MVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK 296
                      |||:::||::   :||:  :|  :::|:| ::::|
     183 .............KVQELERQNIFLMEQVRQLQKLTAQTSSRAAQTSTCV 219
                  .
     297 VIILVILLLLIMPLFS                                   312
         :|:|  | |:|:| ||
     220 LILLFSLALIILPSFS                                   235



Sequence name: swall:CIK1_HUMAN

Sequence documentation:
ID   CIK1_HUMAN     STANDARD;      PRT;   495 AA.
AC   Q09470;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.1 (HUKI) (HBK1).
GN   KCNA1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x CIK1_HUMAN   ..

Alignment segment 1/1:

                     Quality:   96.00                      Escore:       0
```

```
             Matching length:       97                Total length:     113
 Matching Percent Similarity:   70.10   Matching Percent Identity:   29.90
    Total Percent Similarity:   60.18      Total Percent Identity:   25.66
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      64 DDIRIIAFDKRADDGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFF 110
         |:  :|  ::   | | ::  :|:| ||:|:::   |   :||:|::::
     206 DNTTVIYNSNIFTDPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMN 255
                  .         .         .         .         .
     111 WLDIISTVSQILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIR 160
         ::||:: :  ::             ::::: |:| :: :  : ||   :|
     256 FIDIVAIIPYFI.............TLGTEIAEQEGNQKGEQATSLAILR 292
                  .
     161 VVRLVRLIRIVKL                                      173
         |:||||::|| ||
     293 VIRLVRVFRIFKL                                      305



Sequence name: swall:Q9DEY6

Sequence documentation:
ID   Q9DEY6      PRELIMINARY;      PRT;   283 AA.
AC   Q9DEY6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   PUTATIVE POTASSIUM CHANNEL KV1 (FRAGMENT).
OS   Pagothenia bernacchii (Emerald rockcod) (Trematomus bernacchii).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;

Alignment of: EMBOSS_001 x Q9DEY6   ..

Alignment segment 1/1:

                     Quality:   96.00                      Escore:       0
             Matching length:       97                Total length:     113
 Matching Percent Similarity:   70.10   Matching Percent Identity:   29.90
    Total Percent Similarity:   60.18      Total Percent Identity:   25.66
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      64 DDIRIIAFDKRADDGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFF 110
         |:  :|  ::   | | ::  :|:| ||:|:::   |   :||:|::::
     124 DNTTVIYNSNIFTDPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMN 173


                  .         .         .         .         .
     111 WLDIISTVSQILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIR 160
         ::||:: :  ::             ::::: |:| :: :  : ||   :|
     174 FIDIVAIIPYFI.............TLGTEIAEQEGNQKGEQATSLAILR 210
                  .
     161 VVRLVRLIRIVKL                                      173
         |:||||::|| ||
     211 VIRLVRVFRIFKL                                      223
```

```
Sequence name: swall:Q9GPD5

Sequence documentation:
ID   Q9GPD5      PRELIMINARY;      PRT;   771 AA.
AC   Q9GPD5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9GPD5   ..

Alignment segment 1/1:

                     Quality:   95.00                      Escore:       0
             Matching length:     132                Total length:     132
 Matching Percent Similarity:   69.70   Matching Percent Identity:   19.70
    Total Percent Similarity:   69.70      Total Percent Identity:   19.70
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :   :
      31 RVDKISPELQLQLLRYADSQSQSQSQSASQSESNASSQSQAQVQNNRLLE 80
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
           :   :|:| :::::: :| |::::::|::::::|::: :||||::::
      81 NPTVSESQQNAESQSETQSQSESQNQSQSESQNQSQSQAQRQQQIQTQLQ 130
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSD                   291
          ::|::::| :: :  :|:: :::| |   |:
     131 ILRQLQQKSNEQSAAQSASQIQSQRQSDSQSN                   162



Sequence name: swall:Q9TUK5

Sequence documentation:
ID   Q9TUK5      PRELIMINARY;      PRT;   274 AA.
AC   Q9TUK5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL ISOFORM 5 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q9TUK5   ..

Alignment segment 1/1:

                     Quality:   95.00                      Escore:       0
```

```
             Matching length:      84                Total length:     100
 Matching Percent Similarity:   73.81   Matching Percent Identity:   32.14
    Total Percent Similarity:   62.00      Total Percent Identity:   27.00
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | ::  :|:| ||:|:::   |   :||:|:::: ::||:: :  ::
     135 DPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMNFIDIVAIIPYFI. 183
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                     ::::: |:| :: :  : ||   :||:||||::|| ||
     184 ............TLGTEIAEQEGNQRGEQATSLAILRVIRLVRVFRIFKL 221



Sequence name: swall:Q09078

Sequence documentation:
ID   Q09078      PRELIMINARY;      PRT;   274 AA.
AC   Q09078;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL PROTEIN (CLONE KC4) (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q09078   ..

Alignment segment 1/1:

                     Quality:   95.00                      Escore:       0
             Matching length:      84                Total length:     100
 Matching Percent Similarity:   73.81   Matching Percent Identity:   32.14
    Total Percent Similarity:   62.00      Total Percent Identity:   27.00
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | ::  :|:| ||:|:::   |   :||:|:::: ::||:: :  ::
     135 DPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMNFIDIVAIIPYFI. 183
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                     ::::: |:| :: :  : ||   :||:||||::|| ||
     184 ............TLGTEIAEQEGNQRGEQATSLAILRVIRLVRVFRIFKL 221


Sequence name: swall:Q9N583

Sequence documentation:
ID   Q9N583      PRELIMINARY;      PRT;  1724 AA.
AC   Q9N583;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Y110A7A.1 PROTEIN.
```

```
GN   Y110A7A.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x Q9N583   ..

Alignment segment 1/1:

                     Quality:   94.00                      Escore:       0
             Matching length:      98                Total length:      98
 Matching Percent Similarity:   68.37   Matching Percent Identity:   22.45
    Total Percent Similarity:   68.37      Total Percent Identity:   22.45
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     190 IRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQ 239
         ::: :|: |:  |    :   : | : : : :::|:: |::||||: :::
     407 VKSLSTELSSILISSFDFTAPDPGAIPNLDAEQNEEEDEEEEGEDEEEEE 456
                  .         .         .         .
     240 ENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSK   287
         ||::::  | :::| |:|  : :  ::::::: :: | |||| ::|::
     457 ENEQDDVAVKEEEQSDKSDEENDGDNEENVSKKKEEKKKEKKAKEVKE   504



Sequence name: swall:Q9GPD8

Sequence documentation:
ID   Q9GPD8      PRELIMINARY;      PRT;   773 AA.
AC   Q9GPD8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9GPD8   ..

Alignment segment 1/1:

                     Quality:   94.00                      Escore:       0
             Matching length:     131                Total length:     133
 Matching Percent Similarity:   67.94   Matching Percent Identity:   23.66
    Total Percent Similarity:   66.92      Total Percent Identity:   23.31
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :  :
      31 RVDKISPELQLQLLRYADSQSQSQSQSASQSESNASSQSQAQVQNN..RL 78
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
         ::|  |:::: | ::::  | ::::::::|:: |:||:  :|:| :::|
      79 LENPTVSESQQNAESQSETQSQSQSESQNQSQSESQNQSQSQSQRQQQIQ 128
```

```
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                 292
            || || ::: :: :| :: ::  |:| |||
     129 TQLQILRQLQQKSNEQSAAQSASQIQSQRQSDS                 161



Sequence name: swall:Q9GPD7

Sequence documentation:
ID   Q9GPD7      PRELIMINARY;      PRT;   770 AA.
AC   Q9GPD7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9GPD7   ..

Alignment segment 1/1:

                    Quality:   94.00                      Escore:       0
             Matching length:     131                Total length:     133
 Matching Percent Similarity:   67.94   Matching Percent Identity:   23.66
    Total Percent Similarity:   66.92      Total Percent Identity:   23.31
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :  :
      31 RVDKISPELQLQLLRYADSQSQSQSASQSESNASSQSQAQVQNN..RL 78
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
         ::|  |:::: | ::::  | ::::::::|:: |:||:  :|:| :::|
      79 LENPTVSESQQNAESQSETQSQSQSESQNQSQSESQNQSQSQSQRQQQIQ 128
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                 292
            || || ::: :: :| :: ::  |:| |||
     129 TQLQILRQLQQKSNEQSAAQSASQIQSQRQSDS                 161



Sequence name: swall:Q9GPD4

Sequence documentation:
ID   Q9GPD4      PRELIMINARY;      PRT;   772 AA.
AC   Q9GPD4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
```

```
Alignment of: EMBOSS_001 x Q9GPD4   ..

Alignment segment 1/1:

                     Quality:   94.00                      Escore:       0
             Matching length:     131                Total length:     133
 Matching Percent Similarity:   67.94   Matching Percent Identity:   23.66
    Total Percent Similarity:   66.92      Total Percent Identity:   23.31
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :  :
      31 RVDKISPELQLQLLRYADSQSQSQSQSASQSESNASSQSQAQVQNN..RL 78
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
         ::|  |:::: | ::::  | ::::::::|:: |:||:  :|:| :::|
      79 LENPTVSESQQNAESQSETQSQSQSESQNQSQSESQNQSQSQSQRQQQIQ 128
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                  292
            || || ::: :: :| :: ::  |:| |||
     129 TQLQILRQLQQKSNEQSAAQSASQIQSQRQSDS                  161



Sequence name: swall:Q9GPD3

Sequence documentation:
ID   Q9GPD3      PRELIMINARY;      PRT;   770 AA.
AC   Q9GPD3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9GPD3   ..

Alignment segment 1/1:

                     Quality:   94.00                      Escore:       0
             Matching length:     131                Total length:     133
 Matching Percent Similarity:   67.94   Matching Percent Identity:   23.66
    Total Percent Similarity:   66.92      Total Percent Identity:   23.31
                        Gaps:       1

Alignment:



                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :  :
      31 RVDKISPELQLQLLRYADSQSQSQSQSASQSESNASSQSQAQVQNN..RL 78
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
         ::|  |:::: | ::::  | ::::::::|:: |:||:  :|:| :::|
      79 LENPTVSESQQNAESQSETQSQSQSESQNQSQSESQNQSQSQSQRQQQIQ 128
```

```
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                  292
            || || ::: :: :| :: ::  |:| |||
     129 TQLQILRQLQQKSNEQSAAQSASQIQSQRQSDS                  161



Sequence name: swall:AAK68883

Sequence documentation:
ID   AAK68883    PRELIMINARY;      PRT;  1758 AA.
AC   AAK68883;
DT   05-OCT-2001 (EMBLrel. 63, Created)
DT   05-OCT-2001 (EMBLrel. 63, Last sequence update)
DT   05-OCT-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL PROTEIN Y110A7A.1B.
GN   Y110A7A.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x AAK68883   ..

Alignment segment 1/1:

                     Quality:   94.00                      Escore:       0
             Matching length:      98                Total length:      98
 Matching Percent Similarity:   68.37   Matching Percent Identity:   22.45
    Total Percent Similarity:   68.37      Total Percent Identity:   22.45
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     190 IRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQ 239
         ::: :|: |:  |    :   : | : : : :::|:: |::||||: :::
     407 VKSLSTELSSILISSFDFTAPDPGAIPNLDAEQNEEEDEEEEGEDEEEEE 456
                  .         .         .         .
     240 ENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSK   287
         ||::::  | :::| |:|  : :  ::::::: :: | |||| ::|::
     457 ENEQDDVAVKEEEQSDKSDEENDGDNEENVSKKKEEKKKEKKAKEVKE   504



Sequence name: swall:Q9VWS5

Sequence documentation:
ID   Q9VWS5      PRELIMINARY;      PRT;  1895 AA.
AC   Q9VWS5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG15040 PROTEIN.
GN   CG15040.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9VWS5   ..

Alignment segment 1/1:
```

```
                     Quality:   93.00                     Escore:       0
             Matching length:     117               Total length:     118
 Matching Percent Similarity:   75.21   Matching Percent Identity:   20.51
    Total Percent Similarity:   74.58      Total Percent Identity:   20.34
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     180 SQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQ 229
         |::|: ::: |:::  :::: :|  |  || |:| :|::: |:::|| ::
    1476 SKDQSNSQDNIQSKDQSNNQINI.QSKNENNNQGNIESNDQNNNQNNIQS 1524
                  .         .         .         .         .
     230 REGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKE 279
         :: ::: | |:::||:|:   |:| :|:| ::::  ::: :| :  : :|
    1525 KNENNQNNIQRKNENNNQNNIQSQNENNIQSMNENSNHNNNENNIQSKNE 1574
                  .
     280 KKESRVSKRLSDSTTKKV                                297
         ::::::::: :::: :::
    1575 NNQNNIQSQNENNNQNNI                                1592



Sequence name: swall:Q9VGN4

Sequence documentation:
ID   Q9VGN4      PRELIMINARY;      PRT;   399 AA.
AC   Q9VGN4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG14703 PROTEIN.
GN   CG14703.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9VGN4   ..

Alignment segment 1/1:

                     Quality:   93.00                     Escore:       0
             Matching length:     109               Total length:     115
 Matching Percent Similarity:   71.56   Matching Percent Identity:   22.02
    Total Percent Similarity:   67.83      Total Percent Identity:   20.87
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     187 RQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNA 236
         ::|: | :: |:|    |   :: ::   |:|: :::|:::|::|:|:::
     138 KKPVATTSALKKKKEPTPPPESSSEEEESEEEEETEEETESEEEESEEES 187


                  .         .         .         .         .
     237 NSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVS 286
         :|:|:::::: :: :::: : : |: |:|      |: |: | : ||:::
     188 ESEEEESEEEDTTVDEKVPEHVKKAAQLR......PTTVNKKPEPESELT 231
                  .
     287 KRLSDSTTKKVIILV                                   301
         :|:: :  |   ::|
     232 QRFTMEKPKLRSVVV                                   246
```

```
Sequence name: swall:Q9XSQ0

Sequence documentation:
ID   Q9XSQ0      PRELIMINARY;      PRT;   260 AA.
AC   Q9XSQ0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL ISOFORM 1 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q9XSQ0   ..

Alignment segment 1/1:

                    Quality:   92.00                     Escore:       0
            Matching length:      84               Total length:     100
 Matching Percent Similarity:   72.62  Matching Percent Identity:   32.14
    Total Percent Similarity:   61.00     Total Percent Identity:   27.00
                       Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | ::  :|:| ||:|:::   |   :||:|:::: : ||:: :  ::
     121 DPFFIVETLCIIWFSFELVVRFFACPSKTDFFKNIMNFTDIVAIIPYFI. 169
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                     ::::: |:| :: :  : ||   :||:||||::|| ||
     170 ............TLGTEIAEQEGNQKGEQATSLAILRVIRLVRVFRIFKL 207



Sequence name: swall:CEC1_CAEEL

Sequence documentation:
ID   CEC1_CAEEL     STANDARD;      PRT;   304 AA.
AC   P34618;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROTEIN CEC-1.
GN   CEC-1 OR ZK1236.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x CEC1_CAEEL   ..

Alignment segment 1/1:

                    Quality:   91.00                     Escore:       0
            Matching length:     127               Total length:     127
 Matching Percent Similarity:   59.84  Matching Percent Identity:   25.20
    Total Percent Similarity:   59.84     Total Percent Identity:   25.20
```

```
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     196 KKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQ 245
         |||||    :| :: :  | :::::::: :: |: ||||: :||| :: :
     137 KKSKAEDEEEVEDDEEPVPKKKKEVQEEPEEEESVEGEDEEESQEVEDLK 186
                  .         .         .         .         .
     246 NMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTK 295
         :    :::  ::   ||   ::: || :: |::||||:: ::: ::
     187 EDEKMEEDEKEEEEDVQLESEKNEKEEEEEKVEEKKEEEEEEEEEEEIQLV 236
                  .         .
     296 KVIILVILLLLIMPLFSSDYYFEPSYS                         322
          |   ||   :: |  ::    ||| |
     237 IVEKTVIETTIVEPAVATPEPSEPSSS                         263



Sequence name: swall:MLH_TETTH

Sequence documentation:
ID   MLH_TETTH      STANDARD;      PRT;   633 AA.
AC   P40631;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MICRONUCLEAR LINKER HISTONE POLYPROTEIN (MIC LH) [CONTAINS:
DE   MICRONUCLEAR LINKER HISTONE-ALPHA; MICRONUCLEAR LINKER HISTONE-BETA;
DE   MICRONUCLEAR LINKER HISTONE-DELTA; MICRONUCLEAR LINKER HISTONE-GAMMA].
GN   MLH.
OS   Tetrahymena thermophila.

Alignment of: EMBOSS_001 x MLH_TETTH   ..

Alignment segment 1/1:

                     Quality:   90.00                     Escore:       0
             Matching length:     149               Total length:     149
 Matching Percent Similarity:   63.76   Matching Percent Identity:   19.46
    Total Percent Similarity:   63.76      Total Percent Identity:   19.46
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     136 SVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAF 185
         | ||:::::  |:  :|:::|:: ::    |  :   : :  | |::::
     448 SNAAQNSNNTHQTADSSENASSTTQTRTRGRQREQKDMVNEKSNSKSSSK 497
                  .         .         .         .         .
     186 KRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDN 235
          ::  :::| :|||:    :  :| :::  :  | ::: :: :: |::::
     498 GKKNSKSNTRSKSKSKSASKSRKNASKSKKDTTNHGRQTRSKSRSESKSK 547



                  .         .         .         .
     236 ANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESR  284
         ::: :: :|:  | :| : ::|  | :: | || |: |: |:|::::|:
     548 SEAPNKPSNKMEVIEQPKEESSDRKRRESRSQSAKKTSDKKSKNRSDSK  596



Sequence name: swall:O60983
```

```
Sequence documentation:
ID   O60983      PRELIMINARY;      PRT;  1038 AA.
AC   O60983;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REP-LIKE.
GN   REP.
OS   Dictyostelium discoideum (Slime mold).
OG   Plasmid Ddp5.
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

Alignment of: EMBOSS_001 x O60983   ..

Alignment segment 1/1:

                     Quality:   90.00                      Escore:       0
             Matching length:      81                Total length:      81
 Matching Percent Similarity:   77.78   Matching Percent Identity:   22.22
    Total Percent Similarity:   77.78      Total Percent Identity:   22.22
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         ||:: :::::: ||:| ::: ::||::|:|::  :|:| :::  : |: :
     944 EQRERREQQEQQEQQEQQEQQEQQEQQEQQEQQEQQEQQEQQEQREQEEQ 993
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                    296
         ::: :: :| : :|::|:: ::: :::| ||
     994 EEQREQEEQEEQEEQEEQEEEEEEEEQTKKK                    1024



Sequence name: swall:Q9GPD6

Sequence documentation:
ID   Q9GPD6      PRELIMINARY;      PRT;   770 AA.
AC   Q9GPD6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACP36DE (FRAGMENT).
GN   ACP36DE.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9GPD6   ..

Alignment segment 1/1:

                     Quality:   90.00                      Escore:       0
             Matching length:     131                Total length:     133
 Matching Percent Similarity:   67.94   Matching Percent Identity:   22.90
    Total Percent Similarity:   66.92      Total Percent Identity:   22.56
                        Gaps:       1

Alignment:
```

```
                  .         .         .         .         .
     160 RVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYEN 209
         || ::   :::  |  | | ||:|: : :  :::::::|:| :  :  :
      31 RVDKISPELQLQLLRYADSQSQSQSQSASQSESNASSQSQAQVQNN..RL 78
                  .         .         .         .         .
     210 INNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIP 259
         ::|  |:::: | ::::  | ::::::::|:: |:||:  :|:: :::|
      79 LENPTVSESQQNAESQSETQSQSQSESQNQSQSESQNQSQSQSRRQQQIQ 128
                  .         .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                  292
            || || ::: :: :| :: ::  |:| |||
     129 TQLQILRQLQQKSNEQSAAQSASQIQSQRQSDS                  161



Sequence name: swall:Q98QA1

Sequence documentation:
ID   Q98QA1      PRELIMINARY;      PRT;   622 AA.
AC   Q98QA1;
DT   01-SEP-2001 (TrEMBLrel. 18, Created)
DT   01-SEP-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   LIPOPROTEIN.
GN   MYPU_4650.
OS   Mycoplasma pulmonis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.

Alignment of: EMBOSS_001 x Q98QA1   ..

Alignment segment 1/1:

                   Quality:   90.00                     Escore:       0
           Matching length:      98               Total length:      99
 Matching Percent Similarity:   72.45   Matching Percent Identity:   20.41
    Total Percent Similarity:   71.72      Total Percent Identity:   20.20
                      Gaps:       1

Alignment:
                  .         .         .         .         .
     194 TTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRE 243
         ::|::: :  |:  || ||:  :|:| : ::||:: : :::  ::: |:|
     136 SSKEDEKNKNPKDNENSNNNSSDQKNDELQKNNSD.KLNDNVQDEKANKE 184
                  .         .         .         .
     244 NQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDS  292
         |:|   :::: |::  |:|| ::|: :: :  : :::|::::::: :::
     185 NSNSNDSKEKNDENTNKIQQDQNQQKEDQDTKQIEQEKQNNLNSNKNEN  233



Sequence name: swall:Q16322

Sequence documentation:
ID   Q16322      PRELIMINARY;      PRT;   511 AA.
AC   Q16322;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CYCLIC GMP GATED POTASSIUM CHANNEL.
GN   KCN1 OR C2-3A.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x Q16322   ..

Alignment segment 1/1:

                     Quality:   89.00                      Escore:       0
             Matching length:      88                Total length:     108
 Matching Percent Similarity:   70.45   Matching Percent Identity:   26.14
    Total Percent Similarity:   57.41      Total Percent Identity:   21.30
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLV...RTDYFNSFFFWLDIISTVSQILD 123
         | | ::   |:: |::|:::  :|   :||:|::::  :|||| :  :
     269 DPFFMVESTCIVWFTFELVLRFVVCPSKTDFFRNIMNIIDIISIIPYF.. 316
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                        |   ::| | :: |  :: :: ::|::||:|: ::
     317 ...............ATLITELVQETEPSAQQNMSLAILRIIRLVRVFRI 351

     174 YKAASYSQ                                           181
         :| : :|:
     352 FKLSRHSK                                           359



Sequence name: swall:P91642

Sequence documentation:
ID   P91642      PRELIMINARY;      PRT;   454 AA.
AC   P91642;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   CDK8.
GN   CDK8 OR CG10572.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x P91642   ..

Alignment segment 1/1:

                     Quality:   88.00                      Escore:       0
             Matching length:     109                Total length:     113
 Matching Percent Similarity:   69.72   Matching Percent Identity:   24.77
    Total Percent Similarity:   67.26      Total Percent Identity:   23.89
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     153 KTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATI 202
         | :|||::::: : |:  :|   : :  |:| |:::|  || :  :
     303 KPDSKAFHLLQKLLLMDPIKRITSEQAMQDQYFREEPQPTQDVFAGCPIP 352
```

```
                  .         .         .         .         .
     203 YPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQ 252
         ||:  | :::   |::: ||:::: :|:: ::: ::|:::::|:   :||
     353 YPK.REFLTDDDQEDKSDNKRQQQQQQQQQQQQQQQQQQQQQQQ...QQQ 398
                  .
     253 QIDNSIPKVQQIR                                     265
         |  |: |:::::|
     399 QQMNAEPNAKRVR                                     411



Sequence name: swall:O96923

Sequence documentation:
ID   O96923      PRELIMINARY;   PRT;   1087 AA.
AC   O96923;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GELSOLIN-RELATED PROTEIN.
GN   GRP120.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O96923   ..

Alignment segment 1/1:

                     Quality:   88.00                      Escore:       0
             Matching length:     116                Total length:     119
 Matching Percent Similarity:   74.14   Matching Percent Identity:   18.10
    Total Percent Similarity:   72.27      Total Percent Identity:   17.65
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKK.SKATIYPSVYENINNGPVEQENLNKKENNTEQ 229
         :|:  |:: :: : :|: :|     :: |::|:  :| :::|: |:::::
     767 KEEEVKEEEVKEEEVKEVAKEETKEEIKEEVNDEATEVKEVNQVEEEVKE 816
                  .         .         .         .         .
     230 REGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQ..SMKEPSQVKA 277
         :| ::::: : ::|: :  ||:::: ::  | ::::::  ::||  : ::
     817 EEVKEEVKVEVKEEEVKGEAKEEEVKEEEVKEEEVKEEVKEVKEEVKEEV 866
                  .
     278 KEKKESRVSKRLSDSTTKK                               296
         |: ||::|:::::::| ::
     867 KQDKEEEVNEEIKEETKEE                               885



Sequence name: swall:Q9N4C7

Sequence documentation:
ID   Q9N4C7      PRELIMINARY;      PRT;   252 AA.
AC   Q9N4C7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL PROTEIN Y76B12C.4.
GN   Y76B12C.4.
```

```
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x Q9N4C7   ..

Alignment segment 1/1:

                     Quality:   88.00                      Escore:       0
             Matching length:     122                Total length:     132
 Matching Percent Similarity:   64.75   Matching Percent Identity:   27.87
    Total Percent Similarity:   59.85      Total Percent Identity:   25.76
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     175 KAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKE 224
         | |: ::|:  |     |:  :|:|     :: ::::::| |:|:::||:
      94 KEAEKEEEKEEKPPKEVTEEKEKEKEREKTKTKDELEEAPREKETMEKKQ 143
                  .         .         .         .         .
     225 NNTEQREGE.........DNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         :: :: ||:          :| ::|:: :::   | ::||:|    ::
     144 KSKDEDEGDLSKEEILKIKKALKKEKKLRDRERLKTEKIDKSTETKNSTT 193
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKKV                   297
          :: || |::| | |||: |:   | |: |||
     194 PSTEKEKSSMK.KTKKEAAVATGTSGSAKKKV                   224



Sequence name: swall:O35788

Sequence documentation:
ID   O35788      PRELIMINARY;      PRT;  1339 AA.
AC   O35788;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNIT.
GN   CNG4.1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Alignment of: EMBOSS_001 x O35788   ..

Alignment segment 1/1:

                     Quality:   87.00                      Escore:       0
             Matching length:      85                Total length:      85
 Matching Percent Similarity:   75.29   Matching Percent Identity:   18.82
    Total Percent Similarity:   75.29      Total Percent Identity:   18.82
                        Gaps:       0

Alignment:

                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |:|: ::||:: :::| |:|::::|::|:::   :: : :::  : :: :
     403 EEEKKKEKEEEKKEKEEEENGEEEEKEEKEEKEEEEGKEEKEEKEEKEEK 452
```

```
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVIIL                300
         ::: || :: : ||:||:: ::: :::  : :::|
     453 EEEEKEEKEEEEKEEKEEEEEEEEEEEEEEPIVLL                487



Sequence name: swall:O44948

Sequence documentation:
ID   O44948      PRELIMINARY;      PRT;   253 AA.
AC   O44948;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   C17F3.3 PROTEIN.
GN   C17F3.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x O44948   ..

Alignment segment 1/1:

                     Quality:   86.00                      Escore:       0
             Matching length:     107                Total length:     107
 Matching Percent Similarity:   61.68   Matching Percent Identity:   21.50
    Total Percent Similarity:   61.68      Total Percent Identity:   21.50
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     190 IRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQ 239
         ||    | : :|| |: :   : :||  :  :|||:: |:::  |: :::
      37 IRKPPMKPACSTIQPTDFPPASAAPVAPKVEEKKEEKKEEEKKADDEKKK 86
                  .         .         .         .         .
     240 ENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRL 289
          ::::::   | :: |:   |  |  ::: |: :: | :::|::: :::
      87 TEEKDDKKSKKTEEKDKISVKKTQETKSERKDKKDEKKEDEKKEESKEKS 136

     290 SDSTTKK                                            296
         :|:  ||
     137 KDEEKKK                                            143




Sequence name: swall:Q9U0J9

Sequence documentation:
ID   Q9U0J9      PRELIMINARY;      PRT;   842 AA.
AC   Q9U0J9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 94.5 KDA PROTEIN.
```

```
GN   MAL4P2.23.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;

Alignment of: EMBOSS_001 x Q9U0J9   ..

Alignment segment 1/1:

                     Quality:   86.00                     Escore:       0
             Matching length:      83                Total length:      84
 Matching Percent Similarity:   72.29   Matching Percent Identity:   26.51
    Total Percent Similarity:   71.43      Total Percent Identity:   26.19
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         :| |:   : || |::::: ||:::::| |::|::|:|: |  ::| |::
     119 DNNNSDEEQIENSNNNNSDEEQNDSSSNDNNDEENEEQDDVMDNDQNDKK 168
                  .         .         .
     258 IPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSD                291
         |    :: ::| |: :: ::||:|: ::   ::|
     169 IKHSFNLANES.KHTKEERVKEEKKLKIYDFIND                201



Sequence name: swall:Q91781

Sequence documentation:
ID   Q91781      PRELIMINARY;      PRT;   489 AA.
AC   Q91781;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.

Alignment of: EMBOSS_001 x Q91781   ..

Alignment segment 1/1:

                     Quality:   86.00                     Escore:       0
             Matching length:      84                Total length:     100
 Matching Percent Similarity:   71.43   Matching Percent Identity:   30.95
    Total Percent Similarity:   60.00      Total Percent Identity:   26.00
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | |:  :|:| ||:|:::   |   : ::|:::: ::||:: :  ::
     213 DPFFVVETLCIIWFSFELVVRFFACPSKPEFFKNIMNFIDIVAIIPYFI. 261
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                     ::::: |:| :  :  : ||   :||:||||::|| ||
     262 ............TLGTEMAEQEGPQKGEQATSLAILRVIRLVRVFRIFKL 299
```

```
Sequence name: swall:CIK2_RABIT

Sequence documentation:
ID   CIK2_RABIT     STANDARD;      PRT;   273 AA.
AC   Q09081;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.2 (KC22) (FRAGMENT).
GN   KCNA2.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

Alignment of: EMBOSS_001 x CIK2_RABIT   ..

Alignment segment 1/1:

                     Quality:   85.00                     Escore:       0
             Matching length:      88               Total length:     108
 Matching Percent Similarity:   70.45   Matching Percent Identity:   26.14
    Total Percent Similarity:   57.41      Total Percent Identity:   21.30
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | ::  :|:| ||:|:::   |   :: :|::::  :|||| :  :
     135 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIISIIPYL.. 182
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                        |   ::| | : |:: :: :: ::|::||:|: ::
     183 ...............ATLITELVQEDGAQRQQNMSLAILRIIRLVRVFRI 217

     174 YKAASYSQ                                           181
         :| : :|:
     218 FKLSRHSK                                           225



Sequence name: swall:Q9NQS7

Sequence documentation:
ID   Q9NQS7      PRELIMINARY;      PRT;   919 AA.
AC   Q9NQS7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   INNER CENTROMERE PROTEIN INCENP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;

Alignment of: EMBOSS_001 x Q9NQS7   ..

Alignment segment 1/1:

                     Quality:   85.00                     Escore:       0
```

```
             Matching length:      81                Total length:      81
 Matching Percent Similarity:   71.60   Matching Percent Identity:   23.46
    Total Percent Similarity:   71.60      Total Percent Identity:   23.46
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |||: :::::: |:|| | :  :||:|::|: :  ::::::    ::  :
     700 EQERREQERREQERREQERQLAEQERRREQERLQAERELQEREKALRLQK 749
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                    296
         :| ::| :: | ||:::  ::::|:::  ||
     750 EQLQRELEEKKKKEEQQRLAERQLQEEQEKK                    780



Sequence name: swall:O76853

Sequence documentation:
ID   O76853      PRELIMINARY;      PRT;   420 AA.
AC   O76853;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SRF RELATED PROTEIN.
GN   SRFA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O76853   ..

Alignment segment 1/1:

                     Quality:   85.00                      Escore:       0
             Matching length:      99                Total length:      99
 Matching Percent Similarity:   64.65   Matching Percent Identity:   21.21
    Total Percent Similarity:   64.65      Total Percent Identity:   21.21
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     186 KRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDN 235
         | ||: |:   |:      :: :| |:  : ::| |:::|  :|:: ::
     127 KLQPLITRPEGKNLIQSCLNTPDNPNSPSMANQNSNNNSNLLQQQQQQL 176
                  .         .         .         .
     236 ANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESR  284
          ::|: :::|::  :||| :::  : ||  :|:::: :| :| ::: ::
     177 QQQQQLQQQQQQQQQQQQQQQQQQQQAAQQAVQQQQQQQAAQQQAAQ  225



Sequence name: swall:Q9U0K5

Sequence documentation:
ID   Q9U0K5      PRELIMINARY;      PRT;  1111 AA.
AC   Q9U0K5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   HYPOTHETICAL 132.5 KDA PROTEIN.
GN   MAL4P2.17.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;

Alignment of: EMBOSS_001 x Q9U0K5   ..

Alignment segment 1/1:

                     Quality:   85.00                      Escore:       0
             Matching length:      76                Total length:      80
 Matching Percent Similarity:   76.32   Matching Percent Identity:   30.26
    Total Percent Similarity:   72.50      Total Percent Identity:   28.75
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQ....ENRENQNMVAKQQQ 253
         |:::|: :| :: :| ||:|:::: ::| :|:    ||: ::|   |:::
     874 EQMENSEAETSENDKDENKTSENDINENKTSENDINENKTSENNKNKEDS 923
                  .         .         .
     254 IDNSIPKVQQIRRQSMKEPSQVKAKEKKES                     283
         |:|:  | ::|:::: || |  : |:|:|:
     924 IQNDKNKEESIQNDKNKEESIQNDKNKEEA                     953



Sequence name: swall:BAB58163

Sequence documentation:
ID   BAB58163    PRELIMINARY;      PRT;   154 AA.
AC   BAB58163;
DT   14-JUN-2001 (EMBLrel. 63, Created)
DT   14-JUN-2001 (EMBLrel. 63, Last sequence update)
DT   14-JUN-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL 17.8 KDA PROTEIN.
GN   SAV2001.
OS   Staphylococcus aureus subsp. aureus Mu50.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus; Staphylococcus aureus;

Alignment of: EMBOSS_001 x BAB58163   ..

Alignment segment 1/1:

                     Quality:   85.00                      Escore:       0
             Matching length:      62                Total length:      62
 Matching Percent Similarity:   75.81   Matching Percent Identity:   29.03
    Total Percent Similarity:   75.81      Total Percent Identity:   29.03
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     213 GPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQ 262
         |  |::: :||| :::::: | : :::::::::: || ||| ||:  :
      19 GNDENQEESKKEVKSKEKKIEKEKENKSKKDKEKEVATQQQPDNQTVEQP 68
                  .
     263 QIRRQSMKEPSQ                                       274
         | ::||:::|:|
      69 QSQEQSVQQPQQ                                       80
```

```
Sequence name: swall:AAK85503

Sequence documentation:
ID   AAK85503    PRELIMINARY;      PRT;   679 AA.
AC   AAK85503;
DT   05-OCT-2001 (EMBLrel. 63, Created)
DT   05-OCT-2001 (EMBLrel. 63, Last sequence update)
DT   05-OCT-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL PROTEIN Y55B1BR.3.
GN   Y55B1BR.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x AAK85503   ..

Alignment segment 1/1:

                     Quality:   85.00                      Escore:       0
             Matching length:     135                Total length:     149
 Matching Percent Similarity:   70.37   Matching Percent Identity:   24.44
    Total Percent Similarity:   63.76      Total Percent Identity:   22.15
                        Gaps:       3

Alignment:
                  .         .         .         .         .
     143 SQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRT 192
         |:::: :| || |:|: :  :  :  | |:  :: : ::|:  :::
     252 SEMERRRKKSKKSKKSKKFKKSEKRKRAVNDSSSDDEDEEEKPEKR.... 297
                  .         .         .         .         .
     193 QTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENR 242
           :|||| ::  |  |:      |:|: :||::: :::|:::: ::::::
     298 ..SKKSKKAVIDSSSED.....EEEEKSSKKRSKKSKKESDEEQQASDSE 340
                  .         .         .         .
     243 ENQNMVAKQQQIDNSIPK...VQQIRRQSMKEPSQVKAKEKKESRVSKR  288
         |:   | |:::  :: ||   |::  ::|  : ::  :|:||:|:::||
     341 EEVVEVKKNSKSPKKTPKKTAVKEESEESSGDEEEEVVKKKKSSKINKR  389



Sequence name: swall:NPL3_HUMAN

Sequence documentation:
ID   NPL3_HUMAN     STANDARD;      PRT;   506 AA.
AC   Q99457;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   NUCLEOSOME ASSEMBLY PROTEIN 1-LIKE 3.
GN   NAP1L3 OR BNAP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x NPL3_HUMAN   ..

Alignment segment 1/1:

                     Quality:   84.00                      Escore:       0
```

```
             Matching length:      51                Total length:      54
 Matching Percent Similarity:   78.43   Matching Percent Identity:   37.25
    Total Percent Similarity:   74.07      Total Percent Identity:   35.19
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
         |||:: | :|| |   : |::::: ::||:|:: :::  ||  :|||:||
     189 EGEEEENPKENPE...VKAEEKEVPKEIPEVKDEEKEVAKEIPEVKAEEK 235

     281 KESR                                               284
          :|:
     236 ADSK                                               239



Sequence name: swall:P137_HUMAN

Sequence documentation:
ID   P137_HUMAN     STANDARD;      PRT;   649 AA.
AC   Q14444; Q15074;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   GPI-ANCHORED PROTEIN P137 (P137GPI).
GN   M11S1 OR GPIP137.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x P137_HUMAN   ..

Alignment segment 1/1:

                     Quality:   84.00                      Escore:       0
             Matching length:      96                Total length:      96
 Matching Percent Similarity:   66.67   Matching Percent Identity:   19.79
    Total Percent Similarity:   66.67      Total Percent Identity:   19.79
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     219 NLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQS 268
         ||:||::: :: : : | ::: |::: : |:| |:: |::  :::::|:
      14 NLEKKKGKLDDYQERMNKGERLNQDQLDAVSKYQEVTNNLEFAKELQRSF 63
                  .         .         .         .
     269 MKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVILLLLIMPLFSSD     314
         |   ::::   || :| :: : ::: :| :  |: |  ::  :::|
      64 MALSQDIQKTIKKTARREQLMREEAEQKRLKTVLELQYVLDKLGDD     109



Sequence name: swall:Q9BV09

Sequence documentation:
ID   Q9BV09      PRELIMINARY;      PRT;   656 AA.
AC   Q9BV09;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   MEMBRANE COMPONENT, CHROMOSOME 11, SURFACE MARKER 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;

Alignment of: EMBOSS_001 x Q9BV09   ..

Alignment segment 1/1:

                     Quality:   84.00                      Escore:       0
             Matching length:      96                Total length:      96
 Matching Percent Similarity:   66.67   Matching Percent Identity:   19.79
    Total Percent Similarity:   66.67      Total Percent Identity:   19.79
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     219 NLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQS 268
         ||:||::: :: : : | ::: |::: : |:| |:: |::  :::::|:
      14 NLEKKKGKLDDYQERMNKGERLNQDQLDAVSKYQEVTNNLEFAKELQRSF 63
                  .         .         .         .
     269 MKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVILLLLIMPLFSSD     314
         |   ::::   || :| :: : ::: :| :  |: |  ::  :::|
      64 MALSQDIQKTIKKTARREQLMREEAEQKRLKTVLELQYVLDKLGDD     109



Sequence name: swall:O77306

Sequence documentation:
ID   O77306      PRELIMINARY;      PRT;  1338 AA.
AC   O77306; O77319;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   PFC0105W PROTEIN.
GN   PFC0105W.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O77306   ..

Alignment segment 1/1:

                     Quality:   84.00                      Escore:       0
             Matching length:      70                Total length:      73
 Matching Percent Similarity:   80.00   Matching Percent Identity:   28.57
    Total Percent Similarity:   76.71      Total Percent Identity:   27.40
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     209 NINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSI 258
         ||::: ::::::|:| |||::::  :|:::::: |::::| :|| ::::|
     427 NIQTNNINDNTVNEKINNTSKKDMLNNTQNNNDSEKNDVVIEQQLVNEDI 476
                  .         .
     259 PK...VQQIRRQSMKEPSQVKAK                            278
          |    |  ::::::||  || |
     477 LKKKNKQTKKKKNINEPPYVKHK                            499
```

```
Sequence name: swall:Q9PP74

Sequence documentation:
ID   Q9PP74      PRELIMINARY;      PRT;   719 AA.
AC   Q9PP74;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL PROTEIN CJ0849C.
GN   CJ0849C.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.

Alignment of: EMBOSS_001 x Q9PP74   ..

Alignment segment 1/1:

                     Quality:   84.00                      Escore:       0
             Matching length:     117                Total length:     118
 Matching Percent Similarity:   66.67   Matching Percent Identity:   23.93
    Total Percent Similarity:   66.10      Total Percent Identity:   23.73
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     175 KAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKE 224
         | :: :||:   ::  ::: : |::  :| :  || ::  :: :: | :
     393 KNSTPNQEKIKDEKQEKSKENIKENPKFYETKTENKTSINTNTNTSNPNT 442
                  .         .         .         .         .
     225 NNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQ 274
         |||:: :: :| :|::|:  ||:  :|: | ::| |   :: ::: : :|
     443 NNTQNLNNTQNIQSNNNQTMQNIFKNQEFIKQNIVKNLAFNVENL.DLEQ 491
                  .
     275 VKAKEKKESRVSKRLSDS                                 292
         |:  :|: |::|:||::|
     492 VQDLSKNLSNLSRRLNES                                 509



Sequence name: swall:P137_MOUSE

Sequence documentation:
ID   P137_MOUSE     STANDARD;      PRT;   656 AA.
AC   Q60865; Q60758; Q61620;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   GPI-ANCHORED PROTEIN P137 (P137GPI).
GN   M11S1 OR GPIAP OR GPIP137.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x P137_MOUSE   ..

Alignment segment 1/1:

                     Quality:   83.00                      Escore:       0
```

```
             Matching length:      96                Total length:      96
 Matching Percent Similarity:   66.67   Matching Percent Identity:   19.79
    Total Percent Similarity:   66.67      Total Percent Identity:   19.79
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     219 NLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQS 268
         ||:||::: :: : : | ::: |::: : |:| |:: |::  :::::|:
      14 NLEKKKGKLDDYQERMNKGERLNQDQLDAVSKYQEVTNNLGFAKELQRSF 63
                  .         .         .         .
     269 MKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVILLLLIMPLFSSD     314
         |   ::::   || :| :: : ::: :| :  |: |  ::  :::|
      64 MALSQDIQKTIKKTARREQLMREEAEQKRLKTVLELQYVLDKLGDD     109



Sequence name: swall:O60788

Sequence documentation:
ID   O60788      PRELIMINARY;      PRT;   506 AA.
AC   O60788;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DJ32F7.1 (FRAGMENT).
GN   NAP1L3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x O60788   ..

Alignment segment 1/1:

                     Quality:   83.00                         Escore:       0
             Matching length:      51                Total length:      54
 Matching Percent Similarity:   78.43   Matching Percent Identity:   37.25
    Total Percent Similarity:   74.07      Total Percent Identity:   35.19
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
         |||:: | :|| |   : |::::: ::||:|:: :::  ||  :|||:||
     189 EGEEEENPKENPE...VKAEEKEVPKEIPEVKDEEKEVPKEIPEVKAEEK 235

     281 KESR                                               284
          :|:
     236 ADSK                                               239



Sequence name: swall:Q9NW13

Sequence documentation:
ID   Q9NW13      PRELIMINARY;      PRT;   759 AA.
AC   Q9NW13;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   CDNA FLJ10377 FIS, CLONE NT2RM2001989, WEAKLY SIMILAR TO NUCLEOLAR
DE   PROTEIN NOP4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x Q9NW13   ..

Alignment segment 1/1:

                     Quality:   83.00                     Escore:       0
             Matching length:      76                Total length:      81
 Matching Percent Similarity:   77.63   Matching Percent Identity:   26.32
    Total Percent Similarity:   72.84      Total Percent Identity:   24.69
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMV.....AKQQQIDNSIPKV 261
         ||:::|| :: |: | |:| ::::: :::: |      ::::|:::: |
     215 QESVKKKGREEEDMEEEENDDDDDDDDEEDGVFDDEDEEEENIESKVTKP 264
                  .         .         .
     262 QQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                    292
          ||:::::|:|: :|::: :|:  : : |||
     265 VQIQKRAVKRPAPAKSSDHSEEDSDLEESDS                    295



Sequence name: swall:Q26597

Sequence documentation:
ID   Q26597      PRELIMINARY;      PRT;   512 AA.
AC   Q26597;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL PROTEIN.
GN   SKV1.1.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC   Rhabditophora; Eulecithophora; Revertospermata; Mediofusata;

Alignment of: EMBOSS_001 x Q26597   ..

Alignment segment 1/1:

                     Quality:   83.00                     Escore:       0
             Matching length:      86                Total length:      98
 Matching Percent Similarity:   66.28   Matching Percent Identity:   33.72
    Total Percent Similarity:   58.16      Total Percent Identity:   29.59
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      76 DDGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDIT 125
         |: | :|  :|:: || |:::        |: |   :::|::|| |  :
     235 DQPFFIIETFCIVWFSCELLVRFASSPKKFEFFKVLMNVIDVVSIIPYFI 284
                  .         .         .         .
     126 SFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL   173
         ::       |:|   : :|::|:     ||  ::||:||||::|| ||
     285 TL.......GAVIIDDPKQINQT.....TSLAVLRVIRLVRVFRIFKL   320
```

```
Sequence name: swall:Q9T069

Sequence documentation:
ID   Q9T069      PRELIMINARY;      PRT;   532 AA.
AC   Q9T069;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 59.1 KDA PROTEIN.
GN   T28I19.100 OR AT4G37820.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

Alignment of: EMBOSS_001 x Q9T069   ..

Alignment segment 1/1:

                     Quality:   83.00                      Escore:       0
             Matching length:      75                Total length:      81
 Matching Percent Similarity:   78.67   Matching Percent Identity:   30.67
    Total Percent Similarity:   72.84      Total Percent Identity:   28.40
                        Gaps:       2

Alignment:
                 .         .         .         .         .
     223 KENNTEQREGEDNANSQENR....ENQNMVAKQQQIDNSI..PKVQQIRR 266
         ||:: |:|| ||::::|::     ||::  |:::| :|:|   :::: ::
     349 KEEEPEKREKEDSSSQEESKEEEPENKEKEASSSQEENEIKETEIKEKEE 398
                 .         .         .
     267 QSMKEPSQVKAKEKKESRVSKRLSDSTTKKV                    297
         :| :| :: |  |||:|: ::: : :: ||:
     399 SSSQEGNENKETEKKSSESQRKENTNSEKKI                    429



Sequence name: swall:Q54718

Sequence documentation:
ID   Q54718      PRELIMINARY;      PRT;   558 AA.
AC   Q54718;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   M PROTEIN.
GN   EMM19.1.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.

Alignment of: EMBOSS_001 x Q54718   ..

Alignment segment 1/1:

                     Quality:   83.00                      Escore:       0
             Matching length:     124                Total length:     126
 Matching Percent Similarity:   62.90   Matching Percent Identity:   23.39
    Total Percent Similarity:   61.90      Total Percent Identity:   23.02
                        Gaps:       1
```

```
Alignment:
                  .         .         .         .         .
     175 KAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKE 224
         | |  |::||:::|  : :| |::      :: :| |:   : :: :||
     167 KQALESRKQALEEQNKQLSTQKETLERQVQEAQHNNNELTEKLDETQKKL 216
                  .         .         .         .         .
     225 NNTEQREGEDNANSQE..NRENQNMVAKQQQIDNSIPKVQQIRRQSMKEP 272
          | :|:: | : : :|  ::  :: :||:|: :::|  :::|  :::|:
     217 VNKQQESEETKKTLNELLDKTVKDKLAKEQENQETIGTLKKILDETVKDK 266
                  .         .
     273 SQVKAKEKKESRVSKRLSDSTTKKVI                         298
            : |:|::  : |:: |:|:|  |
     267 IAKEQKSKQDIGALKKILDETVKDKI                         292



Sequence name: swall:Q54447

Sequence documentation:
ID   Q54447      PRELIMINARY;      PRT;   563 AA.
AC   Q54447;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   GLUCAN-BINDING PROTEIN PRECURSOR (GBP).
OS   Streptococcus mutans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1309;

Alignment of: EMBOSS_001 x Q54447   ..

Alignment segment 1/1:

                     Quality:   83.00                      Escore:       0
             Matching length:      75                Total length:      82
 Matching Percent Similarity:   65.33   Matching Percent Identity:   28.00
    Total Percent Similarity:   59.76      Total Percent Identity:   25.61
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     315 YYFEPSYSLAYAAEYVRVVAEIPNTKLTEINQTIYFVIDQHKSFDTPVGY 364
         |||: ::|:   :|| :: :::|:  ::||:   ||       |||   :
     447 YYFKDDHSIKAKSEYSQIGGSVPDDGFAEIDGDGYF.......FDTQGQF 489
                  .         .         .
     365 ITNPFTEIENYETPSYQYLRESSKSYYFELVD                   396
         :|| |::  :|::  | |  ::::   :: :|
     490 VTNRFVRKYDYSNIWYYYGSDGKRVSGWQTID                   521


Sequence name: swall:AAH13889

Sequence documentation:
ID   AAH13889    PRELIMINARY;      PRT;   759 AA.
AC   AAH13889;
DT   05-OCT-2001 (EMBLrel. 63, Created)
DT   05-OCT-2001 (EMBLrel. 63, Last sequence update)
DT   05-OCT-2001 (EMBLrel. 63, Last annotation update)
DE   UNKNOWN (PROTEIN FOR MGC:11192).
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;

Alignment of: EMBOSS_001 x AAH13889   ..

Alignment segment 1/1:

                     Quality:   83.00                      Escore:       0
             Matching length:      76                Total length:      81
 Matching Percent Similarity:   77.63   Matching Percent Identity:   26.32
    Total Percent Similarity:   72.84      Total Percent Identity:   24.69
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMV.....AKQQQIDNSIPKV 261
         ||:::|| :: |: | |:| ::::: :::: |      ::::|:::: |
     215 QESVKKKGREEEDMEEEENDDDDDDDDEEDGVFDDEDEEEENIESKVTKP 264
                  .         .         .
     262 QQIRRQSMKEPSQVKAKEKKESRVSKRLSDS                    292
          ||:::::|:|: :|::: :|:  : : |||
     265 VQIQKRAVKRPAPAKSSDHSEEDSDLEESDS                    295



Sequence name: swall:Q9XYL2

Sequence documentation:
ID   Q9XYL2      PRELIMINARY;      PRT;   757 AA.
AC   Q9XYL2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DEVELOPMENT PROTEIN DG1124 (FRAGMENT).
GN   DG1124.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9XYL2   ..

Alignment segment 1/1:

                     Quality:   82.00                      Escore:       0
             Matching length:      86                Total length:      92
 Matching Percent Similarity:   73.26   Matching Percent Identity:   26.74
    Total Percent Similarity:   68.48      Total Percent Identity:   25.00
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     206 VYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQID 255
         : :| ||:  ::|| |:::||:| :  ::| |:::| :|:: ::: :: :
     483 INNNDNNNNNQNENNNQNDNNNEYNIESNNYNNNDNSSNNTNTSNTSSSS 532
                  .         .         .         .
     256 NSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKV         297
          : |::      | |: |: :: |:::| |::: |:||:  :
     533 INSPNL......SNKDDSNRNSAENQRSSVTSNASSSTSTTI         568
```

```
Sequence name: swall:Q9V3R1

Sequence documentation:
ID   Q9V3R1      PRELIMINARY;      PRT;   912 AA.
AC   Q9V3R1; Q9U9Y5; O46198;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ACCESSORY GLAND PROTEIN ACP36DE PRECURSOR.
GN   ACP36DE OR CG7157.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9V3R1   ..

Alignment segment 1/1:

                     Quality:   82.00                      Escore:       0
             Matching length:      96                Total length:      96
 Matching Percent Similarity:   67.71   Matching Percent Identity:   20.83
    Total Percent Similarity:   67.71      Total Percent Identity:   20.83
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     170 IVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQEN 219
         :::| | :: :|:|  ::  :| :: ::| : :  ::  :::|    |:|
     463 LLQLKKLTEVQQKQLAEQPTLRPSSKSQSPGQLEQQILLQLQNLLQFQQN 512
                  .         .         .         .
     220 LNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR     265
           |::::|:::  |:::|| :: ::|::   | | |::: :::||:
     513 QLKSDTQTQSQLQESKSNSLSQSQSQSQEQLQLQRDQNLRQLEQIK     558



Sequence name: swall:Q9TUK7

Sequence documentation:
ID   Q9TUK7      PRELIMINARY;      PRT;   279 AA.
AC   Q9TUK7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL ISOFORM 2 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q9TUK7   ..

Alignment segment 1/1:

                     Quality:   82.00                      Escore:       0
             Matching length:      88                Total length:      97
 Matching Percent Similarity:   67.05   Matching Percent Identity:   34.09
    Total Percent Similarity:   60.82      Total Percent Identity:   30.93
                        Gaps:       2
```

```
Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | ::  :|:| ||:|:::  ::  :  : |   ::||:||:    |
     139 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDTVA....IIP 184
                  .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
         : |::| : :   ::|:| :||      |   :||:||||::|| ||
     185 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    226



Sequence name: swall:Q9L6B1

Sequence documentation:
ID   Q9L6B1      PRELIMINARY;      PRT;  2110 AA.
AC   Q9L6B1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE FILAMENTOUS HEMAGGLUTININ.
GN   PFHAB1.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.

Alignment of: EMBOSS_001 x Q9L6B1   ..

Alignment segment 1/1:

                     Quality:   82.00                      Escore:       0
             Matching length:      59                Total length:      59
 Matching Percent Similarity:   76.27   Matching Percent Identity:   25.42
    Total Percent Similarity:   76.27      Total Percent Identity:   25.42
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     224 ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPS 273
         ::: :|::  :: :::|:|: |: :||| :| ::: :|::||::: :
     538 QKSEQQEKRVEERKQEEKRQAQDKIAKQVEIAKEMQRVEEIRQREKQLAI 587

     274 QVKAKEKKE                                          282
         |:: :|||:
     588 QLQEEEKKQ                                          596



Sequence name: swall:Q9L6B0

Sequence documentation:
ID   Q9L6B0      PRELIMINARY;      PRT;  2095 AA.
AC   Q9L6B0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE FILAMENTOUS HEMAGGLUTININ (FRAGMENT).
GN   PFHAB2.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
```

```
OC   Pasteurella.

Alignment of: EMBOSS_001 x Q9L6B0   ..

Alignment segment 1/1:

                    Quality:   82.00                      Escore:       0
            Matching length:      59                Total length:      59
 Matching Percent Similarity:   76.27   Matching Percent Identity:   25.42
    Total Percent Similarity:   76.27      Total Percent Identity:   25.42
                       Gaps:       0

Alignment:
                  .         .         .         .         .
     224 ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPS 273
         ::: :|::  :: :::|:|: |: :||| :| ::: :|::||::: :
    1191 QKSEQQEKRVEERKQEEKRQAQDKIAKQVEIAKEMQRVEEIRQREKQLAI 1240

     274 QVKAKEKKE                                          282
         |:: :|||:
    1241 QLQEEEKKQ                                          1249



Sequence name: swall:Q9CPH9

Sequence documentation:
ID   Q9CPH9      PRELIMINARY;      PRT;  3919 AA.
AC   Q9CPH9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PFHB2.
GN   PFHB2 OR PM0059.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Pasteurella.

Alignment of: EMBOSS_001 x Q9CPH9   ..

Alignment segment 1/1:

                    Quality:   82.00                      Escore:       0
            Matching length:      59                Total length:      59
 Matching Percent Similarity:   76.27   Matching Percent Identity:   25.42
    Total Percent Similarity:   76.27      Total Percent Identity:   25.42
                       Gaps:       0

Alignment:
                  .         .         .         .         .
     224 ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPS 273
         ::: :|::  :: :::|:|: |: :||| :| ::: :|::||::: :
    1191 QKSEQQEKRVEERKQEEKRQAQDKIAKQVEIAKEMQRVEEIRQREKQLAI 1240

     274 QVKAKEKKE                                          282
         |:: :|||:
    1241 QLQEEEKKQ                                          1249



Sequence name: swall:AAK77711
```

```
Sequence documentation:
ID   AAK77711    PRELIMINARY;      PRT;  1279 AA.
AC   AAK77711;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   ORF42.
OS   White spot syndrome virus (WSSV).
OC   Viruses; Unassigned viruses.
OX   NCBI_TaxID=92652;
RN   [1]

Alignment of: EMBOSS_001 x AAK77711   ..

Alignment segment 1/1:

                     Quality:   82.00                      Escore:       0
             Matching length:     189                Total length:     204
 Matching Percent Similarity:   57.67   Matching Percent Identity:   22.75
    Total Percent Similarity:   53.43      Total Percent Identity:   21.08
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      79 FDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITSFN 128
         ||: ||  ::     | ::  : :   :| :   || :|:::|:|:   :
     491 FDIATIYGQMSDCNAIPLSINIGVPLHKSRMNMQDIEKTIQHIVDMGLAK 540
                  .         .         .         .         .
     129 IAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAAS 178
         | |: :: || :   : :: :||::  : |          |  :|    :
     541 IRVAEENRVARRLLHRRKMQEKAARERAAA..........RQRRLRGEEE 580
                  .         .         .         .         .
     179 YSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTE 228
          ::|:  : : :: :  : : : :  | |:: ::   |:|: :::|:: |
     581 EEEEEEEEGEEMEEEEEEEAGTSGVNGSGYDQEEEDDGEEEEEEEEEEDEE 630
                  .         .         .         .         .
     229 QREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAK 278
         : ||||      | ||::   |:    :|  :  | |::: : | ::|:|
     631 DSEGED.....MNGENNSRKRKNTGNTSSTQQPPQKRQRGKNAPISTKGK 675

     279 EKKE                                               282
         : ||
     676 RLKE                                               679



Sequence name: swall:RBP2_PLAVB

Sequence documentation:
ID   RBP2_PLAVB     STANDARD;      PRT;  1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
```

```
Alignment of: EMBOSS_001 x RBP2_PLAVB   ..

Alignment segment 1/1:

                     Quality:   81.00                      Escore:       0
             Matching length:      77                Total length:      80
 Matching Percent Similarity:   71.43   Matching Percent Identity:   27.27
    Total Percent Similarity:   68.75      Total Percent Identity:   26.25
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     201 TIYPSVYENINNGPVEQENLNKKENNTE...QREGEDNANSQENRENQNM 247
         ||  :: :| :|:   :|| : | |:|:   || :    ::: :::| ::
     185 TIAKEISNNTQNALGFRENAKTKLNKTDELLQRVAAMIEEAKAHKNNIDI 234
                  .         .         .
     248 VAKQQQIDNSIPKVQQIRRQSMKEPSQVKA                     277
         : :: |||::: |::||:|: |:: :::|:
     235 ALEDAQIDTEVSKIEQINREIMNKKDEIKS                     264



Sequence name: swall:Q9W4J3

Sequence documentation:
ID   Q9W4J3      PRELIMINARY;      PRT;   630 AA.
AC   Q9W4J3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG3546 PROTEIN.
GN   CG3546.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9W4J3   ..

Alignment segment 1/1:

                     Quality:   81.00                      Escore:       0
             Matching length:      64                Total length:      64
 Matching Percent Similarity:   75.00   Matching Percent Identity:   21.88
    Total Percent Similarity:   75.00      Total Percent Identity:   21.88
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     221 NKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMK 270
         ||:::| :|:: : | |:|:::::|::  ::|| :::   ::| :: :::
     314 NKNKKNQKQKQKQKNKNQQQKNKRQQQKKRKQQAQDDEQPAKQKKKPQQQ 363
                  .
     271 EPSQVKAKEKKESR                                     284
         : :| | ::| :::
     364 DKDQKKKQDKDQKK                                     377
```

```
Sequence name: swall:Q9N2M3

Sequence documentation:
ID   Q9N2M3      PRELIMINARY;      PRT;  2867 AA.
AC   Q9N2M3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RETICULOCYTE-BINDING PROTEIN 2.
GN   RBP-2.
OS   Plasmodium vivax.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5855;

Alignment of: EMBOSS_001 x Q9N2M3   ..

Alignment segment 1/1:

                    Quality:   81.00                     Escore:       0
            Matching length:      77               Total length:      80
 Matching Percent Similarity:   71.43   Matching Percent Identity:   27.27
    Total Percent Similarity:   68.75      Total Percent Identity:   26.25
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     201 TIYPSVYENINNGPVEQENLNKKENNTE...QREGEDNANSQENRENQNM 247
         ||  :: :| :|:   :|| : | |:|:   || :    ::: :::| ::
    1373 TIAKEISNNTQNALGFRENAKTKLNKTDELLQRVAAMIEEAKAHKNNIDI 1422
                  .         .         .
     248 VAKQQQIDNSIPKVQQIRRQSMKEPSQVKA                     277
         : :: |||::: |::||:|: |:: :::|:
    1423 ALEDAQIDTEVSKIEQINREIMNKKDEIKS                     1452



Sequence name: swall:Q9VT57

Sequence documentation:
ID   Q9VT57      PRELIMINARY;      PRT;   454 AA.
AC   Q9VT57;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   CDK8 PROTEIN.
GN   CDK8 OR CG10572.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9VT57   ..

Alignment segment 1/1:

                    Quality:   81.00                     Escore:       0
            Matching length:     109               Total length:     113
 Matching Percent Similarity:   68.81   Matching Percent Identity:   24.77
    Total Percent Similarity:   66.37      Total Percent Identity:   23.89
                       Gaps:       2

Alignment:
```

```
                  .         .         .         .         .
     153 KTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATI 202
         | :|||::::: : |:   |   : :  |:| |:::|  || :  :
     303 KPDSKAFHLLQKLLLMDPNKRITSEQAMQDQYFQEEPQPTQDVFAGCPIP 352
                  .         .         .         .         .
     203 YPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQ 252
         ||:  | :::   |::: ||:::: :|:: ::: ::|:::::|:   :||
     353 YPK.REFLTDDDQEDKSDNKRQQQQQQQQQQQQQQQQQQQQQQQ...QQQ 398
                  .
     253 QIDNSIPKVQQIR                                     265
         |  |: |:::::|
     399 QQMNAEPNAKRVR                                     411



Sequence name: swall:Q9LH98

Sequence documentation:
ID   Q9LH98      PRELIMINARY;      PRT;  2081 AA.
AC   Q9LH98;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   GENOMIC DNA, CHROMOSOME 3, BAC CLONE: T19N8.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Alignment of: EMBOSS_001 x Q9LH98   ..

Alignment segment 1/1:

                     Quality:   81.00                        Escore:       0
             Matching length:      88                  Total length:      97
 Matching Percent Similarity:   78.41   Matching Percent Identity:   23.86
    Total Percent Similarity:   71.13      Total Percent Identity:   21.65
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKE.....NNTEQREGEDNANSQENRENQNMVAKQQ 252
         |:::    |::: ||:|     |::::::|:|: ::::: :|:||  |::
     912 ESVKYKKDEKKEGNKEENKDTINTSSKQKGKDKKKKKKESKNSNMKKKEE 961
                  .         .         .         .
     253 Q....IDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTK    295
         :    ::|:: | :: :::: |: ::   :|:|::: :|: :||::|
     962 DKKEYVNNELKKQEDNKKETTKSENSKLKEENKDNKEKKESEDSASK    1008



Sequence name: swall:Q99QZ4

Sequence documentation:
ID   Q99QZ4      PRELIMINARY;      PRT;   619 AA.
AC   Q99QZ4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   SA2437 PROTEIN.
GN   SA2437.
```

```
OS   Staphylococcus aureus subsp. aureus N315.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.

Alignment of: EMBOSS_001 x Q99QZ4   ..

Alignment segment 1/1:

                   Quality:   81.00                    Escore:       0
           Matching length:      94              Total length:     100
 Matching Percent Similarity:   70.21   Matching Percent Identity:   24.47
    Total Percent Similarity:   66.00      Total Percent Identity:   23.00
                      Gaps:       1

Alignment:
                  .         .         .         .         .
     205 SVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQ......NMV 248
         | ||: :||  : :: ||:::|: ::: : ::::|:: :||      |
     134 SDYEQPRNGEKSTNDSNKNSDNSIKNDTDTQSSKQDKADNQKAPKSNNTK 183
                  .         .         .         .         .
     249 AKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVI 298
          : :: : : ||  | :::: :  |: ||::|::|: :: :|||:  :::
     184 PSTSNKQPNSPKPTQPNQSNSQPASDDKANQKSSSKDNQSMSDSALDSIL 233



Sequence name: swall:BAB58806

Sequence documentation:
ID   BAB58806    PRELIMINARY;      PRT;   619 AA.
AC   BAB58806;
DT   14-JUN-2001 (EMBLrel. 63, Created)
DT   14-JUN-2001 (EMBLrel. 63, Last sequence update)
DT   14-JUN-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL 69.2 KDA PROTEIN.
GN   SAV2644.
OS   Staphylococcus aureus subsp. aureus Mu50.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus; Staphylococcus aureus;

Alignment of: EMBOSS_001 x BAB58806   ..

Alignment segment 1/1:

                   Quality:   81.00                    Escore:       0
           Matching length:      94              Total length:     100
 Matching Percent Similarity:   70.21   Matching Percent Identity:   24.47
    Total Percent Similarity:   66.00      Total Percent Identity:   23.00
                      Gaps:       1

Alignment:
                  .         .         .         .         .
     205 SVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQ......NMV 248
         | ||: :||  : :: ||:::|: ::: : ::::|:: :||      |
     134 SDYEQPRNGEKSTNDSNKNSDNSIKNDTDTQSSKQDKADNQKAPKSNNTK 183
                  .         .         .         .         .
     249 AKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVI 298
          : :: : : ||  | :::: :  |: ||::|::|: :: :|||:  :::
     184 PSTSNKQPNSPKPTQPNQSNSQPASDDKANQKSSSKDNQSMSDSALDSIL 233
```

```
Sequence name: swall:AAK77753

Sequence documentation:
ID   AAK77753    PRELIMINARY;      PRT;  2314 AA.
AC   AAK77753;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   ORF84.
OS   White spot syndrome virus (WSSV).
OC   Viruses; Unassigned viruses.
OX   NCBI_TaxID=92652;
RN   [1]

Alignment of: EMBOSS_001 x AAK77753   ..

Alignment segment 1/1:

                     Quality:   81.00                      Escore:       0
             Matching length:     115                Total length:     118
 Matching Percent Similarity:   71.30   Matching Percent Identity:   16.52
    Total Percent Similarity:   69.49      Total Percent Identity:   16.10
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     167 LIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVE 216
         | : :::::  : :|:| |::| :: |  ::::  :   : ::|::
     983 LSEALQIFQQQQQQQQQQFQQQLLQQQQDQQNQQQL...LQQQIEEQQRV 1029
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRR 266
         ||: :::::: :|:| ::: ::|::::::::  ::|| :::  : :|:|:
    1030 QEQQQQQRDQQQQQEQQQREQQQQQQQREQQQQREQQQQQQQQQSDQFRQ 1079
                  .
     267 QSMKEPSQVKAKEKKESR                                 284
         | ::: :| :   ::::|
    1080 QLLQQQQQFQQLLQQQGR                                 1097



Sequence name: swall:AAL06398

Sequence documentation:
ID   AAL06398    PRELIMINARY;      PRT;   231 AA.
AC   AAL06398;
DT   05-OCT-2001 (EMBLrel. 63, Created)
DT   05-OCT-2001 (EMBLrel. 63, Last sequence update)
DT   05-OCT-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL 27.1 KDA PROTEIN.
OS   Francisella tularensis subsp. tularensis.
OC   Bacteria; Proteobacteria; gamma subdivision; Francisella group;
OC   Francisella; Francisella tularensis.
OX   NCBI_TaxID=119856;

Alignment of: EMBOSS_001 x AAL06398   ..

Alignment segment 1/1:

                     Quality:   81.00                      Escore:       0
             Matching length:      81                Total length:      81
```

```
 Matching Percent Similarity:   72.84   Matching Percent Identity:   18.52
    Total Percent Similarity:   72.84      Total Percent Identity:   18.52
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         :::: |:::|| ::::::|| ::::|::|::   :::: ||:  | :: :
      88 QKQQQNSQNNNKDNKDNKDNKDNKDNKDNKDNKDNKDNKDNKDNKDNKDN 137
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                    296
         :::  : ::   |::|::: :|  :|:  :|
     138 KDNKDNKDNKDNKDNKDNKDNKDNKDNKDNK                    168



Sequence name: swall:INVO_CANFA

Sequence documentation:
ID   INVO_CANFA     STANDARD;      PRT;   285 AA.
AC   P18174;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   INVOLUCRIN.
GN   IVL.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

Alignment of: EMBOSS_001 x INVO_CANFA   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:      66                Total length:      66
 Matching Percent Similarity:   69.70   Matching Percent Identity:   25.76
    Total Percent Similarity:   69.70      Total Percent Identity:   25.76
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         ||:: ::|| : ||:: ::: ::||:::::::  |:|:  |   :::||
     145 EQQESQEKELHLEQEQQKEELQQQEQQQGKEQCEKHQEAKNLEQQLEQIG 194
                  .
     266 RQSMKEPSQVKAKEKK                                   281
          |  :: ::   :|||
     195 AQRKQQQKEQLEQEKK                                   210



Sequence name: swall:PHPA_PLACH

Sequence documentation:
ID   PHPA_PLACH     STANDARD;      PRT;   441 AA.
AC   Q02752;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   ACIDIC PHOSPHOPROTEIN PRECURSOR (50 KDA ANTIGEN).
```

```
GN   PCEMA1.
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;

Alignment of: EMBOSS_001 x PHPA_PLACH   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:      81                Total length:      81
 Matching Percent Similarity:   71.60   Matching Percent Identity:   19.75
    Total Percent Similarity:   71.60      Total Percent Identity:   19.75
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     205 SVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQI 254
         |: :  :|:  :|||::| ::|  :::|::| ::|::::::    :::::
     336 SLLDYDENSNENQENVKKGNENEGEQKGNENEGEQKGKKKKAKEKSKKKV 385
                  .         .         .
     255 DNSIPKVQQIRRQSMKEPSQVKAKEKKESRV                    285
          |:   : : :::: |: :: : ::|||::|
     386 KNKPTMTTKKKKKKEKKKKKKEKEKKKEKKV                    416



Sequence name: swall:Q23847

Sequence documentation:
ID   Q23847      PRELIMINARY;      PRT;   720 AA.
AC   Q23847;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   GLUTAMINE-ASPARAGINE RICH PROTEIN (FRAGMENT).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]

Alignment of: EMBOSS_001 x Q23847   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:     108                Total length:     108
 Matching Percent Similarity:   68.52   Matching Percent Identity:   16.67
    Total Percent Similarity:   68.52      Total Percent Identity:   16.67
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQR 230
         ::: :|:: :: :  |:::  :     |:::   ::||:|:::: : ||
     460 KQEQLKKEQLKQEQIKQEQLKLEQIKQEQLKLEQLKQEELKQEQLKQEQL 509
                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
         : ::  ::||:  :|::: :|| ::::  : || ::|:: | :| : :::
     510 KQQQIKQQQEKSIQQQQLLEQQLLEQQQHQQQQQHQQLLEQQQQQHQQQ 559
```

```
     281 KESRVSKR                                           288
         :::: :::
     560 QHQQYQQQ                                           567



Sequence name: swall:Q26223

Sequence documentation:
ID   Q26223      PRELIMINARY;      PRT;  2269 AA.
AC   Q26223;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   RHOPTRY PROTEIN.
OS   Plasmodium berghei yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5862;
RN   [1]

Alignment of: EMBOSS_001 x Q26223   ..

Alignment segment 1/1:

                     Quality:   80.00                   Escore:       0
             Matching length:      85             Total length:      85
 Matching Percent Similarity:   69.41   Matching Percent Identity:   25.88
    Total Percent Similarity:   69.41      Total Percent Identity:   25.88
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     202 IYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQ 251
         :: :| ||::|    :|:|:  : :   :||: : ::: :: |::::| :
     584 VFEKVKENVTNLNKIKEKLKHYDFSDFGKEGNIKYTDKIKKINDDIMAVS 633
                  .         .         .
     252 QQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVS                286
         ||||: |  :::|:::| :  |::| : :| ::||
     634 QQIDQHINGLDDIQKKSESYVSEMKEQINKLEKVS                668



Sequence name: swall:Q9Y0C9

Sequence documentation:
ID   Q9Y0C9      PRELIMINARY;      PRT;   838 AA.
AC   Q9Y0C9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RAS INTERACTING PROTEIN RIPA.
GN   RIPA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9Y0C9   ..

Alignment segment 1/1:

                     Quality:   80.00                   Escore:       0
```

```
              Matching length:     102              Total length:     104
 Matching Percent Similarity:    71.57   Matching Percent Identity:    18.63
    Total Percent Similarity:    70.19      Total Percent Identity:    18.27
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQR 230
         |:|  ::|   : : || | : : ::  ::|:  ::::: :::::| :|:
     275 QQQQQQQQQQTSTSPKKCKIVTFTGTSVKLNSKTLQKSDKTSEKENKQQQ 324
                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
            |::::|:::: |:: ::|||  ::  :  || ::|:::: :| : :::
     325 ..PDSSKTQQQQQAQQQQSQQQQQAQQQQQQQQQQQQQQQQQQQQQQQQQ 372

     281 KESR                                               284
         ::::
     373 QQQQ                                               376



Sequence name: swall:Q9W3U2

Sequence documentation:
ID   Q9W3U2      PRELIMINARY;      PRT;  1086 AA.
AC   Q9W3U2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG14431 PROTEIN.
GN   CG14431.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9W3U2   ..

Alignment segment 1/1:

                      Quality:   80.00                    Escore:       0
              Matching length:      57              Total length:      57
 Matching Percent Similarity:    71.93   Matching Percent Identity:    31.58
    Total Percent Similarity:    71.93      Total Percent Identity:    31.58
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     192 TQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQEN 241
         || :: | |  | :    :||   ::|: :::::::|::: :|: :||||
     348 TQEVETSTAYSYSTSTTVVNNVEESREEERENSEESEREDEQDQEQSQEN 397

     242 RENQNMV                                            248
         :|||::|
     398 QENQEVV                                            404
```

```
Sequence name: swall:Q9U9S7

Sequence documentation:
ID   Q9U9S7      PRELIMINARY;      PRT;  2123 AA.
AC   Q9U9S7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   ADENYLYL CYCLASE.
GN   ACRA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9U9S7   ..

Alignment segment 1/1:

                     Quality:   80.00                        Escore:       0
             Matching length:      82                  Total length:      82
 Matching Percent Similarity:   78.05   Matching Percent Identity:   17.07
    Total Percent Similarity:   78.05      Total Percent Identity:   17.07
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         :|:: :::::: :|:: ::: ::|:::::|::  :|||  :   : ||::
    2008 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQHHHHHHQQQQFQ 2057
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKKV                   297
         :||::: :| : ::::::: |:: |:: ::::
    2058 QQSQQSQQQSQQQQQQQQQQSQQQSQQQSQQI                   2089



Sequence name: swall:O97271

Sequence documentation:
ID   O97271      PRELIMINARY;      PRT;   466 AA.
AC   O97271;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 54.8 KDA PROTEIN.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]

Alignment of: EMBOSS_001 x O97271   ..

Alignment segment 1/1:

                     Quality:   80.00                        Escore:       0
             Matching length:     100                  Total length:     101
 Matching Percent Similarity:   67.00   Matching Percent Identity:   22.00
    Total Percent Similarity:   66.34      Total Percent Identity:   21.78
                        Gaps:       1

Alignment:
```

```
                  .         .         .         .         .
     194 TTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRE 243
         | |::  : : :   || ||  ::::  ||: | :||:| :|  :::::|
     196 TLKRNPDAPFVTKPGNILNGK.KKNRRYKKNANKRQRKGGENKMKSQKEE 244
                  .         .         .         .         .
     244 NQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDST 293
         :::| :|:::  :   | :: :::: |: ::  :::||| ::::: ::
     245 EKKMDSKKEEEKKMDSKKKEEKKMDSKKKEDKMGNQKKECKINNQKKERK 294

     294 T                                                  294
         |
     295 T                                                  295



Sequence name: swall:Q9NB35

Sequence documentation:
ID   Q9NB35      PRELIMINARY;      PRT;  1323 AA.
AC   Q9NB35;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MATURE PARASITE-INFECTED ERYTHROCYTE SURFACE ANTIGEN.
GN   MESA.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x Q9NB35   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:     118                Total length:     129
 Matching Percent Similarity:   70.34   Matching Percent Identity:   26.27
    Total Percent Similarity:   64.34      Total Percent Identity:   24.03
                        Gaps:       4

Alignment:
                  .         .         .         .         .
     169 RIVKLYKAASYSQEQAFKRQPIRTQTTKKSK...ATIYPSVYENINNGPV 215
         :::::|: ::|::  :  |:  ::: |::||   ::   : ||| :|: |
     329 KVTRIYEETKYTKITSEFRETENVKITEESKDREGNKVSGPYENSENSNV 378
                  .         .         .         .         .
     216 .....EQENLNKKENNTEQREGED.NANSQENRENQNMVAKQQQIDNSIP 259
              | ::| :||:|  :: ||: | :::|| |: : : :|:  :|:
     379 TSESEETKKLAEKEENEGEKLGENVNDGASENSEDPKKLPEQE..ENGTK 426
                  .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKR                      288
         : ::  :::  | :: || :||:|: :|:
     427 ESSEETKDDKPEENEKKADNKKKSKKKKK                      455



Sequence name: swall:Q9GSG6

Sequence documentation:
ID   Q9GSG6      PRELIMINARY;      PRT;   757 AA.
AC   Q9GSG6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
```

```
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   AARDVARK.
GN   AAR.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9GSG6   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:     133                Total length:     141
 Matching Percent Similarity:   64.66   Matching Percent Identity:   23.31
    Total Percent Similarity:   60.99      Total Percent Identity:   21.99
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQ. 245
         || :|    : || ||:  :::| |:::||::::::::| |:::| :::
     113 KSTTTSNYILEENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSSSSS 162
                  .         .         .         .         .
     246 .......NMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKR 288
                | : :::::: :  : ||:::|: :  :| :  |  |:: :
     163 SSILSKFNKLEEDNELELQKKQKQQLEQQEEELFNQFNFLEGIEDQNDFL 212
                  .         .         .         .
     289 LSDSTTKKVIILVILLLLIMPLFSSDYYFEPSYSLAYAAEY          329
          :::| :|: :|: :    |  :||   : || | :| : :
     213 SEQETIQKIKFLIKMTAKSMSNYSSPNTLIPSVSKTYISPF          253



Sequence name: swall:Q28649

Sequence documentation:
ID   Q28649      PRELIMINARY;      PRT;   725 AA.
AC   Q28649;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CGMP-GATED POTASSIUM CHANNEL.
GN   KCN1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

Alignment of: EMBOSS_001 x Q28649   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:      88                Total length:     108
 Matching Percent Similarity:   69.32   Matching Percent Identity:   23.86
    Total Percent Similarity:   56.48      Total Percent Identity:   19.44
                        Gaps:       2

Alignment:
```

```
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLV...RTDYFNSFFFWLDIISTVSQILD 123
         | | ::   |:: |::|:::  :|   :| ::::::  :|||| :  :
     483 DPFFMVESTCIVWFTLELVLRFVVCPSKTAFYRNIMNIIDIISIIPYF.. 530
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
                        |   ::| | :: |  :: :: ::|::||:|: ::
     531 ...............ATLITELVQETEPSAQQNMSLAILRIIRLVRVFRI 565



     174 YKAASYSQ                                           181
         :| : :|:
     566 FKLSRHSK                                           573



Sequence name: swall:Q44782

Sequence documentation:
ID   Q44782      PRELIMINARY;      PRT;   218 AA.
AC   Q44782;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPB.
GN   ERPB.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-1.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x Q44782   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:        0
             Matching length:      75                Total length:       75
 Matching Percent Similarity:   76.00   Matching Percent Identity:    21.33
    Total Percent Similarity:   76.00      Total Percent Identity:    21.33
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         :::::: :: |: ::|::|||:|: :: |:::|::::::   ::|: :::
     123 KKVEESEAKVEGKEEKQENTEERNKQELAKQEEEQQKRKAEQEKQKREEE 172
                  .         .
     258 IPKVQQIRRQSMKEPSQVKAKEKKE                          282
           : :: ::|: |  :: :|||| |
     173 QERQKREEEQERKAKAEKEAKEKAE                          197






Sequence name: swall:Q44791

Sequence documentation:
ID   Q44791      PRELIMINARY;      PRT;   338 AA.
AC   Q44791;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPD.
GN   ERPD.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-2.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x Q44791   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:      93                Total length:      95
 Matching Percent Similarity:   74.19   Matching Percent Identity:   22.58
    Total Percent Similarity:   72.63      Total Percent Identity:   22.11
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     201 TIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAK 250
         |: | : :  ::|  ::|: :| |:: |::|::::  :::|:|:::   |
      92 TLVPIAKSAEQSGDQKEEKSGKVEEKKEKQESKEEKVEEQNQEKKQ..KK 139
                  .         .         .         .
     251 QQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTK      295
         :::  ::  : |: :||:::| :::||:::|::| :|: :::  |
     140 EERNVKEEKQKQEEERQKQQEEARAKAEKEKREREEKQKQEEEKK      184



Sequence name: swall:O31328

Sequence documentation:
ID   O31328      PRELIMINARY;      PRT;   378 AA.
AC   O31328;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPJ.
GN   ERPJ.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-5.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x O31328   ..

Alignment segment 1/1:

                     Quality:   80.00                      Escore:       0
             Matching length:      75                Total length:      75
 Matching Percent Similarity:   76.00   Matching Percent Identity:   21.33
    Total Percent Similarity:   76.00      Total Percent Identity:   21.33
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         :::::: :: |: ::|::|||:|: :: |:::|::::::   ::|: :::
     123 KKVEESEAKVEGKEEKQENTEERNKQELAKQEEEQQKRKAEQEKQKREEE 172
```

```
                  .         .
     258 IPKVQQIRRQSMKEPSQVKAKEKKE                          282
           : :: ::|: |  :: :|||| |
     173 QERQKREEEQERKAKAEKEAKEKAE                          197



Sequence name: swall:O08264

Sequence documentation:
ID   O08264      PRELIMINARY;      PRT;   378 AA.
AC   O08264;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPB2 PROTEIN.
GN   ERPB2 OR BBL40 OR BBP39.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-1, and Plasmid cp32-8.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x O08264   ..

Alignment segment 1/1:

                     Quality:   80.00                    Escore:       0
              Matching length:      75               Total length:      75
 Matching Percent Similarity:   76.00   Matching Percent Identity:   21.33
    Total Percent Similarity:   76.00      Total Percent Identity:   21.33
                         Gaps:       0

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         :::::: :: |: ::|::|||:|: :: |:::|::::::   ::|: :::
     123 KKVEESEAKVEGKEEKQENTEERNKQELAKQEEEQQKRKAEQEKQKREEE 172
                  .         .
     258 IPKVQQIRRQSMKEPSQVKAKEKKE                          282
           : :: ::|: |  :: :|||| |
     173 QERQKREEEQERKAKAEKEAKEKAE                          197



Sequence name: swall:Q9QR71

Sequence documentation:
ID   Q9QR71      PRELIMINARY;      PRT;  1129 AA.
AC   Q9QR71;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LATENT NUCLEAR ANTIGEN.
GN   ORF 73.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.

Alignment of: EMBOSS_001 x Q9QR71   ..

Alignment segment 1/1:
```

```
                    Quality:   80.00                  Escore:       0
            Matching length:      69            Total length:      69
 Matching Percent Similarity:   76.81   Matching Percent Identity:   26.09
    Total Percent Similarity:   76.81      Total Percent Identity:   26.09
                       Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |||: :::|:: ||:: |:: ::|:: |:|:   :||: :::  : || :
     648 EQEQQEEQEQQEEQQQDEQQQDEQQQDEQQQDEQEQQDEQQQDEQQQQDE 697
                  .
     266 RQSMKEPSQVKAKEKKESR                                284
         :::::| :| : :|::|::
     698 QEQQEEQEQQEEQEQQEEQ                                716


Sequence name: swall:PTP3_DICDI

Sequence documentation:
ID   PTP3_DICDI     STANDARD;      PRT;   989 AA.
AC   P54637;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   PROTEIN-TYROSINE PHOSPHATASE 3 (EC 3.1.3.48) (PROTEIN-TYROSINE-
DE   PHOSPHATE PHOSPHOHYDROLASE 3).
GN   (PTPC1 OR PTP3) AND (PTPC2 OR PTP3).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

Alignment of: EMBOSS_001 x PTP3_DICDI   ..

Alignment segment 1/1:

                    Quality:   79.00                  Escore:       0
            Matching length:     100            Total length:     101
 Matching Percent Similarity:   71.00   Matching Percent Identity:   26.00
    Total Percent Similarity:   70.30      Total Percent Identity:   25.74
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     209 NINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSI 258
         | ||:  :::| |:::||::::::::| |:::|::|:|   ::::|: ::
     146 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSNSNIEINV 195
                  .         .         .         .         .
     259 PKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVILLLLIM 308
         |::|  :: :| | ::|   :  ::::   |:  :||::    | || |:
     196 PSIQFDNEPAM.EVDSVAPLNVPSNHTRTTLAMHNTKSLSTSNIGLLNIL 244

     309 P                                                  309
         |
     245 P                                                  245



Sequence name: swall:Q9HI57

Sequence documentation:
ID   Q9HI57      PRELIMINARY;      PRT;   295 AA.
```

```
AC   Q9HI57;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL PROTEIN TA1488.
GN   TA1488.
OS   Thermoplasma acidophilum.
OC   Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC   Thermoplasma.

Alignment of: EMBOSS_001 x Q9HI57   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:      95                Total length:     101
 Matching Percent Similarity:   73.68   Matching Percent Identity:   23.16
    Total Percent Similarity:   69.31      Total Percent Identity:   21.78
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     203 YPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQ 252
         | :: |: | : |:|:::: ||:: ::   :::: :  : |::::| : :
      89 YKKITEEANYASVNQRDIKAKERELQKLIKKQETMQLPKAEEDKVVQEIK 138
                  .         .         .         .         .
     253 QIDNSIPKVQQIRRQSMKEPSQVKA......KEKKESRVSKRLSDSTTKK 296
         :::|:| |::: |:::::| ::|||      |||  :|  |: ::: :::
     139 RLSNEIKKMKEDREKQLSENEKVKALLAEIDKEKTLARDLKKRAEEISNQ 188

     297 V                                                  297
         :
     189 I                                                  189



Sequence name: swall:P91570

Sequence documentation:
ID   P91570      PRELIMINARY;      PRT;   312 AA.
AC   P91570;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   COSMID ZK354.
GN   ZK354.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x P91570   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:     106                Total length:     113
 Matching Percent Similarity:   64.15   Matching Percent Identity:   19.81
    Total Percent Similarity:   60.18      Total Percent Identity:   18.58
                        Gaps:       1

Alignment:
```

```
                  .         .         .         .         .
     190 IRTQTTKKSKATIYPSVYENINNGPV.......EQENLNKKENNTEQREG 232
         ||    | : :|| |: :   : :|:       ::|: :::|:: |::::
      37 IRKPPMKPACSTIKPTDFPPASAAPIAPKGEEKKEEEKKEEEKKEEEKKA 86
                  .         .         .         .         .
     233 EDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKE 282
         :|: :: |::::::    :::   |: |:|: : :   : :: |  :|||
      87 DDEKKKTEEKDDKKSKKTEEKDKLSVKKTQETKSERKDKKDERKEDDKKE 136
                  .
     283 SRVSKRLSDSTTK                                      295
         :: :|  :::  |
     137 ENKEKSKDEEKKK                                      149



Sequence name: swall:O76719

Sequence documentation:
ID   O76719      PRELIMINARY;      PRT;   335 AA.
AC   O76719;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   F36H12.3 PROTEIN.
GN   F36H12.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x O76719   ..

Alignment segment 1/1:

                     Quality:   79.00                       Escore:       0
             Matching length:      67                 Total length:      67
 Matching Percent Similarity:   74.63    Matching Percent Identity:   20.90
    Total Percent Similarity:   74.63       Total Percent Identity:   20.90
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |::: ::||:: :::| : : ::::|:::::   |::: |::  | :: :
     132 EDKKEDEKEKDDDKKEDKKDDKKSDNKDDKDDEKKEEKKDDKEEKKEEKK 181
                  .
     266 RQSMKEPSQVKAKEKKE                                  282
         ::: :| :: | :||||
     182 EEKKEEKKEEKKEEKKE                                  198



Sequence name: swall:Q17909

Sequence documentation:
ID   Q17909      PRELIMINARY;   PRT;   385 AA.
AC   Q17909;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C11G6.3 PROTEIN.
GN   C11G6.3.
```

```
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x Q17909   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:      82                Total length:      82
 Matching Percent Similarity:   71.95   Matching Percent Identity:   20.73
    Total Percent Similarity:   71.95      Total Percent Identity:   20.73
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     222 KKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKE 271
         |:::: :::| | : :::::||:::   :::: :::  | :: :::  ||
     210 KERSERKEKERELEREKEKSREKEREKEREKEREKEREKEREKQKEREKE 259
                  .         .         .
     272 PSQVKAKEKKESRVSKRLSDSTTKKVIILVIL                   303
          :: : ||||::: ::| :::::  |||  :|
     260 REKEREKEKKREEEARRKKEEASTPVIIRPLL                   291



Sequence name: swall:Q9NG85

Sequence documentation:
ID   Q9NG85      PRELIMINARY;      PRT;   362 AA.
AC   Q9NG85;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CUT PROTEIN (FRAGMENT).
GN   CT.
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9NG85   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:      75                Total length:      79
 Matching Percent Similarity:   65.33   Matching Percent Identity:   30.67
    Total Percent Similarity:   62.03      Total Percent Identity:   29.11
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGED....NANSQENR 242
         || :|     : | ||:  ::|:|:::|:: |: | ||    ||: |:|
     176 KSGSTTSNVNHTNSNNSHQDEEELDEEEEDEEEDEDEDDEEENASMQSNA 225
                  .         .
     243 ENQNMVAKQQQIDNSIPKVQQIRRQSMKE                      271
         ::::: |:|:   :    :|| ::|:  |
     226 DDMELDAQQETRTEPSATTQQXQQQQDAE                      254
```

```
Sequence name: swall:Q9GTW3

Sequence documentation:
ID   Q9GTW3      PRELIMINARY;      PRT;   682 AA.
AC   Q9GTW3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   GLUTAMIC ACID-RICH PROTEIN.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x Q9GTW3   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:      77                Total length:      81
 Matching Percent Similarity:   77.92   Matching Percent Identity:   23.38
    Total Percent Similarity:   74.07      Total Percent Identity:   22.22
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |||:::||::: |::| : : ::::::|::    :::| :::: | ::|:
     280 EQEKIEKKKKKQEEKEKKKQEKERKKQEKK....ERKQKEKEMKKQKKIE 325
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                   296
         ::  |: :: | |: ||:: : :  |:|:::
     326 KERKKKEEKEKKKHDKENEETMQQPDQTSEE                   356



Sequence name: swall:Q97FK1

Sequence documentation:
ID   Q97FK1      PRELIMINARY;      PRT;  1163 AA.
AC   Q97FK1;
DT   01-SEP-2001 (TrEMBLrel. 18, Created)
DT   01-SEP-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   ATPASE INVOLVED IN DNA REPAIR.
GN   CAC2736.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.

Alignment of: EMBOSS_001 x Q97FK1   ..

Alignment segment 1/1:

                     Quality:   79.00                      Escore:       0
             Matching length:     117                Total length:     125
 Matching Percent Similarity:   66.67   Matching Percent Identity:   27.35
    Total Percent Similarity:   62.40      Total Percent Identity:   25.60
                        Gaps:       3
```

```
Alignment:
                  .         .         .         .         .
     172 KLYKAASYSQEQAFK.RQPIRTQTTKKSKATIYPSVYENINNGPVEQENL 220
         :|::  :|::| :||  : || :  |::  :   : |||||:  :::::
     170 RLFNLQEYGDELSFKLARKIRKEREKENVLVGELKGYENINEDVLKERRE 219
                  .         .         .         .         .
     221 NKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMK 270
           ||||   :|:: :  : |:: |::  :   ||:  | : :::|:: |:
     220 LLKENNDFFNEASKEYLKAEEEYNEGKEVWGLQIE..IEEKNRVRKDLME 267
                  .         .
     271 EPSQVKAKEKKESRVSKRLSDSTTK                          295
         : :::  |||:      ||::|::|
     268 KKDEIDLKEKR.....ARLGESSSK                          287



Sequence name: swall:YHV4_YEAST

Sequence documentation:
ID   YHV4_YEAST     STANDARD;      PRT;  1070 AA.
AC   P38850;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   HYPOTHETICAL 123.0 KDA PROTEIN IN SPO16-REC104 INTERGENIC REGION.
GN   YHR154W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.

Alignment of: EMBOSS_001 x YHV4_YEAST   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      80                Total length:      83
 Matching Percent Similarity:   75.00   Matching Percent Identity:   30.00
    Total Percent Similarity:   72.29      Total Percent Identity:   28.92
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTE.QREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQI 264
         :::||:||::|:: : | |  ::::|:||:::::    : |   || |:|
     582 KKSNLQKKDSNSDISMETEVFCEGHEKREEKEFTKPITEYD..APKKQEI 629
                  .         .         .
     265 RRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKV                  297
         |:|| |: :    ||::|:::: :|:: | :::
     630 REQSRKKNDIDYKKEEEETELQVQLGQRTKREI                  662



Sequence name: swall:MNN4_YEAST

Sequence documentation:
ID   MNN4_YEAST     STANDARD;      PRT;  1178 AA.
AC   P36044; P36043; P89095;
DT   01-JUN-1994 (Rel. 29, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MNN4 PROTEIN.
```

```
GN   MNN4 OR YKL200C/YKL201C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.

Alignment of: EMBOSS_001 x MNN4_YEAST   ..

Alignment segment 1/1:

                    Quality:   78.00                        Escore:       0
            Matching length:     123                  Total length:     128
 Matching Percent Similarity:   65.85   Matching Percent Identity:   21.14
    Total Percent Similarity:   63.28      Total Percent Identity:   20.31
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     174 YKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKK 223
         |  |:  :|:  :::  : :  ||:|     :  |: ::   |:|: :||
    1028 YAYAKLLEERKRREKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKK 1077
                  .         .         .         .         .
     224 ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPS 273
         |:: :::: |:: ::||::|:::   :::: :::  |:::  ::: |: :
    1078 EEEEKKKKEEEEKKKQEEEEKKKKEEEEKKKQEEGEKMKNEDEENKKNED 1127
                  .         .
     274 QVKAK.....EKKESRVSKRLSDSTTKK                       296
         : | |     :||::: :|: :|:  ||
    1128 EEKKKNEEEEKKKQEEKNKKNEDEEKKK                       1155



Sequence name: swall:KAPC_DICDI

Sequence documentation:
ID   KAPC_DICDI     STANDARD;      PRT;   648 AA.
AC   P34099;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT (EC 2.7.1.37).
GN   PKAC OR PK2 OR PK3.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x KAPC_DICDI   ..

Alignment segment 1/1:

                    Quality:   78.00                        Escore:       0
            Matching length:      69                  Total length:      69
 Matching Percent Similarity:   78.26   Matching Percent Identity:   18.84
    Total Percent Similarity:   78.26      Total Percent Identity:   18.84
                       Gaps:       0

Alignment:


                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         :|:: ::::::  |:: : : : |:|:::|::  :|||::::: : || :
     154 QQQQPQQQQQQQPQQQQQPQQQLQQNNQQQQQQLQQQQLQQQLQQQQQQQ 203
```

```
                  .
     266 RQSMKEPSQVKAKEKKESR                                284
         :|:::: :| : |::::::
     204 QQQQQQQQKQQKQQQQQQ                                 222



Sequence name: swall:O43988

Sequence documentation:
ID   O43988      PRELIMINARY;      PRT;   800 AA.
AC   O43988;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   HOMEOBOX-CONTAINING PROTEIN WARIAI (FRAGMENT).
GN   WARIAI.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O43988   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      82                Total length:      88
 Matching Percent Similarity:   76.83   Matching Percent Identity:   26.83
    Total Percent Similarity:   71.59      Total Percent Identity:   25.00
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     206 VYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQID 255
         :|:| ||      | |:::||::::::::|:|:::|::|:| : :::: :
      70 MYNNNNN......NNNNNNNNNNNNNNNNNSNNNNNNNNNNNINNNNNNN 113
                  .         .         .
     256 NSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDST             293
         |:  : :: ::| ::: :|::    :::  || || ||
     114 NNNNNNNNNNNQHLSQSQQLSPTPYSSNSFSKLLSRST             151



Sequence name: swall:O00910

Sequence documentation:
ID   O00910      PRELIMINARY;      PRT;   707 AA.
AC   O00910;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   STAT PROTEIN.
GN   DSTA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O00910   ..

Alignment segment 1/1:
```

```
                     Quality:   78.00                      Escore:       0
             Matching length:      95                Total length:     102
 Matching Percent Similarity:   75.79   Matching Percent Identity:   23.16
    Total Percent Similarity:   70.59      Total Percent Identity:   21.57
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     179 YSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTE 228
         :|::|:::::  :: :::::  :   :: :| ||: ::::| |:: ||::
     139 FSRQQSYNNNNSNNTSSSQNYNNNNININNNNNNNNTNNNNNNNNGNNSN 188
                  .         .         .         .         .
     229 QREGEDNANSQ.......ENRENQNMVAKQQQIDNSIPKVQQIRRQSMKE 271
          ::|::| |::       :|::|:|:  :||| :::  : || ::|:: :
     189 GNNGNNNNNNNNNNNNNTNNNNNNNQQQQQQQQQQQQQQQQQQQQQQQGN 238

     272 PS                                                 273
         |:
     239 PN                                                 240



Sequence name: swall:O96133

Sequence documentation:
ID   O96133      PRELIMINARY;   PRT;   1979 AA.
AC   O96133;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   HYPOTHETICAL 237.7 KDA PROTEIN.
GN   PFB0145C.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O96133   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      66                Total length:      66
 Matching Percent Similarity:   77.27   Matching Percent Identity:   24.24
    Total Percent Similarity:   77.27      Total Percent Identity:   24.24
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQR 230
         ::::: :: |:|:  :|:|     |: :|||:  ::::| |:::||::::
      37 KSDSWYKKIIETKGKSKTKYKNDNSLDDNINEDIINNNNNNNNDNNNDNN 86
                  .
     231 EGEDNANSQENRENQN                                   246
         : ::| |:::|::::|
      87 NDNNNDNNNDNNNDNN                                   102
```

```
Sequence name: swall:O96229

Sequence documentation:
ID   O96229      PRELIMINARY;   PRT;   665 AA.
AC   O96229;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 78.6 KDA PROTEIN.
GN   PFB0680W.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O96229   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      61                Total length:      61
 Matching Percent Similarity:   77.05   Matching Percent Identity:   26.23
    Total Percent Similarity:   77.05      Total Percent Identity:   26.23
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         || |::  ::|: |::||::|:::|::| |::||::::|:   ::  : :
     493 ENGNENENKNESENENENENENENGNENENEKENEKDKNIKEIENVTNAN 542
                  .
     258 IPKVQQIRRQS                                        268
           : ::|:::|
     543 KENYEKINKNS                                        553



Sequence name: swall:Q9VA29

Sequence documentation:
ID   Q9VA29      PRELIMINARY;      PRT;   578 AA.
AC   Q9VA29;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG12071 PROTEIN.
GN   CG12071.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9VA29   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      78                Total length:      78
 Matching Percent Similarity:   71.79   Matching Percent Identity:   23.08
    Total Percent Similarity:   71.79      Total Percent Identity:   23.08
                        Gaps:       0

Alignment:
```

```
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRR 266
         |  :: |:::|:|:::: :||:::| :|:: : :||| :::  : || ::
      25 QAAIQFKQQQTSQQQNSPTANNNNNSQNNGNMQQQQQQQQQQQQQQQQQQ 74
                  .         .
     267 QSMKEPSQVKAKEKKESRVSKRLSDSTT                       294
         |::   : :|: : ::|:|:   : ::|
      75 QQQHYAATIKVPQISSSNVAAAAAGEST                       102



Sequence name: swall:Q9TXB8

Sequence documentation:
ID   Q9TXB8      PRELIMINARY;      PRT;   648 AA.
AC   Q9TXB8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   SERINE/THREONINE PROTEIN KINASE.
OS   Dictyostelium.
OC   Eukaryota; Mycetozoa; Dictyosteliida.
OX   NCBI_TaxID=5782;
RN   [1]

Alignment of: EMBOSS_001 x Q9TXB8   ..

Alignment segment 1/1:

                     Quality:   78.00                     Escore:       0
             Matching length:      69               Total length:      69
 Matching Percent Similarity:   78.26   Matching Percent Identity:   18.84
    Total Percent Similarity:   78.26      Total Percent Identity:   18.84
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         :|:: ::::::  |:: : : : |:|:::|::  :|||::::: : || :
     154 QQQQPQQQQQQQPQQQQQPQQQLQQNNQQQQQQLQQQQLQQQLQQQQQQQ 203
                  .
     266 RQSMKEPSQVKAKEKKESR                                284
         :|:::: :| : |::::::
     204 QQQQQQQQKQQKQQQQQQ                                 222



Sequence name: swall:O97242

Sequence documentation:
ID   O97242      PRELIMINARY;      PRT;  1114 AA.
AC   O97242;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 132.2 KDA PROTEIN.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
```

```
Alignment of: EMBOSS_001 x O97242   ..

Alignment segment 1/1:

                     Quality:   78.00                       Escore:       0
             Matching length:      87                 Total length:      99
 Matching Percent Similarity:   70.11   Matching Percent Identity:   24.14
    Total Percent Similarity:   61.62      Total Percent Identity:   21.21
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKV.... 261
         ::|: | |::: :::| :|| |::: :||:|   :: : ||:  |:
     779 DDEDDNDKDEDDNDKEEDDNDNDNDIEENDNTEDSDYREDNDDEKISGNS 828
                  .         .         .         .
     262 ........QQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVI  302
                 :: : ||:|: |: | |:||:::: ::::       :::::
     829 ENTNMMTHSDEESQSVKQMSNKKKKKKKKKKILENINKYIIMPEFVIIL  877


Sequence name: swall:Q9GZ76

Sequence documentation:
ID   Q9GZ76      PRELIMINARY;      PRT;  1387 AA.
AC   Q9GZ76;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RETICULOCYTE-BINDING PROTEIN 2 HOMOLOG A (FRAGMENT).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x Q9GZ76   ..

Alignment segment 1/1:

                     Quality:   78.00                       Escore:       0
             Matching length:      81                 Total length:      83
 Matching Percent Similarity:   76.54   Matching Percent Identity:   19.75
    Total Percent Similarity:   74.70      Total Percent Identity:   19.28
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSI..PKVQQ 263
         ::: : |:|:: ::||  ::|:::|:|:::: : |::|  : :   | :|
    1030 DEKLILKEEEERKERERLEKAKQEEERKERERIEKEKQEKERLEREKQEQ 1079
                  .         .         .
     264 IRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                  296
         ::::::|: :| : :::::::: || :::  :|
    1080 LKKEALKKQEQERQEQQQKEEALKRQEQERLQK                  1112
```

```
Sequence name: swall:Q06166

Sequence documentation:
ID   Q06166  PRELIMINARY;      PRT;  1661 AA.
AC   Q06166;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MATURE PARASITE-INFECTED ERYTHROCYTE SURFACE ANTIGEN (ANTIGENIC
DE   PROTEIN PFEMP2).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x Q06166   ..

Alignment segment 1/1:

                   Quality:   78.00                      Escore:       0
           Matching length:     118                Total length:     129
 Matching Percent Similarity:   71.19   Matching Percent Identity:   26.27
    Total Percent Similarity:   65.12      Total Percent Identity:   24.03
                      Gaps:       4

Alignment:
                  .         .         .         .         .
     169 RIVKLYKAASYSQEQAFKRQPIRTQTTKKSK...ATIYPSVYENINNGPV 215
         :::::|: ::|::  :  |:  ::: |::||   ::   : ||| :|: |
     323 KVTRIYEETKYTKITSEFRETENVKITEESKDREGNKVSGPYENSENSNV 372
                  .         .         .         .         .
     216 .....EQENLNKKENNTEQREGED.NANSQENRENQNMVAKQQQIDNSIP 259
              | ::| :||:|  :: ||: | :::|| |: : :::|::  |:
     373 TSESEETKKLAEKEENEGEKLGENVNDGASENSEDPKKLTEQEK..NGTK 420
                  .         .
     260 KVQQIRRQSMKEPSQVKAKEKKESRVSKR                     288
         : ::  :::  | :: || :||:|: :|:
     421 ESSEETEDDKPEENEKKADNKKKSKKKKK                     449



Sequence name: swall:Q9BK45

Sequence documentation:
ID   Q9BK45      PRELIMINARY;      PRT;  3254 AA.
AC   Q9BK45;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 HOMOLOG B.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x Q9BK45   ..

Alignment segment 1/1:

                   Quality:   78.00                      Escore:       0
           Matching length:      81                Total length:      83
 Matching Percent Similarity:   76.54   Matching Percent Identity:   19.75
```

```
    Total Percent Similarity:   74.70      Total Percent Identity:   19.28
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSI..PKVQQ 263
         ::: : |:|:: ::||  ::|:::|:|:::: : |::|  : :   | :|
    2666 DEKLILKEEEERKERERLEKAKQEEERKERERIEKEKQEKERLEREKQEQ 2715
                  .         .         .
     264 IRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                  296
         ::::::|: :| : :::::::: || :::  :|
    2716 LKKEALKKQEQERQEQQQKEEALKRQEQERLQK                  2748



Sequence name: swall:Q04111

Sequence documentation:
ID   Q04111      PRELIMINARY;      PRT;   891 AA.
AC   Q04111;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SURFACE EXCLUSION PROTEIN.
GN   PRGA.
OS   Enterococcus faecalis (Streptococcus faecalis).
OG   Plasmid pCF10.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;

Alignment of: EMBOSS_001 x Q04111   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:     172                Total length:     185
 Matching Percent Similarity:   67.44   Matching Percent Identity:   22.67
    Total Percent Similarity:   62.70      Total Percent Identity:   21.08
                        Gaps:       5

Alignment:
                  .         .         .         .         .
     115 ISTVSQILDITSFNIAVGLQGSVAAKSASQ.LSQANKASKTSSKAIRVVR 163
         :::::::||::: |:    | :|| |:::: ||: ::|: | :||:: ::
     671 LNSLKEVLDLAKENLN...QKQVALKTSTRSLSRLENAQPTYEKALNELN 717
                  .         .         .         .         .
     164 LVR..LIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYE... 208
          ::  :::  : |: :  | |:  ::| : | :  :::  :  :  |
     718 KAEAAVVQAQEAYENSMKSLEELKEQQAVATLAYAQAQEDLSNAKLELQQ 767
                  .         .         .         .         .
     209 ...NINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQID 255
             :::  ::| : :::|   ||   |::  ::| ::||::||:  : |
     768 YQGVLRDLEAQQAEQRRQEALQEQVAKEQQRLEREAKQNQTLVASATSAD 817
                  .         .         .
     256 NSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLS                290
         :: | :||:  :::||  :::|   :| | :|:::
     818 KT.PGLQQLSFSKQKEQPKAQALTHSEPRKTKQVA                851
```

```
Sequence name: swall:Q9S035

Sequence documentation:
ID   Q9S035      PRELIMINARY;      PRT;   343 AA.
AC   Q9S035;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPB2 PROTEIN.
GN   BBN39.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-9.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x Q9S035   ..

Alignment segment 1/1:

                     Quality:   78.00                        Escore:       0
             Matching length:      85                  Total length:      93
 Matching Percent Similarity:   74.12   Matching Percent Identity:   28.24
    Total Percent Similarity:   67.74      Total Percent Identity:   25.81
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     204 PSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQ 253
         | : |||:|: :  : ::::::: |::  | :|: :||:|      ||::
      87 PVLPENIHNNALVLKAIEQSDGQQEKKVEEAEAKVEENKE......KQEN 130
                  .         .         .         .
     254 IDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK        296
          :::| : : | :|: :|   :||||::::: :|| :::  :|
     131 TEENIKEKEIIDEQNKQE..LAKAKEEEQQKEQKRHQEEQQRK        171



Sequence name: swall:AAK54091

Sequence documentation:
ID   AAK54091    PRELIMINARY;      PRT;  2062 AA.
AC   AAK54091;
DT   14-JUN-2001 (EMBLrel. 63, Created)
DT   14-JUN-2001 (EMBLrel. 63, Last sequence update)
DT   14-JUN-2001 (EMBLrel. 63, Last annotation update)
DE   HISTIDINE KINASE DHKJ.
GN   DHKJ.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x AAK54091   ..

Alignment segment 1/1:

                     Quality:   78.00                        Escore:       0
             Matching length:     105                  Total length:     111
 Matching Percent Similarity:   64.76   Matching Percent Identity:   26.67
    Total Percent Similarity:   61.26      Total Percent Identity:   25.23
                        Gaps:       2

Alignment:
```

```
                  .         .         .         .         .
     175 KAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKE 224
         |:|:| :: ::  |   | ||  : :|  ||  : ::    :::|||:::
    1867 KSANYINSTSYYGQLTPTTTTTTTTTALPSPQKILS..IEDDKNLNSND 1914
                  .         .         .         .         .
     225 NNTEQREGEDNANSQ....ENRENQNMVAKQQQIDNSIPKVQQIRRQSMK 270
         || :::::::| |:|    :| :|||   : : |:|:  :    :::: :
    1915 NNEKDNNNQNNNNNQNDNNKNDNNQNDFDQIKTISNNKESGVGENKKAPR 1964
                  .
     271 EPSQVKAKEKK                                       281
         :||: :   ||
    1965 NPSDNERIPKK                                       1975



Sequence name: swall:AAK54092

Sequence documentation:
ID   AAK54092    PRELIMINARY;      PRT;  1686 AA.
AC   AAK54092;
DT   14-JUN-2001 (EMBLrel. 63, Created)
DT   14-JUN-2001 (EMBLrel. 63, Last sequence update)
DT   14-JUN-2001 (EMBLrel. 63, Last annotation update)
DE   HISTIDINE KINASE DHKL (FRAGMENT).
GN   DHKL.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x AAK54092   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:     112                Total length:     114
 Matching Percent Similarity:   69.64   Matching Percent Identity:   23.21
    Total Percent Similarity:   68.42      Total Percent Identity:   22.81
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     174 YKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKK 223
         |  :|: : :: : :   | ||: :: :   ::|:: :|  |:::| | :
     178 YLKSSFFSGNGSSSSSSSTTTTNINSNNSNTNTYSTSTN..VKKNNSNTN 225
                  .         .         .         .         .
     224 ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPS 273
         :|:::: :|: |:||::|::| |  :::: | |:   ::: ||: :|: |
     226 SNTNNNTNGNTNTNSNSNKKNINSNSSDNIIINDTADIMNRRRDRFKSFS 275
                  .
     274 QVKAKEKKESRVSK                                    287
          : ::::::|: |:
     276 WTIGESNSNSKFSE                                    289



Sequence name: swall:AAK50002

Sequence documentation:
ID   AAK50002    PRELIMINARY;      PRT;  1003 AA.
AC   AAK50002;
DT   06-SEP-2001 (EMBLrel. 63, Created)
```

```
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   ORF73.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=37296;

Alignment of: EMBOSS_001 x AAK50002   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:     111                Total length:     111
 Matching Percent Similarity:   67.57   Matching Percent Identity:   16.22
    Total Percent Similarity:   67.57      Total Percent Identity:   16.22
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQR 230
         |:|   ::| : :  :::     |:  |  :: | :||  :: |:: :::
     450 QQQEPLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQDEQQQDEQ 499


                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
         : ::: ::::::::|::  :||| :::  : ||  :|:::| :: : :::
     500 QQDEQQQDEQQQDEQEQQDEQQQDEQQQDEQQQQDEQEQQEEQEQQEEQE 549
                  .
     281 KESRVSKRLSD                                        291
         :::: :::|::
     550 QQEEQEQELEE                                        560



Sequence name: swall:AAG02259

Sequence documentation:
ID   AAG02259    PRELIMINARY;      PRT;  3254 AA.
AC   AAG02259;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN ANALOG.
GN   RBP.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x AAG02259   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      81                Total length:      83
 Matching Percent Similarity:   76.54   Matching Percent Identity:   19.75
    Total Percent Similarity:   74.70      Total Percent Identity:   19.28
                        Gaps:       1

Alignment:
```

```
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSI..PKVQQ 263
         ::: : |:|:: ::||  ::|:::|:|:::: : |::|  : :   | :|
    2666 DEKLILKEEEERKERERLEKAKQEEERKERERIEKEKQEKERLEREKQEQ 2715
                  .         .         .
     264 IRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                  296
         ::::::|: :| : :::::::: || :::  :|
    2716 LKKEALKKQEQERQEQQQKEEALKRQEQERLQK                  2748



Sequence name: swall:AAK84527

Sequence documentation:
ID   AAK84527    PRELIMINARY;      PRT;   447 AA.
AC   AAK84527;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   HYPOTHETICAL PROTEIN F40F4.8.
GN   F40F4.8.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x AAK84527   ..

Alignment segment 1/1:

                     Quality:   78.00                      Escore:       0
             Matching length:      71                Total length:      71
 Matching Percent Similarity:   74.65   Matching Percent Identity:   25.35
    Total Percent Similarity:   74.65      Total Percent Identity:   25.35
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     214 PVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQ 263
         |: |:| |:::||::::::::|:|| :: :|||:  ::||  ::  : :|
      86 PIPQNNRNNNNNNNNNNNNNQNGNSPNRPRNQNQNNRNQQNMQGQGNFNQ 135
                  .         .
     264 IRRQSMKEPSQVKAKEKKESR                              284
          ::| ::: :| : ::  ::|
     136 NQNQLQNQRTQFRNNQAMRNR                              156



Sequence name: swall:CIKS_DROME

Sequence documentation:
ID   CIKS_DROME     STANDARD;      PRT;   656 AA.
AC   P08510; P08511; Q24521; P08512; P08513; Q24277; Q9VWZ9;
DT   01-AUG-1988 (Rel. 08, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN SHAKER.
GN   SH OR CG12348.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
```

```
Alignment of: EMBOSS_001 x CIKS_DROME   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:     104                Total length:     105
 Matching Percent Similarity:   63.46   Matching Percent Identity:   20.19
    Total Percent Similarity:   62.86      Total Percent Identity:   20.00
                        Gaps:       1

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | :|  :|:| |::|: :  |:  : :|     :::|: :: |  : :
     277 DPFFLIETLCIIWFTFELTVRFLACPNKLNFCRDVMNVIDIIAIIPYFIT 326
                  .         .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKA 176
         :  :|: : ::     : :|  :|:| ::: :: ::|::||:|: :::|
     327 LATVVAEEEDTLNLPKAPVSPQDKSS.NQAMSLAILRVIRLVRVFRIFKL 375

     177 ASYSQ                                              181
         : :|:
     376 SRHSK                                              380



Sequence name: swall:HMCU_DROME

Sequence documentation:
ID   HMCU_DROME     STANDARD;      PRT;  2175 AA.
AC   P10180; Q9W3Q6;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HOMEOBOX PROTEIN CUT.
GN   CT OR CG11387.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x HMCU_DROME   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:      75                Total length:      79
 Matching Percent Similarity:   65.33   Matching Percent Identity:   30.67
    Total Percent Similarity:   62.03      Total Percent Identity:   29.11
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGED....NANSQENR 242
         || :|   : : | ||:  ::|:|: :|:: |: | ||    ||: |:|
     252 KSGSTTSNANHTNSNNSHQDEEELDDEEEDEEEDEDEDDEEENASMQSNA 301
                  .         .
     243 ENQNMVAKQQQIDNSIPKVQQIRRQSMKE                      271
         ::::: |:|:   :    :|| ::|:  |
     302 DDMELDAQQETRTEPSATTQQQHQQQDTE                      330
```

```
Sequence name: swall:CENC_HUMAN

Sequence documentation:
ID   CENC_HUMAN     STANDARD;      PRT;   943 AA.
AC   Q03188; Q9P0M5;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CENTROMERE PROTEIN C (CENP-C) (CENTROMERE AUTOANTIGEN C).
GN   CENPC1 OR CENPC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Alignment of: EMBOSS_001 x CENC_HUMAN   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:     146                Total length:     149
 Matching Percent Similarity:   66.44   Matching Percent Identity:   18.49
    Total Percent Similarity:   65.10      Total Percent Identity:   18.12
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     150 KASKTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAFKRQPIRTQTTKKSK 199
         ::|: |:|::  :  : : : |: |::::|:: :   ::| |:: | |:|
     371 ETSQPSDKTVLDTSYALIDETVNNYRSTKYEMYSKNAEKPSRSKRTIKQK 420

                  .         .         .         .         .
     200 AT...IYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQN 246
              :   : |::: |  ::||:: :: : :: : ::: | :|::|  |
     421 QRRKFMAKPAEEQLDVGQSKDENIHTSHITQDEFQRNSDRNMEEHEEMGN 470

                  .         .         .         .
     247 MVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTK  295
           ::::|:     | :: |::: :: ::  ::|:|:: | :::::::||
     471 DCVSKKQMPPVGSKKSSTRKDKEESKKKRFSSESKNKLVPEEVTSTVTK  519



Sequence name: swall:Q25709

Sequence documentation:
ID   Q25709      PRELIMINARY;      PRT;   374 AA.
AC   Q25709;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   10B ANTIGEN (FRAGMENT).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x Q25709   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:     100                Total length:     111
```

```
 Matching Percent Similarity:   70.00   Matching Percent Identity:   27.00
    Total Percent Similarity:   63.06      Total Percent Identity:   24.32
                        Gaps:       2

Alignment:
                  .         .         .         .         .
     175 KAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKE 224
         |: : :||  :|: | : :: |::|:       :| ||   :::| |:|:
     185 KSKNVQQETPIKHLPNEEENMKQGKT.......KNSNNKNNNNKNNNNKN 227
                  .         .         .         .         .
     225 NNTEQREGEDNANSQE....NRENQNMVAKQQQIDNSIPKVQQIRRQSMK 270
         ||::::::::| |: |    |::|:: :: : : :|   : |: : ::||
     228 NNNKNNNNNNNNNNNLESITKNKNNSQYISTNVENENLNKNEQKGKMNKMK 277
                  .
     271 EPSQVKAKEKK                                       281
           :|   ||::
     278 TKTQRFNKEEQ                                       288



Sequence name: swall:O61335

Sequence documentation:
ID   O61335      PRELIMINARY;      PRT;   522 AA.
AC   O61335;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL.
GN   SHAKER.
OS   Panulirus interruptus (California spiny lobster).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;

Alignment of: EMBOSS_001 x O61335   ..

Alignment segment 1/1:

                     Quality:   77.00                     Escore:       0
             Matching length:     102               Total length:     108
 Matching Percent Similarity:   67.65   Matching Percent Identity:   19.61
    Total Percent Similarity:   63.89      Total Percent Identity:   18.52
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIII...ASLVRTDYFNSFFFWLDIISTVSQILD 123
         | | :|  ||:: || |:::   || ::  :::: :  :||:: :  ::
     234 DPFFLIETICIVWFSTELLVRFFASPLKGAFLKDVMNIIDIVAIIPYFIT 283
                  .         .         .         .         .
     124 ITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL 173
         :::   :|: : |      ::  : :: ::::: :: ::|::||:|: ::
     284 LAT...VVAEEESEELVIPAETLSPHEKGSNQAMSLAILRVIRLVRVFRI 330

     174 YKAASYSQ                                          181
         :| : :|:
     331 FKLSRHSK                                          338
```

```
Sequence name: swall:Q22665

Sequence documentation:
ID   Q22665      PRELIMINARY;      PRT;   167 AA.
AC   Q22665;
DT   01-JAN-1999 (TrEMBLrel. 09, Created)
DT   01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   HYPOTHETICAL 19.6 KDA PROTEIN T22C1.5 IN CHROMOSOME I.
GN   T22C1.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x Q22665   ..

Alignment segment 1/1:

                     Quality:   77.00                        Escore:       0
             Matching length:      71                  Total length:      73
 Matching Percent Similarity:   71.83   Matching Percent Identity:   25.35
    Total Percent Similarity:   69.86      Total Percent Identity:   24.66
                        Gaps:       1

Alignment:

                  .         .         .         .         .
     218 ENLNKKENNTEQREGEDNANSQENR..ENQNMVAKQQQIDNSIPKVQQIR 265
         |:|:: :::| ::::||: :::||:  |: | : :: ::|:  ||: |::
      87 EKLRRMDEETAEQDAEDEKENDENKEPESPNRTLRKTKVDTPRPKTAQLK 136
                  .         .
     266 RQSMKEPSQVKAKEKKESRVSKR                            288
         |: :   |:  :::: : | | |
     137 RKRIIMSSDDDSEDEFSMRSSPR                            159



Sequence name: swall:O96201

Sequence documentation:
ID   O96201      PRELIMINARY;   PRT;   1844 AA.
AC   O96201;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   HYPOTHETICAL 214.7 KDA PROTEIN.
GN   PFB0540W.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O96201   ..

Alignment segment 1/1:

                     Quality:   77.00                        Escore:       0
             Matching length:      68                  Total length:      75
 Matching Percent Similarity:   75.00   Matching Percent Identity:   27.94
    Total Percent Similarity:   68.00      Total Percent Identity:   25.33
                        Gaps:       1
```

```
Alignment:
                  .         .         .         .         .
     197 KSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQN 246
         |::: |   :|:| ||:       |:|::|::| : ::: ::|::|::||
    1273 KKEGQIKNVIYNNNNNN.......NEKDQNADQYQDQNKNHKQDQRQDQN 1315
                  .         .
     247 MVAKQQQIDNSIPKVQQIRRQSMKE                          271
            ||:|: ::  :  |:: |::::
    1316 KSHKQDQLHDQNQNQGQLHDQNQNQ                          1340



Sequence name: swall:Q9NDS4

Sequence documentation:
ID   Q9NDS4      PRELIMINARY;      PRT;  2678 AA.
AC   Q9NDS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   AMIB.
GN   AMIB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9NDS4   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:      76                Total length:      76
 Matching Percent Similarity:   78.95   Matching Percent Identity:   21.05
    Total Percent Similarity:   78.95      Total Percent Identity:   21.05
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRR 266
         ::|:|:::||  ::::::| |:: |::|:::  :||| :::  : || ::
     489 KNNINNNSNNIINNNNNSNINNNINNNNNQQQQQQQQQQQQQQQQQQQQQ 538
                  .         .
     267 QSMKEPSQVKAKEKKESRVSKRLSDS                         292
         |:::: :| : :::::::: :: |::
     539 QQQQQQQQQQQQQQQQQQTINNNSNN                         564



Sequence name: swall:Q9ZU69

Sequence documentation:
ID   Q9ZU69      PRELIMINARY;      PRT;   699 AA.
AC   Q9ZU69;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN-LIKE).
GN   F24H14.11.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

```
Alignment of: EMBOSS_001 x Q9ZU69   ..

Alignment segment 1/1:

                    Quality:   77.00                    Escore:       0
            Matching length:     123              Total length:     125
 Matching Percent Similarity:   65.04   Matching Percent Identity:   18.70
    Total Percent Similarity:   64.00      Total Percent Identity:   18.40
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     172 KLYKAASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLN 221
         |  :|   ::|:| :|:  :|:  |:::        |: ::   |: : :
     447 KREEAKRREEEEAKRREEEETERKKREEEEARKR..EEERKREEEEAKRR 494
                  .         .         .         .         .
     222 KKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKE 271
         ::|:: :::|:|:  : :|:||::: :||::: :::  : ::::|:  :|
     495 EEERKKREEEAEQARKREEEREKEEEMAKKREEERQRKEREEVERKRREE 544
                  .         .
     272 PSQVKAKEKKESRVSKRLSDSTTKK                         296
          :: : :|: ::| ::|  ::   |
     545 QERKRREEEARKREEERKREEEMAK                         569



Sequence name: swall:Q9FPT6

Sequence documentation:
ID   Q9FPT6      PRELIMINARY;      PRT;   533 AA.
AC   Q9FPT6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   NOP58-LIKE PROTEIN F108.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Alignment of: EMBOSS_001 x Q9FPT6   ..

Alignment segment 1/1:

                    Quality:   77.00                    Escore:       0
            Matching length:      98              Total length:     100
 Matching Percent Similarity:   66.33   Matching Percent Identity:   21.43
    Total Percent Similarity:   65.00      Total Percent Identity:   21.00
                       Gaps:       1

Alignment:
                  .         .         .         .         .
     189 PIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANS 238
         | :| :|  ::    |:|  : :||  |::: :||::  :::|:: :  |
     431 PAKTYNTAADSLLQTPTV..DSENGVKEKKDKKKKKKADDEEEAKTEEPS 478
                  .         .         .         .         .
     239 QENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKR 288
         ::: ::::  |: :  ::   | :: :|:  :| ::: ||:|::|: :|:
     479 KKKSNKKKTEAEPETAEEPAKKEKKKKRKHEEEETEMPAKKKEKSEKKKK 528
```

```
Sequence name: swall:Q9ZIU2

Sequence documentation:
ID   Q9ZIU2      PRELIMINARY;      PRT;   460 AA.
AC   Q9ZIU2;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   VIRULENT STRAIN ASSOCIATED LIPOPROTEIN.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]

Alignment of: EMBOSS_001 x Q9ZIU2   ..

Alignment segment 1/1:

                     Quality:   77.00                     Escore:       0
             Matching length:     115               Total length:     117
 Matching Percent Similarity:   66.96  Matching Percent Identity:   19.13
    Total Percent Similarity:   65.81     Total Percent Identity:   18.80
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     180 SQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQ 229
         ::|:  |:|  :    |:::  :  :  |:  :   ::|:|:||::: |
      52 TEEELKKKQQEEELKKKQQEEELKKKQQEEELKKKQQEEELKKKQQEEEL 101
                  .         .         .         .         .
     230 REGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKE 279
         :: ::: : :::::::::  |||  :::: | || :: : |: ::   |:
     102 KKKQQEEELKKKQQEEELKKKQQ..EEELKKKQQEEELKKKQQEEELKKK 149
                  .
     280 KKESRVSKRLSDSTTKK                                 296
         ::|::::|: :::  ||
     150 QQEEELKKKQQEEELKK                                 166



Sequence name: swall:O50870

Sequence documentation:
ID   O50870      PRELIMINARY;      PRT;   451 AA.
AC   O50870;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 54.3 KDA PROTEIN.
GN   BBI16.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp28-4.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x O50870   ..

Alignment segment 1/1:

                     Quality:   77.00                     Escore:       0
```

```
             Matching length:     115                Total length:     117
 Matching Percent Similarity:   66.96   Matching Percent Identity:   19.13
    Total Percent Similarity:   65.81      Total Percent Identity:   18.80
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     180 SQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQ 229
         ::|:  |:|  :    |:::  :  :  |:  :   ::|:|:||::: |
      52 TEEELKKKQQEEELKKKQQEEELKKKQQEEELKKKQQEEELKKKQQEEEL 101
                  .         .         .         .         .
     230 REGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKE 279
         :: ::: : :::::::::  |||  :::: | || :: : |: ::   |:
     102 KKKQQEEELKKKQQEEELKKKQQ..EEELKKKQQEEELKKKQQEEELKKK 149
                  .
     280 KKESRVSKRLSDSTTKK                                 296
         ::|::::|: :::  ||
     150 QQEEELKKKQQEEELKK                                 166



Sequence name: swall:Q9DGL1

Sequence documentation:
ID   Q9DGL1      PRELIMINARY;      PRT;   791 AA.
AC   Q9DGL1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RETINITIS PIGMENTOSA GTPASE REGULATOR-LIKE PROTEIN (FRAGMENT).
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Alignment of: EMBOSS_001 x Q9DGL1   ..

Alignment segment 1/1:

                     Quality:   77.00                      Escore:       0
             Matching length:      81                Total length:      82
 Matching Percent Similarity:   74.07   Matching Percent Identity:   24.69
    Total Percent Similarity:   73.17      Total Percent Identity:   24.39
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |:|: |::|::: ::|||:| ::|:: |:::   ::|:  ::  : :: :
     516 EKEESNEEEQSSTEKEGESNEEEQSSTEKEESNEEEQSSTEKEEESNEEE 565
                  .         .         .
     266 RQSM.KEPSQVKAKEKKESRVSKRLSDSTTKK                  296
         ::|  || |: : :|::||: ::: |::  :|
     566 QSSTEKEESNEEEEEEEEESESNEEESEEEDQK                 597
```

```
Sequence name: swall:CIK2_HUMAN

Sequence documentation:
ID   CIK2_HUMAN     STANDARD;      PRT;   499 AA.
AC   P16389;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.2 (RBK2) (HBK5) (NGK1)
DE   (MK2) (HUKIV).
GN   KCNA2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Alignment of: EMBOSS_001 x CIK2_HUMAN   ..

Alignment segment 1/1:

                     Quality:   76.00                         Escore:       0
             Matching length:      88                   Total length:      97
 Matching Percent Similarity:   65.91   Matching Percent Identity:   32.95
    Total Percent Similarity:   59.79      Total Percent Identity:   29.90
                        Gaps:       2

Alignment:
                 .         .         .         .         .
     77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
        | | ::  :|:| ||:|:::  ::  :  : |   ::||: |:    |
    220 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIVA....IIP 265


                 .         .         .         .
    127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
        : |::| : :   ::|:| :||      |   :||:||||::|| ||
    266 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    307



Sequence name: swall:CIK2_MOUSE

Sequence documentation:
ID   CIK2_MOUSE     STANDARD;      PRT;   499 AA.
AC   P15386; Q02010;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.2 (RBK2) (RCK5) (NGK1)
DE   (MK2) (HUKIV).
GN   KCNA2.
OS   Mus musculus (Mouse), and
OS   Rattus norvegicus (Rat).

Alignment of: EMBOSS_001 x CIK2_MOUSE   ..

Alignment segment 1/1:

                     Quality:   76.00                         Escore:       0
             Matching length:      88                   Total length:      97
 Matching Percent Similarity:   65.91   Matching Percent Identity:   32.95
    Total Percent Similarity:   59.79      Total Percent Identity:   29.90
                        Gaps:       2
```

```
Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | ::  :|:| ||:|:::  ::  :  : |   ::||: |:    |
     220 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIVA....IIP 265
                  .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
         : |::| : :   ::|:| :||      |   :||:||||::|| ||
     266 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    307



Sequence name: swall:CIK4_MOUSE

Sequence documentation:
ID   CIK4_MOUSE     STANDARD;      PRT;   654 AA.
AC   Q61423;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.4.
GN   KCNA4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Alignment of: EMBOSS_001 x CIK4_MOUSE   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      95                Total length:     106
 Matching Percent Similarity:   69.47   Matching Percent Identity:   23.16
    Total Percent Similarity:   62.26      Total Percent Identity:   20.75
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      76 DDGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDIT 125
         :| | ::  :|:: ||:|:::  :: |:    |   ::||: ||    |
     369 NDPFFIVETVCIVWFSFEFVVRCFACTSQALFFKNIMNIIDIVS....IL 414
                  .         .         .         .         .
     126 SFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYK 175
          : |::|        : :| : :::::: :: :: ::|::||:|: :::|
     415 PYFITLG.......TDLAQQQGGGNGQQQQAMSFAILRIIRLVRVFRIFK 457

     176 AASYSQ                                             181
          : :|:
     458 LSRHSK                                             463



Sequence name: swall:Q25920

Sequence documentation:
ID   Q25920      PRELIMINARY;      PRT;  1510 AA.
AC   Q25920;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MATURE-PARASITE-INFECTED ERYTHROCYTE SURFACE ANTIGEN.
```

```
GN   MESA.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x Q25920   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:     115                Total length:     126
 Matching Percent Similarity:   70.43   Matching Percent Identity:   26.96
    Total Percent Similarity:   64.29      Total Percent Identity:   24.60
                        Gaps:       4

Alignment:
                  .         .         .         .         .
     172 KLYKAASYSQEQAFKRQPIRTQTTKKSK...ATIYPSVYENINNGPV... 215
         ::|: ::|::  :  |:  ::: |::||   ::   : ||| :|: |
     310 RIYEETKYNKITSEFRETENVKITEESKDREGNKVSGPYENSENSNVTSE 359
                  .         .         .         .         .
     216 ..EQENLNKKENNTEQREGED.NANSQENRENQNMVAKQQQIDNSIPKVQ 262
           | ::| :||:|  :: ||: | :::|| |: : :::|:  :|:  : :
     360 SEETKKLAEKEENEGEKLGENVNDGASENSEDPKKLTEQE..ENGTKESS 407
                  .         .
     263 QIRRQSMKEPSQVKAKEKKESRVSKR                         288
         :  :::  | :: || :||:|: :|:
     408 EETKDDKPEENEKKADNKKKSKKKKK                         433



Sequence name: swall:Q19974

Sequence documentation:
ID   Q19974      PRELIMINARY;      PRT;   754 AA.
AC   Q19974;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CODED FOR BY C. ELEGANS CDNA YK69D11.5.
GN   F32E10.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x Q19974   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:     138                Total length:     157
 Matching Percent Similarity:   74.64   Matching Percent Identity:   22.46
    Total Percent Similarity:   65.61      Total Percent Identity:   19.75
                        Gaps:       3

Alignment:
                  .         .         .         .         .
     139 AKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAFKRQ 188
         |::|:| ::::| :| |:|:  :: |:   |: ||::: ::: ::: ::
     463 ASAANQKAETKKEKKKSKKSDAATSLLADDRFKKLFESEDFEVDETSEQF 512
                  .         .         .         .         .
```

```
     189 PIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANS 238
           ::: ::| ||           :|||| |: ::::::::  ::|::|::
     513 LKNASIAQKMKA...........KGPVEVESEGEESEDNDASDSENEASG 551
                  .         .         .         .         .
     239 QENRENQN...MVAKQQQID.....NSIPKVQQIRRQSMKEPSQVKAKEK 280
         :::: : :   || :::::|     :|  : ::| ::   : ::|:| :|
     552 SDSEPKGSASWMVEENREVDMASDLESASSDEDIDQSLKAKRAKVRAAKK 601

     281 KESRVSK                                            287
         :| :::|
     602 SERNIQK                                            608



Sequence name: swall:O01900

Sequence documentation:
ID   O01900      PRELIMINARY;      PRT;  1621 AA.
AC   O01900;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CODED FOR BY C. ELEGANS CDNA YK144H11.5.
GN   F59E12.9.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.

Alignment of: EMBOSS_001 x O01900   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      99                Total length:     104
 Matching Percent Similarity:   67.68   Matching Percent Identity:   21.21
    Total Percent Similarity:   64.42      Total Percent Identity:   20.19
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     196 KKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQ 245
         |:|| ::  :  |::::   |:|  :|||:: |:|:::     :|::: :
     100 KRSKRALENDSDEDVGDDGKEEEMKKKKEKEEEERRAK.....EERKKAK 144
                  .         .         .         .         .
     246 NMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTK 295
         :   :|:::::   |:::   :  |: ::|:|::|:: :  :  ||:  |
     145 KEKQRQKELEERNRKAEEAASKRRKREAEVEAERKRKRKEISVDSDDDDK 194

     296 KVII                                               299
          :::
     195 PALL                                               198



Sequence name: swall:O15763

Sequence documentation:
ID   O15763      PRELIMINARY;      PRT;  1969 AA.
AC   O15763;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYBRID HISTIDINE KINASE DHKB.
GN   DHKB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O15763   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      77                Total length:      81
 Matching Percent Similarity:   74.03   Matching Percent Identity:   25.97
    Total Percent Similarity:   70.37      Total Percent Identity:   24.69
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     209 NINNGPVEQENLNKKENNTEQREGEDNANSQENREN....QNMVAKQQQI 254
         | ||: ::::  |: |||::::::::| |:::|::|    :| : |:::
      14 NNNNNNIDSNINNNLENNNNKNNNNNNNNNNNNNNNNNNNIENSIDKNNKE 63
                  .         .         .
     255 DNSIPKVQQIRRQSMKEPSQVKAKEKKESRV                    285
         |||:  |:: |::  :  |:    |:: ::::
      64 DNSLVGVNSHRKHRTRLKSKKGNKHETKKEI                    94



Sequence name: swall:O96146

Sequence documentation:
ID   O96146      PRELIMINARY;   PRT;   508 AA.
AC   O96146;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   HYPOTHETICAL 61.3 KDA PROTEIN.
GN   PFB0225C.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O96146   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      72                Total length:      72
 Matching Percent Similarity:   75.00   Matching Percent Identity:   20.83
    Total Percent Similarity:   75.00      Total Percent Identity:   20.83
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     217 QENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRR 266
         ::|  |||:|::::| : | ::::::::::   |::: :::  | :: ::
     363 KKNEAKKEKNNKKKEKQKNKKNEKEKNKKKEKEKNKKKEKEKNKKKEKEK 412
```

```
                  .         .
     267 QSMKEPSQVKAKEKKESRVSKR                                288
         :: || :: | |||:::: :::
     413 SKKKEKEKNKKKEKEKNKKKEK                                434



Sequence name: swall:O96242

Sequence documentation:
ID   O96242      PRELIMINARY;   PRT;   1351 AA.
AC   O96242;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   HYPOTHETICAL 161.8 KDA PROTEIN.
GN   PFB0745W.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;

Alignment of: EMBOSS_001 x O96242   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      70                Total length:      71
 Matching Percent Similarity:   75.71   Matching Percent Identity:   27.14
    Total Percent Similarity:   74.65      Total Percent Identity:   26.76
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     228 EQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKA 277
         |::| |:: ::::::|:::   |::: |::  | :: :::: || :: |
     868 EEEKEEEEEKEKEKEEEEEDEKEEEEDKEKEKEEEKEEEKDKEEEDEKD 917
                  .         .
     278 KEK.KESRVSKRLSDSTTKKV                                 297
         ||| ||::::|::: :  |||
     918 KEKEKEEEIQKKVKKEIQKKV                                 938



Sequence name: swall:O97140

Sequence documentation:
ID   O97140      PRELIMINARY;   PRT;   1231 AA.
AC   O97140;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   PRESPORE-SPECIFIC PROTEIN (FRAGMENT).
GN   PSPC.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x O97140   ..

Alignment segment 1/1:
```

```
                     Quality:   76.00                  Escore:       0
             Matching length:      92            Total length:      96
 Matching Percent Similarity:   80.43  Matching Percent Identity:   21.74
    Total Percent Similarity:   77.08     Total Percent Identity:   20.83
                        Gaps:       1

Alignment:
                  .         .         .         .         .
     211 NNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPK 260
         ||:| :: | | ::||:: :::::| |:::|::|:|   :::: :|:
     828 NNNPNNNPNNNPNNNNNNNPNNNNNNNNNNNNNNNNNNNNNNNNNNN... 874
                  .         .         .         .
     261 VQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVIILVILLLL     306
          :: : ::::: :: : :::|::: ::: :::|:: :||::|:||:
     875 .NNTNYNNFNNTNNNNNNSNKNNNNNNNNNNNTTNNTIIIIIILLI     919



Sequence name: swall:Q9Y114

Sequence documentation:
ID   Q9Y114      PRELIMINARY;      PRT;   656 AA.
AC   Q9Y114;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   BCDNA:GH10229 PROTEIN.
GN   BCDNA:GH10229 OR CG8042.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9Y114   ..

Alignment segment 1/1:

                     Quality:   76.00                  Escore:       0
             Matching length:     133            Total length:     145
 Matching Percent Similarity:   63.91  Matching Percent Identity:   25.56
    Total Percent Similarity:   58.62     Total Percent Identity:   23.45
                        Gaps:       3

Alignment:
                  .         .         .         .         .
     136 SVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAASYSQEQAF 185
         ::||:: :  |::: :: ||: | |     :||  | |   |     |
     222 ALAAEKRAAASKVEGSATTSGAAARNFATNNLIPAVPLMTPAVPVPTQ.. 269
                  .         .         .         .         .
     186 KRQPIRTQTTKKSKATI..YPSVYENINNGPVEQENLNKKENNTEQREGE 233
           :|:::| : :|:  :    :: |: ::  ||:|:  :|||: ::|:
     270 ..RPLEAQDNTESEERLAEVRNILEQKRKERVEEEKRMEKENELRRRRDG 317
                  .         .         .         .
     234 DNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAK      278
          :|:||: |      ||:|:: |   :::: |::::   ::::|:
     318 REAQSQQAR......AKEQELKNMQEQIKRERQEELAARERIRAQ      356
```

```
Sequence name: swall:Q9VB01

Sequence documentation:
ID   Q9VB01      PRELIMINARY;      PRT;  1010 AA.
AC   Q9VB01;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG12877 PROTEIN.
GN   CG12877.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Alignment of: EMBOSS_001 x Q9VB01   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      81                Total length:      81
 Matching Percent Similarity:   74.07   Matching Percent Identity:   14.81
    Total Percent Similarity:   74.07      Total Percent Identity:   14.81
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         |:::::   :| |::|: : :::::::|:::  :|::: | :  |:::::
     147 EEQKIDTVAENEEDQENTQLTDARDSHEKDRKPSKERERDIEKKKIRELK 196
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTKK                     296
         |:::|: :: : ::::::: :::  |   :|
     197 REKIKDRDRDRDRDREKDKEKSKERDREKEK                     227



Sequence name: swall:Q9U0K4

Sequence documentation:
ID   Q9U0K4      PRELIMINARY;      PRT;   622 AA.
AC   Q9U0K4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 73.3 KDA PROTEIN.
GN   MAL4P2.18.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;

Alignment of: EMBOSS_001 x Q9U0K4   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      70                Total length:      75
 Matching Percent Similarity:   80.00   Matching Percent Identity:   31.43
    Total Percent Similarity:   74.67      Total Percent Identity:   29.33
                        Gaps:       2

Alignment:
```

```
                  .         .         .         .         .
     208 ENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNS 257
         :||||:  :: | |:: ||:|: :::|| |:::| :|:: : ::::|:||
     156 DNINNN..DNINNNENINNNENINNNDNINNNDNINNNDNINNNDNINNS 203
                  .         .
     258 IPKVQQIRRQSMKEPSQVKAKEKKE                          282
             : : |: :|:  ::::|:||:
     204 ...NRPVVRKVIKRVIRTRVKNKKD                          225



Sequence name: swall:Q9GV71

Sequence documentation:
ID   Q9GV71      PRELIMINARY;      PRT;  1543 AA.
AC   Q9GV71;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MKPA PROTEIN.
GN   MKPA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x Q9GV71   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      51                Total length:      51
 Matching Percent Similarity:   80.39   Matching Percent Identity:   31.37
    Total Percent Similarity:   80.39      Total Percent Identity:   31.37
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     207 YENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDN 256
         |:: ||   :|:| |:::|| :|:::::| |: :|::| |: :::::::|
     637 YNHSNNDNNSQNNQNNNNNNHNQNNNNNNNNNGNNHNNGNFNSNNNSNFSN 686

     257 S                                                  257
         |
     687 S                                                  687



Sequence name: swall:Q28293

Sequence documentation:
ID   Q28293      PRELIMINARY;      PRT;   499 AA.
AC   Q28293;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   DELAYED RECTIFIER K+ CHANNEL.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
```

```
Alignment of: EMBOSS_001 x Q28293   ..

Alignment segment 1/1:

                     Quality:   76.00                    Escore:      0
             Matching length:      88              Total length:     97
 Matching Percent Similarity:   65.91   Matching Percent Identity:   32.95
    Total Percent Similarity:   59.79      Total Percent Identity:   29.90
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | ::  :|:| ||:|:::  ::  :  : |   ::||: |:    |
     220 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIVA....IIP 265
                  .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
         : |::| : :   ::|:| :||      |   :||:||||::|| ||
     266 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    307



Sequence name: swall:Q9MYX3

Sequence documentation:
ID   Q9MYX3      PRELIMINARY;      PRT;   499 AA.
AC   Q9MYX3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-SEP-2001 (TrEMBLrel. 18, Last annotation update)
DE   POTASSIUM CHANNEL SUBUNIT KV 1.2.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q9MYX3   ..

Alignment segment 1/1:

                     Quality:   76.00                    Escore:      0
             Matching length:      88              Total length:     97
 Matching Percent Similarity:   65.91   Matching Percent Identity:   32.95
    Total Percent Similarity:   59.79      Total Percent Identity:   29.90
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | ::  :|:| ||:|:::  ::  :  : |   ::||: |:    |
     220 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIVA....IIP 265
                  .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
         : |::| : :   ::|:| :||      |   :||:||||::|| ||
     266 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    307
```

```
Sequence name: swall:Q09080

Sequence documentation:
ID   Q09080      PRELIMINARY;      PRT;   279 AA.
AC   Q09080;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL PROTEIN (CLONE KC19) (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

Alignment of: EMBOSS_001 x Q09080   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      88                Total length:      97
 Matching Percent Similarity:   65.91   Matching Percent Identity:   32.95
    Total Percent Similarity:   59.79      Total Percent Identity:   29.90
                        Gaps:       2

Alignment:
                  .         .         .         .         .
      77 DGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQILDITS 126
         | | ::  :|:| ||:|:::  ::  :  : |   ::||: |:    |
     139 DPFFIVETLCIIWFSFEFLVRFFACPSKAGFFTNIMNIIDIVA....IIP 184
                  .         .         .         .
     127 FNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKL    173
         : |::| : :   ::|:| :||      |   :||:||||::|| ||
     185 YFITLGTELAEKPEDAQQGQQA.....MSLAILRVIRLVRVFRIFKL    226




Sequence name: swall:O31329

Sequence documentation:
ID   O31329      PRELIMINARY;      PRT;   363 AA.
AC   O31329;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ERPM.
GN   ERPM OR BBO40.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-7.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

Alignment of: EMBOSS_001 x O31329   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      79                Total length:      80
 Matching Percent Similarity:   74.68   Matching Percent Identity:   24.05
    Total Percent Similarity:   73.75      Total Percent Identity:   23.75
                        Gaps:       1
```

```
Alignment:
                  .         .         .         .         .
     216 EQENLNKKENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIR 265
         ||:: ::| :: :|:: |:: ||:|::::|:  || :   :   :::| :
     138 EQDSKKEKVEKQSQKQKEEERNSKEEQQKQE.EAKARADREREERLKQQE 186
                  .         .         .
     266 RQSMKEPSQVKAKEKKESRVSKRLSDSTTK                     295
         :: ::| ::|||:::|::| ::: :::  |
     187 QKRQQEEARVKAEKEKQEREEQQKQEEEKK                     216



Sequence name: swall:O40947

Sequence documentation:
ID   O40947      PRELIMINARY;      PRT;  1089 AA.
AC   O40947;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ORF 73.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=37296;

Alignment of: EMBOSS_001 x O40947   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:     104                Total length:     104
 Matching Percent Similarity:   67.31   Matching Percent Identity:   17.31
    Total Percent Similarity:   67.31      Total Percent Identity:   17.31
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     181 QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQR 230
         |::  :::| : :  :::     |:  |  :: | :||  :: |:: :::
     530 QREPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQDEQQQDEQ 579
                  .         .         .         .         .
     231 EGEDNANSQENRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEK 280
         : ::: ::::::::|::  :||: :::  : || ::|:: | :| : :|:
     580 QQDEQQQDEQQQDEQQQDEQQQDEQEQQDEQQQDEQQQQDEQEQQEEQEQ 629

     281 KESR                                               284
         :|::
     630 QEEQ                                               633



Sequence name: swall:Q9DDN7

Sequence documentation:
ID   Q9DDN7      PRELIMINARY;      PRT;   651 AA.
AC   Q9DDN7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   KV3.3 POTASSIUM CHANNEL SUBUNIT PRECURSOR.
OS   Apteronotus leptorhynchus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Gymnotiformes; Apteronotidae; Apteronotus.

Alignment of: EMBOSS_001 x Q9DDN7   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      90                Total length:     105
 Matching Percent Similarity:   64.44   Matching Percent Identity:   34.44
    Total Percent Similarity:   55.24      Total Percent Identity:   29.52
                        Gaps:       2

Alignment:
                 .         .         .         .         .
     69 IAFDKRADDGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTV 118
        |:::  :|: :  :  :|:| |:||:::      |  : |   |:||: |
    254 IVYEVVTDSWLTYVEGVCVIWFTIEVFMRVTFCPDKAEFFKSLLNIIDFV 303
                 .         .         .         .         .
    119 SQILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLI 168
        :    |  | : |:|:|            ::||:|     :||||:||::
    304 A....ILPFYLEVALSG...........LSSKAAKDVLGFLRVVRFVRIL 338

    169 RIVKL                                              173
        || ||
    339 RIFKL                                              343



Sequence name: swall:AAK54090

Sequence documentation:
ID   AAK54090    PRELIMINARY;      PRT;  1736 AA.
AC   AAK54090;
DT   14-JUN-2001 (EMBLrel. 63, Created)
DT   14-JUN-2001 (EMBLrel. 63, Last sequence update)
DT   14-JUN-2001 (EMBLrel. 63, Last annotation update)
DE   HISTIDINE KINASE DHKI.
GN   DHKI.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;

Alignment of: EMBOSS_001 x AAK54090   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      88                Total length:      94
 Matching Percent Similarity:   70.45   Matching Percent Identity:   25.00
    Total Percent Similarity:   65.96      Total Percent Identity:   23.40
                        Gaps:       1

Alignment:
                 .         .         .         .         .
    176 AASYSQEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKK.. 223
        :::: ::| :   |  :|  :: : :  |   :| :||  :|::|:|:
    364 SGEFINDQEIIITPTSNQQQSSLSKSNRPRSQSNCSNGNKSQNRLSKNYS 413
```

```
                  .         .         .         .
     224 ....ENNTEQREGEDNANSQENRENQNMVAKQQQIDNSIPKVQQ       263
              :|::::::::| |:::|::|:| :::||| : |  ::||
     414 TTTTTTNNNNNNNNNNNNNNNNNNNNNNSISQQQQTQVSTQQTQQ       457



Sequence name: swall:AAK72247

Sequence documentation:
ID   AAK72247    PRELIMINARY;      PRT;   491 AA.
AC   AAK72247;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   MB2 (FRAGMENT).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x AAK72247   ..

Alignment segment 1/1:

                     Quality:   76.00                      Escore:       0
             Matching length:      68                Total length:      68
 Matching Percent Similarity:   69.12   Matching Percent Identity:   27.94
    Total Percent Similarity:   69.12      Total Percent Identity:   27.94
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     202 IYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQ 251
         | |:  ::|||:   ::|:|:| |: :::  ||: |::|| :|:|   ::
     273 IIPEGKNKINNNIDVKHNINNKLNEINEEPYEDTHNKEENSSNKNDNDEK 322
                  .
     252 QQIDNSIPKVQQIRRQSM                                 269
         :: :|:    : |::::|
     323 RKEENNNITRRYIKNDQM                                 340



Sequence name: swall:AAK72251

Sequence documentation:
ID   AAK72251    PRELIMINARY;      PRT;   414 AA.
AC   AAK72251;
DT   06-SEP-2001 (EMBLrel. 63, Created)
DT   06-SEP-2001 (EMBLrel. 63, Last sequence update)
DT   06-SEP-2001 (EMBLrel. 63, Last annotation update)
DE   MB2 (FRAGMENT).
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]

Alignment of: EMBOSS_001 x AAK72251   ..

Alignment segment 1/1:
```

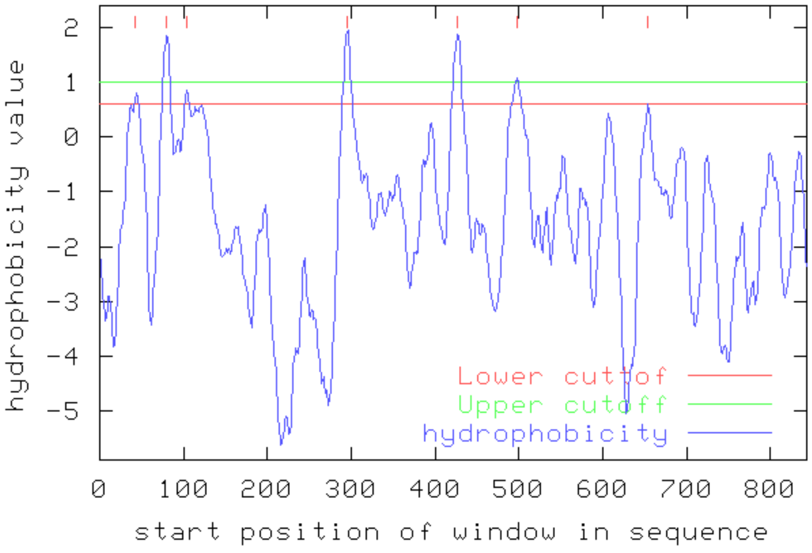
```
                     Quality:   76.00                     Escore:       0
             Matching length:      68               Total length:      68
 Matching Percent Similarity:   69.12   Matching Percent Identity:   27.94
    Total Percent Similarity:   69.12      Total Percent Identity:   27.94
                        Gaps:       0

Alignment:
                  .         .         .         .         .
     202 IYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQENRENQNMVAKQ 251
         | |:  ::|||:   ::|:|:| |: :::  ||: |::|| :|:|   ::
     273 IIPEGKNKINNNIDVKHNINNKLNEINEEPYEDTHNKEENSSNKNDNDEK 322
                  .
     252 QQIDNSIPKVQQIRRQSM                                 269
         :: :|:    : |::::|
     323 RKEENNNITRRYIKNDQM                                 340
```

7.1.22 Toppred prediction mit pAC

Algorithm specific parameters

```
Algorithm specific parameters:

Full window size : 21
Core window size : 11
Wedge window size: 5
Using hydrophobicity file: GES-scale

Cutoff for certain transmembrane segments: 1.00
Cutoff for putative transmembrane segments: 0.60
Critical distance between 2 transmembrane segments: 2

Critical loop length:  60

Kingdom: eucaryote

Using cyt/ext file: CYTEXT-scale
```

Hydrophobicity plot



Sequence

```
Sequence : ,  (865 res)
MSEAQVNPEVDNQSVISEDSNNEDVDVFGQDQEIDKFPLLTSILDSSWSQILVNLLTIYA
LFGDDIRIIAFDKRADDGFDVITIICMIIFSIEIIIASLVRTDYFNSFFFWLDIISTVSQ
ILDITSFNIAVGLQGSVAAKSASQLSQANKASKTSSKAIRVVRLVRLIRIVKLYKAASYS
QEQAFKRQPIRTQTTKKSKATIYPSVYENINNGPVEQENLNKKENNTEQREGEDNANSQE
NRENQNMVAKQQQIDNSIPKVQQIRRQSMKEPSQVKAKEKKESRVSKRLSDSTTKKVIIL
VILLLLIMPLFSSDYYFEPSYSLAYAAEYVRVVAEIPNTKLTEINQTIYFVIDQHKSFDT
```

```
PVGYITNPFTEIENYETPSYQYLRESSKSYYFELVDPVLVGLEYIGDPVILFVSDNSDVE
SKNSIINIVNTLFVSAVLLFGAIAFSNDANNVALKPIDRMIVKVNLIAKNPQLAKEMKLE
SDGTQKETTQIENAIIKIGALLALGFGDAGSAIIGTNMASSGDVDPMLPGKRKWAIYGFC
DIRNFTDATEVLQKDVMLFVNNIAEIVHSMVDRYQGSANKNIGDAFLLVWKINDSQWYED
GNEIKWSNLTFINQLADCALIAFMKIYAKINREPKILEYRNDERLSQRLPGYKVKMGFGL
HIGWGIEGAIGSEFKIDASYLSPNVNMASRLEAATKQYGVAVLISSELHQYFSNEVKKYT
RQIDKVTVKGSVKPIGLFTVEMEADDLPPSKQDYPQEEKQQVMYEKKQIFLQQLESGDFN
AELYIKRNKDLGLITKNVNQEFLHQFGQGFQGYLLGNWKEAHMLFEKAKQIRPNDGPIST
LFNVMGETNFKAPSDWKGYRELTEK
```

Transmembran segments

```
Found: 7 segments

Candidate membrane-spanning segments:

 Helix Begin - End   Score Certainity
     1    43 - 63    0.815 Putative
     2    80 - 100   1.855 Certain
     3   104 - 124   0.863 Putative
     4   296 - 316   1.938 Certain
     5   427 - 447   1.877 Certain
     6   498 - 518   1.092 Certain
     7   654 - 674   0.605 Putative
```

Topologies

```
Total of 8 structures are to be tested

HEADER     START    STOP     LEN  PROB    HP   DARGLYS   DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY      1                   1.00          -8.00      1.53     -1.00     0.62
TOPOLOGY                                        N-out     N-out     N-out
EXT_LOOP       1      79      79               (  4.00)    0.12
TRANSMEM      80     100      21  1.00  1.86
CYT_LOOP     101     295     195               ( 35.00)    0.59
TRANSMEM     296     316      21  1.00  1.94
EXT_LOOP     317     426     110               (  6.00)    1.28
TRANSMEM     427     447      21  1.00  1.88
CYT_LOOP     448     497      50                  8.00  ( -0.29)
TRANSMEM     498     518      21  1.00  1.09
EXT_LOOP     519     865     347               ( 43.00)    0.72


HEADER     START    STOP     LEN  PROB    HP   DARGLYS   DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY      2                   0.54          -8.00      1.45     -1.00     0.83
TOPOLOGY                                        N-out     N-out     N-out
EXT_LOOP       1      42      42                  1.00  (  0.56)
TRANSMEM      43      63      21  0.54  0.81
CYT_LOOP      64      79      16                  3.00  ( -0.85)
TRANSMEM      80     100      21  1.00  1.86
EXT_LOOP     101     103       3                  1.00  ( -1.26)
TRANSMEM     104     124      21  0.66  0.86
CYT_LOOP     125     295     171               ( 34.00)    0.56
TRANSMEM     296     316      21  1.00  1.94
EXT_LOOP     317     426     110               (  6.00)    1.28
TRANSMEM     427     447      21  1.00  1.88
```

```
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    865    347           ( 43.00)    0.72


HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    3                  0.01          -8.00     0.27     -1.00      0.62
TOPOLOGY                                  N-out    N-out     N-out
EXT_LOOP      1     79     79           (  4.00)    0.12
TRANSMEM     80    100     21 1.00 1.86
CYT_LOOP    101    295    195           ( 35.00)    0.59
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    653    135           ( 18.00)    0.24
TRANSMEM    654    674     21 0.01 0.61
CYT_LOOP    675    865    191           ( 25.00)    0.79


HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    4                  0.01          -8.00     0.18     -1.00      0.83
TOPOLOGY                                  N-out    N-out     N-out
EXT_LOOP      1     42     42              1.00  (  0.56)
TRANSMEM     43     63     21 0.54 0.81
CYT_LOOP     64     79     16              3.00  ( -0.85)
TRANSMEM     80    100     21 1.00 1.86
EXT_LOOP    101    103      3              1.00  ( -1.26)
TRANSMEM    104    124     21 0.66 0.86
CYT_LOOP    125    295    171           ( 34.00)    0.56
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    653    135           ( 18.00)    0.24
TRANSMEM    654    674     21 0.01 0.61
CYT_LOOP    675    865    191           ( 25.00)    0.79


HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    5                  0.66           7.00    -1.32     -1.00      0.62
TOPOLOGY                                   N-in     N-in     N-out
CYT_LOOP      1     79     79           (  4.00)    0.12
TRANSMEM     80    100     21 1.00 1.86
EXT_LOOP    101    103      3              1.00  ( -1.26)
TRANSMEM    104    124     21 0.66 0.86
CYT_LOOP    125    295    171           ( 34.00)    0.56
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    865    347           ( 43.00)    0.72
```

```
HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    6                  0.54        7.00    -1.41     -1.00     0.83
TOPOLOGY                                   N-in     N-in     N-out
CYT_LOOP      1     42     42              1.00  (  0.56)
TRANSMEM     43     63     21 0.54 0.81
EXT_LOOP     64     79     16              3.00  ( -0.85)
TRANSMEM     80    100     21 1.00 1.86
CYT_LOOP    101    295    195           ( 35.00)    0.59
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    865    347           ( 43.00)    0.72

HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    7                  0.01        7.00    -0.05     -1.00     0.62
TOPOLOGY                                   N-in     N-in     N-out
CYT_LOOP      1     79     79           (  4.00)    0.12
TRANSMEM     80    100     21 1.00 1.86
EXT_LOOP    101    103      3              1.00  ( -1.26)
TRANSMEM    104    124     21 0.66 0.86
CYT_LOOP    125    295    171           ( 34.00)    0.56
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    653    135           ( 18.00)    0.24
TRANSMEM    654    674     21 0.01 0.61
CYT_LOOP    675    865    191           ( 25.00)    0.79

HEADER    START   STOP    LEN PROB   HP  DARGLYS  DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY    8                  0.01        7.00    -0.14     -1.00     0.83
TOPOLOGY                                   N-in     N-in     N-out
CYT_LOOP      1     42     42              1.00  (  0.56)
TRANSMEM     43     63     21 0.54 0.81
EXT_LOOP     64     79     16              3.00  ( -0.85)
TRANSMEM     80    100     21 1.00 1.86
CYT_LOOP    101    295    195           ( 35.00)    0.59
TRANSMEM    296    316     21 1.00 1.94
EXT_LOOP    317    426    110           (  6.00)    1.28
TRANSMEM    427    447     21 1.00 1.88
CYT_LOOP    448    497     50              8.00  ( -0.29)
TRANSMEM    498    518     21 1.00 1.09
EXT_LOOP    519    653    135           ( 18.00)    0.24
TRANSMEM    654    674     21 0.01 0.61
CYT_LOOP    675    865    191           ( 25.00)    0.79
```