Gene Expression Studies on the Evolution of Development and Innate Immunity in the Nematode *Pristionchus pacificus*

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Amit Sinha
aus Dehra Dun, Indien

Tübingen
2012

Tag der mündlichen Prüfung:	21. Dezember 2012
Dekan:	Prof. Dr. W. Rosenstiel
1. Berichterstatter:	Prof. Dr. Ralf J. Sommer
2. Berichterstatter:	Prof. Dr. Rita Groß-Hardt

List of Publications

Teile der vorliegenden Arbeit wurden bereits verffentlicht oder wurden zur Verffentlichung eingereicht:

1. **Amit Sinha**, Ralf J Sommer, and Christoph Dieterich:
Divergent Gene Expression in the Conserved Dauer Stage of the Nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*.
BMC Genomics 2012, **13** (1): 254. doi:10.1186/1471-2164-13-254.

2. **Amit Sinha**, Robbie Rae, Igor Iatsenko, and Ralf J. Sommer:
System Wide Analysis of the Evolution of Innate Immunity in the Nematode Model Species *Caenorhabditis elegans* and *Pristionchus pacificus*.
PLoS ONE 2012, **7** (9): e44255. doi:10.1371/journal.pone.0044255.

3. Robbie Rae*, **Amit Sinha***, and Ralf J. Sommer:
Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode *Pristionchus pacificus*.
PLoS Pathogens 2012, **8** (8): e1002864. doi:10.1371/journal.ppat.1002864.
* = Authors contributed equally

4. Nadine Borchert*, Karsten Krug*, Florian Gnad, **Amit Sinha**, Ralf J Sommer, and Boris Macek:
Phosphoproteome of *Pristionchus pacificus* Provides Insights into Architecture of Signaling Networks in Nematode Models.
Molecular & Cellular Proteomics 2012, **11** (12):1631-1639. doi:10.1074/mcp.M112.022103.
* = Authors contributed equally

Acknowledgements

Writing these acknowledgements has been the most pleasant part of the entire thesis, but not the easiest one due to limited space. Many different people have contributed towards this thesis in their own special ways, and I am glad that I finally have an opportunity to thank at least some of them.

It is hard to overstate my heartfelt gratitude to my supervisor **Prof. Dr. Ralf J. Sommer**, who not only provided me the dream opportunity of working with him on the topics I liked, but also for being a superb mentor in many ways. He has given me great freedom to try new things and dabble in various projects, opportunities to collaborate with many brilliant people, and many enlightening and entertaining discussions in the lab, over lunch, in our coffee room, at "Sommer" parties and elsewhere. Most of all, I am grateful to him for imparting to me a unique, comprehensive perspective in biology.

Thanks are also due to **Prof. Dr. Rita Groß-Hardt** for accepting to be my co-supervisor and an evaluator of my thesis. I would also like to thank **Prof Boris Maček** and **Dr. Gáspár Jékely** for being examiners on my thesis defense committee. Thanks to **Melanie Mayer** for great help in translating my thesis summary into the German "'Zusammenfassung'.

I have benefited immensely from my collaborations with **Dr. Christoph Dieterich**, with whom I first started working in this lab and continued our projects together even after he moved to MDC Berlin. His inputs and support were a key factor in our success in designing a custom microarray platform for *Pristionchus pacificus*, which I could then use as the basic tool for my subsequent projects. Thanks to Christoph also for hosting me at MDC Berlin at various times which helped in moving things ahead faster. I am also grateful to **Dr. Mitch Levesque** and **Horst Geiger** from

Department 3 for sharing their expertise in running microarray experiments, and for help with troubleshooting the initial problems.

Dr. Robbie Rae has been very generous in sharing many of his great ideas with me, collaborating on many different experiments, as well as giving me a chance to follow-up on various interesting hypothesis that we generated from our microarray experiments on innate immunity and longevity. I am also grateful to **Igor Iatsenko** for collaborating on many of his exciting projects. I cannot thank you guys enough for being great colleagues as well as great friends.

I would also like to thank **Dr. Federico Brown**, **Dr. Nadine Borchert**, **Ruxandra Molnar**, **Gabi Bartelmes** and others with whom I worked on the huge deletion library project. That was one of the most exciting time during my Ph.D. research and I learnt a lot from all of you. **Dr. Akira Ogawa** has always been very helpful in sharing his wisdom on the "dauers". I must have bothered him a hundred times for various small things but he has always been very patient with me. I also appreciate the inputs from **Dr. Christian Rödelsperger**, **Andreas Weller**, **Ruxandra Molnar**, **Dr. Gabriel Markov**, **Dr. Adrian Streit** and the "hardcore" nematologist **Dr. Erik Ragsdale**.

Thanks to a diverse and multidisciplinary lab like ours, I had the unparalleled opportunity to be exposed to various aspects of evolutionary biology and learn from all of you. Interacting with all of you has been a great education for me. Additionally, I can now proudly say that I have friends in at least **15 countries** around the globe!

I have to thank our lab technicians **Hanh Witte** and **Gabi Bartelmes** for so many things that not all can be listed here. **Sandra Mäck** provided me tens of liters of M9 and S-medium for my liquid cultures, **Tamara Eirich** prepared large quantities of 10cm NGM agar plates and boxes of IPTG-RNAi plates. I am grateful to **Lorena Cali-Oezer** and **Kostadinka Krause** for taking taken great care of all administrative tasks, and for many great conversations during coffee-breaks.

I always looked forward to Thursdays to meet with our **evoJC** journal club members. Many thanks to all the people involved, for exciting discus-

sions on the latest happenings in evolutionary biology, and for keeping the group going.

It was a great pleasure to be in the company of many marvelous people: **Ruxandra and Boris**, **Simone Kienle**, **Isabella D'Anna**, **Andreas Weller**, **Dan and Barbara**, **Christian and Charlotte**, **Nadine**, **Erik**, **Alex and Gabi**, **Lisa Schuster**, **Metta Riebesell**, **Dr. Katy Morgan**, **Dr. Angela McGaughran and Niko**. You not only ensured a rocking atmosphere in the lab, but many of you also dragged me out of the lab (and drove me around in your cars) to see and do many wonderful things. I will always remember the times spent with you. I must also express my gratitude to various **"Max-Planck Indians"** and **"Tuebingen Indians"** with whom I spent a fantastic time, especially the weekends and festival celebrations. Most importantly for me, all of you have always been better and more enthusiastic cooks than me, and I have often exploited you for feeding me many delicious dishes!

Then there are people who although not physically here in Tübingen, have always been a source of support, inspiration and happiness, sometimes for more than a decade! It is impossible to list your contributions to all my journeys - scientific and others, so I would just mention your names: **Amar Chandra**, **Aditya**, **Mitesh Jain**, **Harsh Raju**, **Vikrant Kumar**, **Abhishek Bansal**, **JP Singh**, **Vinay Khandelwal**, **Gaurav Sinha** and other **3midians**, and my brother **Sumit Sinha**. I plunged into Biology only during my Masters, and made it through only because of my friends at the IIT Kanpur Fly Lab: **Manish**, **Hina**, **Shubho**, **Atif**, and specially **Sumbul**. I am also very grateful to my teachers **Dr. K. Subramaniam** and **Dr. S. Ganesh** at IIT Kanpur, for their invaluable help during my search for a Ph.D position.

Finally, **my parents**, for their infinite faith and belief in me, for being what they are, and for making me what I am.

Contents

Contents **vii**

Summary **xi**

Zusammenfassung **xiii**

1 Introduction **1**

1.1 Genetic basis of the evolution of biological diversity 1

1.1.1 Discovery of remarkable conservation of developmental control genes 2

1.1.2 Differential regulation of gene expression as a source of morphological diversity 2

1.2 Nematodes as models organisms. 4

1.2.1 Evolutionary history of nematodes 5

1.2.2 The nematode *Pristionchus pacificus* as a model organism to integrate evolution, development and ecology 6

1.3 Genomic approaches to study evolution in *P. pacificus* 8

1.3.1 Transcriptomic approaches to study evolution 11

1.4 The dauer polyphenism in *P. pacificus* and *C. elegans* 13

1.5 Innate immunity and response to pathogens in *P. pacificus* and *C. elegans* 16

1.6 The role of germline and somatic gonad in the regulation of longevity and innate-immunity in *P. pacificus* and *C. elegans* 17

1.7 Evolution of genome structure: Analysis of trans-splicing and operon-like organization of genes in *P. pacificus* genome 19

1.8 Tools for gene expression profiling: Microarrays 21

1.8.1 Design of custom microarrays or *P. pacificus* 22

1.9 Aims of the thesis 23

1.9.1 Phenotypic plasticity and evolutionary novelty: comparing dauer transcriptomes . . . 23
1.9.2 Evolution of innate immunity: genetic basis of differential susceptibility of *P. pacificus* and *C. elegans* to various pathogens . 23
1.9.3 Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in *P. pacificus* . . . 24
1.9.4 Functional genomic investigation of pathways regulated by the somatic gonad thse contribute towards the longevity of germline-deficient *C. elegans* . . . 24
1.9.5 Functional genomic investigation of novel *C. elegans* mutants *Cel-btr-1* and *Cel-btr-2*, which display enhanced resistance to the bacteria pathogen *Bacillus thuringiensis* . . . 24
1.9.6 Functional genomic investigation of enhanced pathogen resistance of *C. elegans* when pre-exposed to a related but non-pathogenic bacteria . . . 25
1.9.7 Characterization of the trans-splicing landscape, and operon like gene-clusters in *P. pacificus* genome using next generation sequencing . . . 25

2 Results **27**

2.1 Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans* . . . 27
2.1.1 Synopsis . . . 27
2.1.2 Contributions . . . 28
2.2 A system wide analysis of the evolution of innate immunity in the nematode model species *Caenorhabditis elegans* and *Pristionchus pacificus* 29
2.2.1 Synopsis . . . 29
2.2.2 Contributions . . . 30
2.3 Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode *Pristionchus pacificus* . . . 31
2.3.1 Synopsis . . . 31
2.3.2 Contributions . . . 32

2.4 Phosphoproteome of *Pristionchus pacificus* provides insights into architecture of signaling networks in nematode models 33
2.4.1 Synopsis 33
2.4.2 Contributions 34
2.5 Genes and developmental pathways regulating innate immunity and lifespan in *C. elegans* 35
2.5.1 Reproductive signaling genes that regulate extreme longevity and innate immunity in *C. elegans* 35
2.5.2 Transcriptome analysis of *C. elegans* mutants *btr-1* and *btr-2* that are resistant to the pathogenic bacteria *Bacillus thuringiensis* DB27 38
2.5.3 *Bacillus subtilis* increases *C. elegans* pathogen survival by activating innate immunity, stress resistance and longevity genes . . 40
2.6 Gene organization into operon-like clusters in *Pristionchus pacificus* is remarkably different from *Caenorhabditis elegans* 43
2.6.1 Synopsis 43
2.6.2 Results 43
2.6.3 Contributions 45

Bibliography **47**

Lebenslauf / Curriculum vitae **65**

Publications **67**

Summary

Evolutionary biology aims to explain the enormous biological diversity by employing diverse approaches including systematics, ecology, population genetics and developmental biology. Comparison of a set of traits across different species is a very fruitful approach in evolutionary biology. With this goal in mind, the nematode *Pristionchus pacificus* has been developed as a model system for comparative analysis vis-a-vis the extensively studied nematode model *Caenorhabditis elegans*, and serves to integrate insights from developmental biology, ecology and population genetics. Comparison of processes such as vulva development, sex-determination pathways, gonad development and developmental polyphenisms such as dauer development and mouth-form dimorphism has revealed that highly conserved genetic pathways can be used in different configurations to generate macroevolutionary changes. The sequencing of the whole genome of *P. pacificus* opened a new avenue for comparative genomics, e.g. it has a larger genome and a higher gene content than *C. elegans*. Many factors contribute to this increase in gene number, including expansion of certain gene families, horizontal gene transfer, and origin of lineage specific "pioneer" genes, many of which might have a potential role in its ecological adaptations related to beetle association. However, function for most of these genes is not yet characterized. Since gene expression is the first manifestation of gene activity and is dynamically regulated during development and response to environment, I took a comparative transcriptomics approach to study evolution across the two species. I first developed a microarray platform for gene expression profiling in *P. pacificus* and used it to characterize transcriptomes under various developmental and ecological contexts. These include dauer development, response to bacterial pathogens, and role of germline and somatic gonad in regulating lifespan and innate immunity. The key questions that I tried to answer are (1) what genes are regulated during a particular process (2) do orthologous genes show expression patterns in the two (3) what do the observed differences tell us about evolutionary process of gene expression evolution. Taken together, the results from these studies further add to the concept of developmental systems drift.

Zusammenfassung

Ein Ziel der Evolutionsbiologie ist die Erklärung der biologischen Vielfalt. Dazu werden unter anderem verschiedene Methode der Systematik, Ökologie, Populationsgenetik und Entwicklungsbiologie kombiniert. Der Vergleich morphologischen Eigenschaften zwischen verschiedenen Arten ist dabei eine sehr produktive Methode. Mit diesem Ziel wurde der Nematode *Pristionchus pacificus* als Modellsystem für Analysen im Vergleich zum ausführlich erforschten Nematodenmodell *Caenorhabditis elegans* entwickelt, was dazu dient, Erkenntnisse aus der Entwicklungsbiologie, ökologie und Populationsgenetik zu integrieren. Der Vergleich von Entwicklungsvorgängen, so wie z.B. die Vulva-Entwicklung, geschlechtsbestimmende Signalwege und Gonadenentwicklung und von Entwicklungs-Polyphenismen, wie Dauer-Entwicklung und Mundform-Dimorphismus, zeigte, dass hoch konservierte genetische Signalwege in verschiedenen Zusammenstellungen verwendet werden können, um makro-evolutionäre Veränderungen zu verursachen. Durch die Sequenzierung des vollständigen Genoms von *P. pacificus* ergaben sich neue Möglichkeiten für die vergleichende Genomik, z.B. hat *P. pacificus* ein größeres Genom und eine höhere Anzahl von Genen als *C. elegans*. Viele Faktoren tragen zu dieser erhöhten Anzahl von Genen bei, darunter Expansion bestimmter Genfamilien, horizontaler Gentransfer und Entstehung abstammungs-spezifischer "Pioniergene". Viele dieser Pioniergene spielen vielleicht eine Rolle bei den ökologischen Anpassungen von *P. pacificus* bezüglich der Käferassoziation. Für viele dieser Gene wurde eine Funktion jedoch noch nicht beschrieben. Da Genexpression die erste Manifestation von Genaktivität ist und während der Entwicklung und der Reaktion auf die Umwelt dynamisch reguliert wird, verwendete ich die vergleichende Transkriptomik um Evolutionprozesse zu untersuchen. Ich entwickelte zuerst eine Mikroarray-Plattform für Genexpressions-Analysen in *P. pacificus* und nutzte sie, um Transkriptome in verschiedenen Entwicklungs- und ökologischen Zusammenhängen zu beschreiben. Dazu gehören Dauer-Entwicklung, Reaktion auf bakterielle Pathogene und die Rolle der Keimbahn und der somatischen Gonade bei der Regulierung der Lebensdauer und der angeborenen Immunität. Die wichtigsten Fragen, die ich zu beantworten versuchte, sind (1) welche Gene werden während eines bestimmten Vorgangs reguliert, (2) zeigen orthologe Gene verschiedene Expressionsmuster in den zwei Arten und (3) was sagen uns die beobachteten Unterschiede über den Vorgang der Evolution der Genexpression. Insgesamt leisten die Ergebnisse dieser Untersuchungen einen weiteren Beitrag zur Idee der Verschiebung der Entwicklungssysteme.

1

Introduction

1.1 Genetic basis of the evolution of biological diversity

About 1.2 million species of various organisms are known to biology as named species and most recent estimates predict ∼8.7 ± 1.3 million eukaryotic species to exist on the earth and oceans (Mora et al., 2011). Advances in evolutionary and developmental biology, genetics and ecology have begun to provide mechanistic insights into how this astounding diversity comes about. For example, the Modern Synthesis in the early 20th century integrated, among other fields of biology, the fields of population genetics with Mendelian genetics, so as to provide a comprehensive theoretical framework explaining the genetic basis of evolution. Although the discovery of DNA as the universal genetic material by Griffiths in 1928 had already underscored the unity of life, an understanding of proximate mechanisms that govern the translation of a genotype into the phenotype of an animal, could only be achieved relatively recently with the rise of modern, molecular developmental genetics and the cloning of first genes regulating embryonic development in Drosophila melanogaster (Lewis, 1978; Nüsslein-Volhard and Wieschaus, 1980). Meanwhile, scientists like Stephen Jay Gould had already begun to analyze the relationship of developmental biological concepts like heterochrony and neoteny to mechanisms of morphological evolution (Gould, 1977), and it would perhaps not be wrong to say, that such ideas provided the impetus to the nascent field of the so called evolutionary developmental biology (or evo-devo in short).

1.1.1 Discovery of remarkable conservation of developmental control genes

After the Modern Synthesis of the early 19th century, it was argued that positive, directional selection was the predominant evolutionary force and it was thus predicted that genes making up organisms of one species would hardly be conserved across another species (Dobzhansky, 1955; Mayr, 1963). In contrast, the neutral theory of molecular evolution (Kimura, 1983) argues that since most mutations in genes would be deleterious and selected against while mutations on the non-coding portions of the genomes will tend to be selectively neutral, the protein coding regions of the genome would thus be under stabilizing or purifying selection. A corollary to this prediction is that genes would be conserved over large evolutionary time-spans, and hence even across different species (Kimura, 1983; Sommer, 2009). Comparative analysis of genes regulating development in different species was then the natural method to help resolve these conflicting viewpoints. Soon after the discovery of genes regulating development in *Drosophila* (Lewis, 1978; Nüsslein-Volhard and Wieschaus, 1980), similar searches for developmental control genes in other model systems such as *Xenopus laevis* (Carrasco et al., 1984), chicken (Rangini et al., 1989), *Caenorhabditis elegans* (Costa et al., 1988; Kenyon, 1986), mice (McGinnis et al., 1984) and also humans (Levine et al., 1984) revealed an unexpected level of conservation of developmental regulatory genes, the Hox gene cluster being the most exemplary (Carroll, 2001). Not only the individual genes, but even entire developmental pathways have now been found to be conserved across various species, at least in their biochemical functions. Such discoveries have now led to an interesting conundrum at the opposite end of the spectrum, namely how can the observed animal diversity be explained, if most of the genes and signaling pathways are so conserved (e.g. (Pires-daSilva and Sommer, 2003)).

1.1.2 Differential regulation of gene expression as a source of morphological diversity

Multicellular organisms also present a similar problem in that how is cellular diversity achieved when all the cells essentially carry the same genome. It is now evident that only a few selected genes are activated in a particular cell-type at a particular time in a well-regulated manner to enable developmental unfolding of genotype to phenotype. Molecular genetic studies have further elucidated the fine structure of genes including

the exons and introns as well as the regulatory regions such as the promoter, 5'-UTR and 3'-UTR regions and cis-acting enhancer elements that could act over large distances. These regions were found to act in combination with trans-acting components such as transcription factors to regulate gene expression in a precise spatial and temporal manner, giving rise to pattern formation and specification of different cell fates during differentiation (Carroll, 2001). Such gene regulatory cascades and their interactions are now viewed as the so-called Gene Regulatory Networks (Davidson, 2006), which have properties such as modularity, redundancy and robustness (Wagner, 2005), providing great flexibility in their combinatorial deployment to generate diverse cell-types within an organism (Davidson, 2006). Comparison of such gene regulatory networks across different model organisms highlighted a deeper homology underlying the apparently diverse external morphology such as the similarities in the antero-posterior coordinate system and segmentation in fly and vertebrates, limb field patterning in fly, chick and mouse, and conserved role of Pax6 in eye development(Wolpert, 2001). Given the high degree of conservation of genes encoding vavrious developmental regulators, cis-regulatory evolution in the non-coding areas of the genome has been proposed as the primary factor generating morphological diversity (Carroll, 2001), almost to the exclusion of variation in protein coding regions as a factor contributing to morphological diversity. However, there do exist several examples where morphological adaptations have come about through changes in the protein coding genes themselves instead of the regulatory regions (e.g. (Hoekstra et al., 2006; Pellegrino et al., 2011; Perry et al., 2007; Steiner et al., 2007). Further, morphological diversity is just one, conspicuous aspect of organismal diversity, but is also complemented by diversity in organismal physiology, behavior, and life-history traits etc. (e.g (Sommer and Ogawa, 2011)). It is quite likely that evolution of the protein coding regions of the genome would play a greater role in such scenarios. Evolutionary studies on different model systems, within a comparative framework will shed further light on different modes of evolution and their relative contributions in the nature. The nematode model *Pristionchus pacificus* has been established keeping these goals in mind, and will be introduced in the following section.

1.2 Nematodes as models organisms.

The phylum Nematoda is one of the most abundant and taxonomically diverse and species rich phylum (Floyd et al., 2002; Lee, 2002). Despite a fairly simple body plan and relatively few organ systems, they have conquered every imaginable ecological niche available on the earth, from terrestrial to marine and fresh water to the arctic. They are also one of the numerically most abundant animal group (Brusca and Brusca, 2003). It has even been remarked that "if all the matter in the universe except the nematodes were swept away, our world would still be dimly recognizable, and if, as disembodied spirits, we could then investigate it, we should find its mountains, hills, vales, rivers, lakes, and oceans represented by a film of nematodes." (Cobb, 1914).

While most nematodes have a quite un-remarkable, simple and very similar body plan, they nonetheless exhibit a fascinating diversity in many other aspects. For example, nematode species reproduction modes exhibit great variations from gonochorists to hermaphrodites to parthenogenetic species (Denver et al., 2011). They also show many variations in life strategies, with many of them as free-living species, others in a phoretic or necromenic association with typically invertebrate hosts, and still others as parasites of diverse hosts such as plants, insects, mammals and even humans (Kiontke and Sudhaus, 2006). The evolution of these diverse life-history traits has been proposed as a key factor for the success of this animal phylum (Sommer and Ogawa, 2011).

The nematode *C. elegans* was adopted as a model system for developmental genetics by Sydney Brenner, because of its simple body plan, transparent body, simplicity of neuronal circuits, short generation time, ease of culturing in lab on bacteria, and hermaphroditic reproduction (Brenner, 1974). Related species in the same genus, e.g. *C. briggsae*, *C. remanei* and *C. brenneri* are also used as comparative model systems in evolution and development (Gupta and Sternberg, 2003; Stein et al., 2003). To enable macroevolutionary comparisons across larger evolutionary distances, the nematode *P. pacificus* has been developed as a model system for comparative studies in evo-devo with *C. elegans* (Sommer, 2009; Sommer et al., 1996). Further, such studies are now also being complemented with studies on natural variation and microevolution, population genetics and ecology of *P. pacificus*, which together would facilitate an integration of different aspects of studying evolutionary biology (Sommer, 2009).

1.2.1 Evolutionary history of nematodes

Placing the model organisms within a robust phylogenetic framework is essential for interpretation of findings in evolutionary biology, to infer the directions and nature of changes (gains / losses / conservation / variation) across evolutionary time-spans. Initial attempts to place nematodes within the tree of life were ambiguous and were considered to belong to a group of soft-bodied invertebrates called Pseudocoelomata (Fitch, 2005).

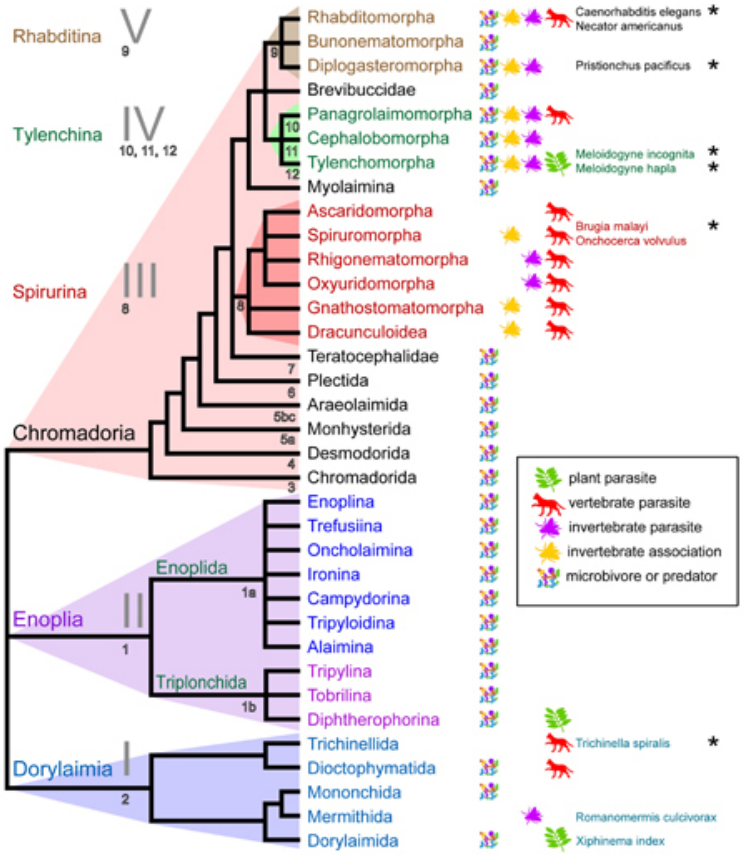


Figure 1.1: Molecular phylogeny of the phylum Nematoda - The molecular phylogeny is based on the analysis of the SSU rRNA gene. The figure has been adapted from Figure 2 of (Blaxter, 2011). Clade names in Roman numerals (I, II, II, IV and V) represent the classical clades as per Blaxter et al. (1998). Some representative species are listed on the right, and the diversity in feeding modes and animal associations are indicated with graphical icons. *C. elegans* and *P. pacificus* both belong to clade V, and are marked by blue and red arrows respectively. Species with sequenced genomes are indicated with asterisks.

Recent phylogenies based on molecular data from ribosomal RNA sequences of

various nematodes have now resolved their position within the superphylum Ecdysozoa, which contains other moulting invertebrates including Arthropoda (Aguinaldo et al., 1997; Blaxter, 2011). Within the phylum, around 25,000 species are known and are organized in a robust molecular phylogenetic framework, consisting of five major clades (see Figure 1.1 above).

Interestingly modes of reproduction have evolved multiple times, so has parasitism, indicating extensive homoplasy (convergent evolution) within this phylum. *P. pacificus* belongs to clade V, family Diplogastridae (Blaxter, 2011) while *C. elegans* belongs to the taxonomic order Rhabditida within the same clade V. The predicted evolutionary distance between these two nematode model species is about 280-430 million years (Dieterich et al., 2008), making them ideal candidates for studies on macroevolution.

1.2.2 The nematode *Pristionchus pacificus* as a model organism to integrate evolution, development and ecology

P. pacificus was originally established as a model system in evo-devo, to analyze evolution of developmental mechanisms in comparison with the related and extensively studied model nematode *C. elegans* (Sommer et al., 1996). It shares many features with *C. elegans*, which make it ideal model organisms for lab studies, such as small size, short generation time, hermaphroditic mode of reproduction and cryopreservation, and a developmental program very similar to *C. elegans* (reviewed in (Hong and Sommer, 2006)). Over the years, tools, techniques and resources such as genetic linkage map, forward and reverse screening methods (Hong and Sommer, 2006), a sequenced genome and well characterized transcriptome and proteome (Borchert et al., 2010; Dieterich et al., 2008) and transgenic methods (Schlager et al., 2009) have been added to the *P. pacificus* toolkit to enable functional and mechanistic studies. It has been further enhanced as a model system for evolutionary biology, ecology and population genetics, by defining its natural history and its ecological niches in the wild (Herrmann et al., 2007). In addition, the isolation of various natural strains from the wild all over the world (Herrmann et al., 2007; Zauner et al., 2007)) as well as outgroups and sister species (Kanzaki et al., 2012) facilitates assessment of natural variation and study of population genetics (Morgan et al., 2012).

Comparative studies of developmental processes with *C. elegans* have revealed many interesting patterns of macroevolutionary alterations and conservation. For example,

investigations of vulva development, one of the best-studied development process in *C. elegans*, formed the starting point for evo-devo studies in *P. pacificus* and revealed that not only the source but also the molecular nature of the inductive signal responsible for vulva formation has changed dramatically over the course of evolution between the two species. The inductive signal in the form of EGFR / Ras signal is provided by the Anchor Cell in the gonad of *C. elegans*, whereas the Wnt pathway is the inductive signal that is relayed from the entire gonad in *P. pacificus* (Tian et al., 2008; Zheng et al., 2005). Recently it was also discovered that the Wnt pathway itself is wired in a fundamentally different manner in *P. pacificus* as compared to *C. elegans*, with the receptor tyrosine kinase *Ppa-lin-18*/Ryk harboring an additional SH3 motif with inhibitory functions (Wang and Sommer, 2011). Similar studies on gonad formation (Rudel et al., 2005, 2008) and sex-determination systems (Pires-daSilva and Sommer, 2004) in the two species have also provided insights into evolution of developmental mechanisms. *C. elegans* and *P. pacificus* also provide an excellent system to study evolution of developmental polyphenisms such as entry into dauer diapause under unfavorable conditions, which was shown to be regulated by a conserved endocrine regulator Dafachronic acid, dependent on conserved roles the nuclear hormone receptor DAF-12, and the forkhead family transcription factor DAF-16 which acts downstream to insulin signaling (Ogawa et al., 2011, 2009). In addition to the dauer polyphenism, *P. pacificus* displays another polyphenism in its mouth, in form of two discrete morphs called eurystomatous and stenostomatous. It has been shown that the regulation of this mouth-form dimorphism is regulated via co-option of some of the molecular genetic modules regulating the dauer polyphenism (Bento et al., 2010).

The nematode *P. pacificus* is found in an necromenic association with several scarab beetles, in which the nematodes live on the host beetle in an arrested dauer stage, and resume development on the micro-organisms that become available when the beetle dies and starts decomposing. This host-specific association has allowed the collection of large number of natural isolates form all over the world (Herrmann et al., 2007; Zauner et al., 2007), and especially from the La Reunion islands (Morgan et al., 2012). The collection of natural isolates provides a rich resource to assay the natural variation and microevolution in *P. pacificus* both at genetic and phenotypic level e.g differences in chemoattraction to the pheromones from host beetles (Hong et al., 2008), sensitivity to species-specific dauer pheromones (Mayer and Sommer, 2011), and differences in

vulva development (Ph.D Thesis (Kienle, 2012) and submitted manuscript Kienle et al 2012) as well as population genetic approaches to understand its evolutionary history (Morgan et al., 2012). The ecology and natural history of *P. pacificus* is further being studied in more detail (Herrmann et al., 2007; Weller et al., 2010) e.g. olfactory cues for attraction to host beetle (Hong et al., 2008), interaction with bacteria (Rae et al., 2010, 2008) and nictation behavior (Brown et al., 2011).

1.3 Genomic approaches to study evolution in *P. pacificus*

Over the last decade, the whole genome sequences from various model organisms have become available. These include *C. elegans*, *D. melanogaster*, chicken, mice and rats(Adams et al. (2000); Consortium (2004); Gibbs et al. (2004); elegans Sequencing Consortium. (1998); Waterston et al. (2002) humans (McPherson et al., 2001; Venter et al., 2001) and the plant model Arabidopsis thaliana (Initiative., 2000). Availability of these whole genome sequences has opened up a new vista in evolutionary studies (reviewed in (Cañestro et al., 2007)). The findings from molecular genetic studies of development that many key regulatory genes are conserved across species boundaries can now be tested and examined at the whole genome level. This allows not only comparisons of components of genetic toolkits but also studies of how the genome structure itself evolves (Cañestro et al., 2007). A sequenced genome is also a basic requirement for mechanistic studies in any model system. The whole genome of *P. pacificus* has now been sequenced (Dieterich et al., 2008) and the transcriptome (the complement of all expressed genes) and the proteome have also been qualitatively characterized (Borchert et al., 2010). The genome of *P. pacificus* exhibits some interesting features that provide insight into its life history and ecology as well genome-evolution. At about 169 mega-bases (Mb), the genome of *P. pacificus* is larger than the 100Mb genome of *C. elegans*, which also translates to a higher number of genes in *P. pacificus* (predicted gene counts between 25,000 to 29,000) as compared to *C. elegans* (20,000 genes). Interestingly, many gene families have undergone contraction or expansion in *P. pacificus* lineage ((Dieterich et al., 2008), see Figure 1.2), with potential bearing on its ecological adaptations, including those needed to survive the host beetles.

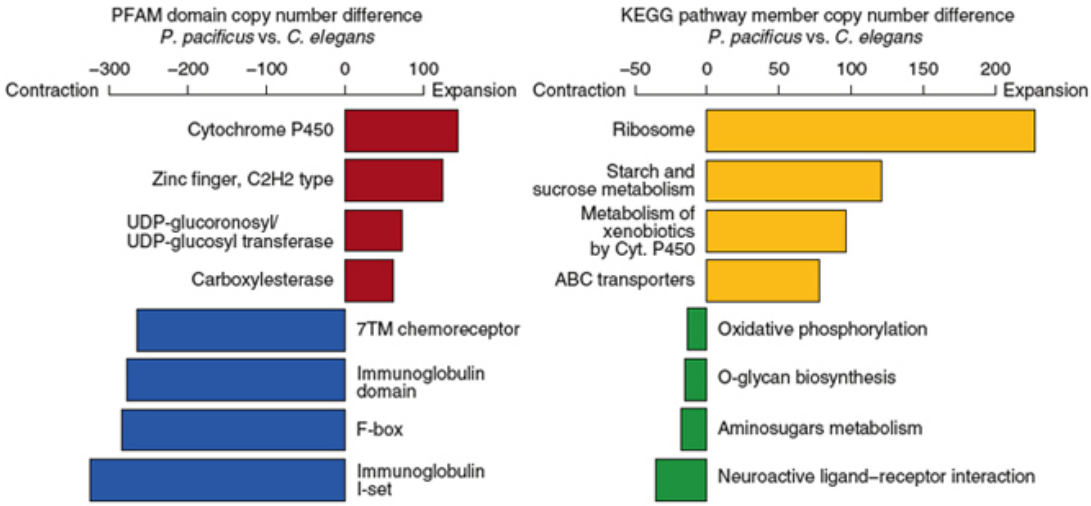


Figure 1.2: Gene families that have undergone expansion or contraction in *P. pacificus* genome relative to the *C. elegans* genome. - Genes encoding for proteins with roles in members of xenobiotic metabolism and detoxification machinery are expanded in *P. pacificus*. Figure taken from (Dieterich et al., 2008)

For example, the greatest increase in copy number in *P. pacificus* with respect to *C. elegans* is found in the gene families involved in metabolism of xenobiotics (Cytochrome P450, UDP-glycosyltransferases and Carboxylesterases), which might be needed to defend against some beetle associated toxic compounds or other toxins present in its ecological niche. However, this observed contraction and expansion of gene families does not by itself explain the increase in gene number in *P. pacificus*. A search of orthologs for all predicted *P. pacificus* genes revealed that not all genes in *P. pacificus* have a homolog in *C. elegans* and vice-versa (see Figure 1.3 below). Many of these genes are most likely to be acquired via horizontal gene transfer, a hitherto under-appreciated aspect of gene content changes in eukaryotic genomes Gogarten (2003) , but observed to be more and more common Keeling (2009), especially in nematode genomes (Dieterich and Sommer, 2009; Haegeman et al., 2011; Mayer et al., 2011; Schuster and Sommer, 2012). The best-characterized example in *P. pacificus* is the acquisition of genes encoding for cellulase enzymes (Dieterich et al., 2008). Even though *P. pacificus* is not a plant parasitic nematode, its genome was found to have 6 cellulase encoding genes, which seem to have been acquired via horizontal gene transfer from slime-moulds (Dieterich et al., 2008), have been stably integrated into host genomes and host biology

and are also present in 8 other species within the genus *Pristionchus* (Mayer et al., 2011). Most of these genes are expressed, display cellulase activity in-vitro, and some of these are also developmentally regulated (Schuster and Sommer, 2012), although their precise function in *P. pacificus* biology is still under investigation.

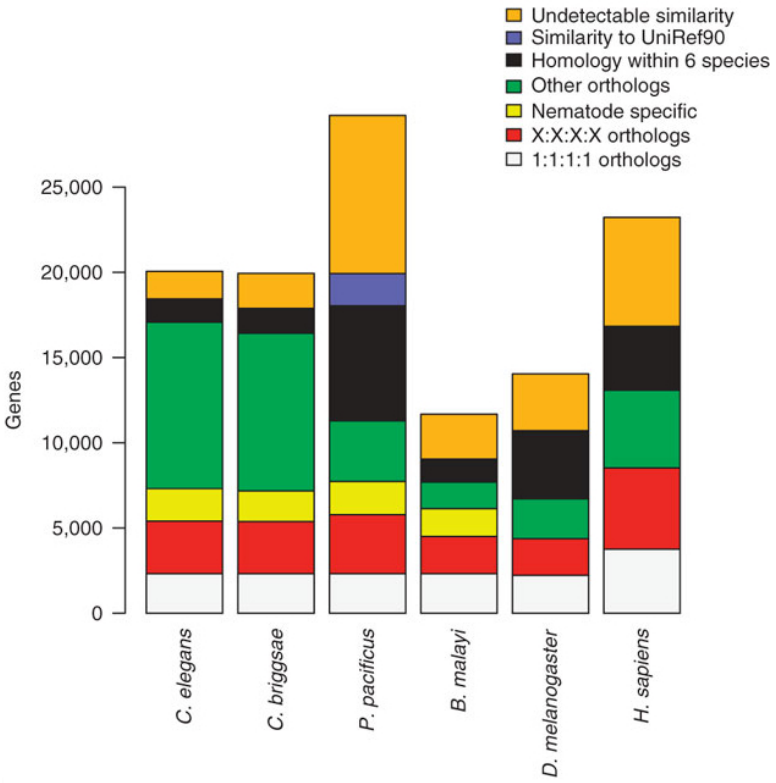


Figure 1.3: *P. pacificus* has a large number of "Pioneer genes" - The homology relations between proteomes of various sequenced organisms show that apart from proteins with one-to-one or one-to-many orthology across different species, there exist many lineage specific genes, e.g. those found only in nematodes. The fraction of genes with "Undetectable similarity" to any of the known genes comprises the so called pioneer genes and makes up about one-third of the predicted proteome of *P. pacificus*. Figure taken from (Dieterich et al., 2008)

Another very interesting class of genes is the so called "Pioneer genes" class, which consist of genes whose protein products do not show any detectable similarity to the known protein universe (Borchert et al., 2010; Dieterich et al., 2008), see Figure 1.3). Expression for many of these genes has been confirmed by cDNA / EST and proteomic analysis (Borchert et al., 2010) suggesting some as yet unknown functions for these genes in *P. pacificus* biology. Computational approaches analyzing codon-usage bias

suggest that many of these genes might have been acquired via HGT from insects possibly through retrotransposon based mechanisms and have subsequently diverged (Rödelsperger and Sommer, 2011).

In summary, the *P. pacificus* genome provides the opportunity for genomic comparisons with *C. elegans* to reveal many interesting features in genome evolution such as changes in gene contents including presence of novel genes, at least some of which must have been selected for ecological adaptations. Further studies will be needed to elucidate their functions.

1.3.1 Transcriptomic approaches to study evolution

As discussed in previous sections, much of evolutionary diversity seems to have been produced by differential regulation of gene expression. Given the availability of nematode genomes, it would then be interesting to ask how similar or diverged are the set of genes expressed under similar conditions or in homologous developmental processes. The presence of pioneer genes in the *P. pacificus* genome also raises the question of what their potential function might be. Expression of a certain gene in a particular developmental stage or under specific environmental conditions might provide first clues to their function. Such questions can be tackled by a systematic analysis of the set of genes expressed under different conditions. Further, comparative transcriptomic data will allow constructing and comparing gene regulatory networks, to understand the evolution of phenotypes at a system level. Use of transcriptomic approaches such as microarrays (Schena et al., 1995) and recently RNA-Seq (Wang et al., 2009) has been most widespread in tackling functional mechanistic questions e.g. in identifying candidate genes involved in a particular biological process. Additionally, these methods have also been used in evolutionary studies to discover patterns of evolution of the transcriptome (reviewed in (Ranz and Machado, 2006)), and will be increasingly useful for answering such questions (Hashimshony and Yanai, 2010). For example, comparison of developmental transcriptome during the metamorphosis of four *D. melanogaster* strains and two related species *D. simulans* and *D. yakuba* found that inter-lineage differences in gene activity mirrored the phylogenetic pattern of the species, and the genes which varied most could be identified as the candidate loci for positive selection during transcriptome evolution (Rifkin et al., 2003). Other studies with similar approaches

within *Drosophila* subgroup have studied how sex-dependent and male-biased gene expression evolves across species (Meiklejohn et al., 2003; Ranz et al., 2003), evolution of gene expression in olfactory system in correlation with dietary preferences (Kopp et al., 2008). Other studies in various models have investigated the extent of transcriptome conservation in homologous developmental processes across related species and found widely varying levels of conservation (Parikh et al., 2010; Yanai and Hunter, 2009; Yanai et al., 2011; Zhang et al., 2007; Zheng-Bradley et al., 2010), prompting a search for mechanisms that govern the evolution of gene expression (Gilad et al., 2006). Other novel applications of transcriptomics in evolutionary biology have been in inferring the phylotypic stages in zebrafish (Domazet-Lošo and Tautz, 2010), *Drosophila* (Kalinka et al., 2010), *C. elegans* (Levin et al., 2012) and different vertebrate species (Irie and Kuratani, 2011) and have provided insights into a molecular basis of the hourglass model originally proposed by Karl von Baer in 1928 (reviewed in (Prud'homme and Gompel, 2010)). In summary, transcriptome-wide studies of comparable life-stages and conditions in related species will help answering question such as (1) whether gene expression always correlates with adaptation or can it evolve by drift (gratuitous expression, (Gerhart and Kirschner, 1997)), (2) do orthologous genes show similar expression patterns across species, and (3) how are new genes integrated into pre-existing gene regulatory networks (Hashimshony and Yanai, 2010). Gene expression programs are organized in hierarchical gene regulatory networks (GRNs in short, (Davidson, 2006)), with key regulators such as signaling components at the top, which affect gene expression ultimately via action of transcription factors, which then control the expression of most downstream genes required for expression of a phenotype. It has been argued that changes to the molecular components at different hierarchical levels in a GRN will have different evolutionary constraints and generate different outcomes at a phenotypic level, with the most downstream components being relatively free to change and even integrate new genes into the networks, to generate diversity of phenotypes (Davidson and Erwin, 2006). Transcriptome measurements record a snapshot of this bottom-most layer of the GRN, and analyzing these data in a comparative context can be used to test the various hypotheses about evolution of GRNs. Furthermore, transcriptome analysis of a given biological process can in itself be used to gain mechanistic insights and identify potential gene candidates mediating the process. It must be noted that such studies will also have to be complemented by classical mechanistic studies of gene function to

build a functionally relevant GRN. In the following sections, aspects of *P. pacificus* and *C. elegans* biology that are amenable to such genome wide gene-expression studies are discussed in more detail.

1.4 The dauer polyphenism in *P. pacificus* and *C. elegans*

Phenotypic plasticity denotes the ability of animals to acquire different phenotypes under different environmental conditions. The response might be continuous (reaction norms) or discrete (polyphenism). Examples of morphological polyphenisms include season dependent wing-morphs in butterflies Beldade et al. (2002) , or sexually dimorphic horns in beetles Moczek and Nijhout (2004) . Phenotypic plasticity can also entail formation of developmentally discrete life stages or life histories, e.g caste determination in hymenopterans Abouheif and Wray (2002) , and temperature dependent sex-determination in reptiles (Ramsey and Crews, 2007). Such plasticity in the phenotype not only allows the organism to survive a range of environmental conditions, but can also provide a background for new evolutionary innovations (West-Eberhard, 2003). An interesting case of phenotypic plasticity observed in many free living nematodes is the formation of the so called dauer larvae, in response to unfavorable conditions like high population density, scarcity of food or high temperature (Sommer and Ogawa, 2011). Under optimal conditions in laboratory, post-embryonic development of *C. elegans* as well as *P. pacificus* proceeds through four larval molts (L1 to L4 and J1 to J4 respectively) finally leading to reproductively mature adult hermaphrodites in about three to four days at 20 °C, thus completing the so called direct life-cycle (Figure 1.4).

However, if the nematodes encounter harsh conditions at the L2 or J2 larval stage, they can bypass development to the L3 / J3 stage and instead divert their development into the dauer diapause (Figure 1.4), which basically is a developmentally quiescent, stress resistant and effectively non-aging life-stage (Klass and Hirsh, 1976). The morphological changes include radial shrinking, formation of a tougher and impermeable cuticle, sealing of buccal and anal orifices Cassada and Russell (1975) , while behavioral changes include cessation of feeding, lethargy as well host-seeking nictation (e.g. (Brown et al., 2011)). Physiological changes include a shift to non-aerobic metabolism, deriving energy by metabolizing fat stores accumulated during the previous larval stage

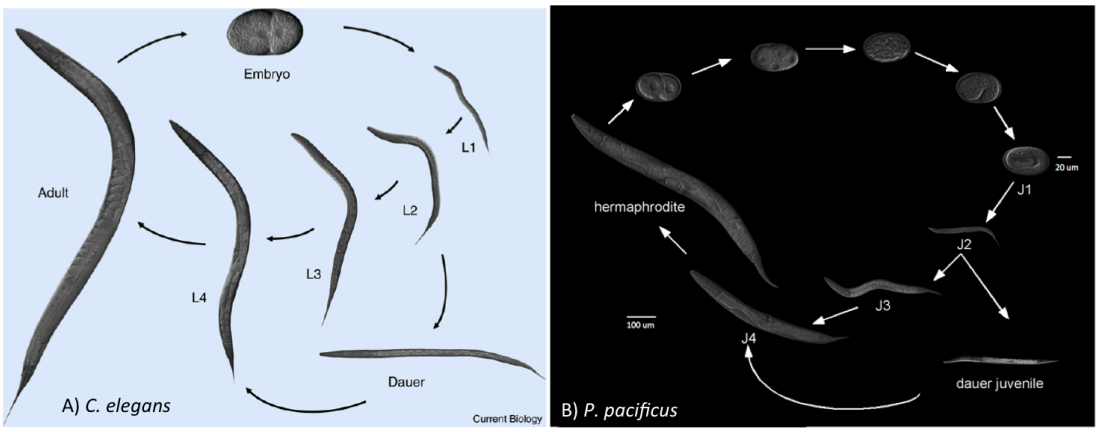


Figure 1.4: Life stages of (A) *C. elegans* and (B) *P. pacificus* during the direct and indirect cycle. - Under unfavorable environmental conditions, both nematodes can switch to an alternative, stress-resistant and non-aging "dauer" stage, and can resume direct development on return of favorable conditions. Figure (A) adapted from Figure 1 of (Sommer and Ogawa, 2011); Figure (B) courtesy of Metta Riebesell

(O'Riordan and Burnell, 1990). Upon the return of favorable conditions like availability of food, dauers then exit the diapause and go on to develop into the L4 / J4 stage worms and subsequently into reproducing adults. Interestingly, the time spent in the dauer stage does not affect the post-dauer lifespan in the L4 / J4 stages and beyond, implying that the dauer larva is essentially a non-aging stage (Klass and Hirsh, 1976). This diapause not only allows surviving harsh conditions but also has behavioral adaptations like nictation that allow dispersal by seeking and attaching to invertebrate hosts that might transport them to a new location with better conditions for survival. The dauer larval stage in free-living nematodes like *C. elegans* and *P. pacificus* is considered developmentally homologous to the infective juvenile stage observed in of many parasitic nematodes, and a possible precursor to the evolutionary innovation of a parasitic life-strategy (Dieterich and Sommer, 2009).

Through various screens for "Dauer formation constitutive" (Daf-c) and "Dauer formation defective" (or Daf-d) phenotypes in *C. elegans*, the molecular genetic basis for this environment dependent developmental switch to dauer stage has been worked out in great detail (Fielenbach and Antebi, 2008; Hu, 2007). Briefly, different environmental cues are sensed by specific neurons, transduced to the inside of the animal and converted into endocrine signals which then regulate organism-wide physiological changes (Figure 1.5). The chemosensory neurons such as ASI, ADF and ASG transduce

the environmental signals, such as pheromone concentration, via a trans-membrane guanylyl cyclase DAF-11, which regulates the levels of cGMP. During normal, reproductive development, cGMP levels result in production of TGF-β ligand DAF-7 in ASI neurons, and insulin like peptides (ILPs) in amphid neurons and other tissues (Fielenbach and Antebi, 2008), which ultimately phosphorylate the respective transcription factors such as SMADs DAF-8 and DAF-14 of the TGF-β pathway, and the forkhead family transcription factor DAF-16 of the Insulin signaling pathway.

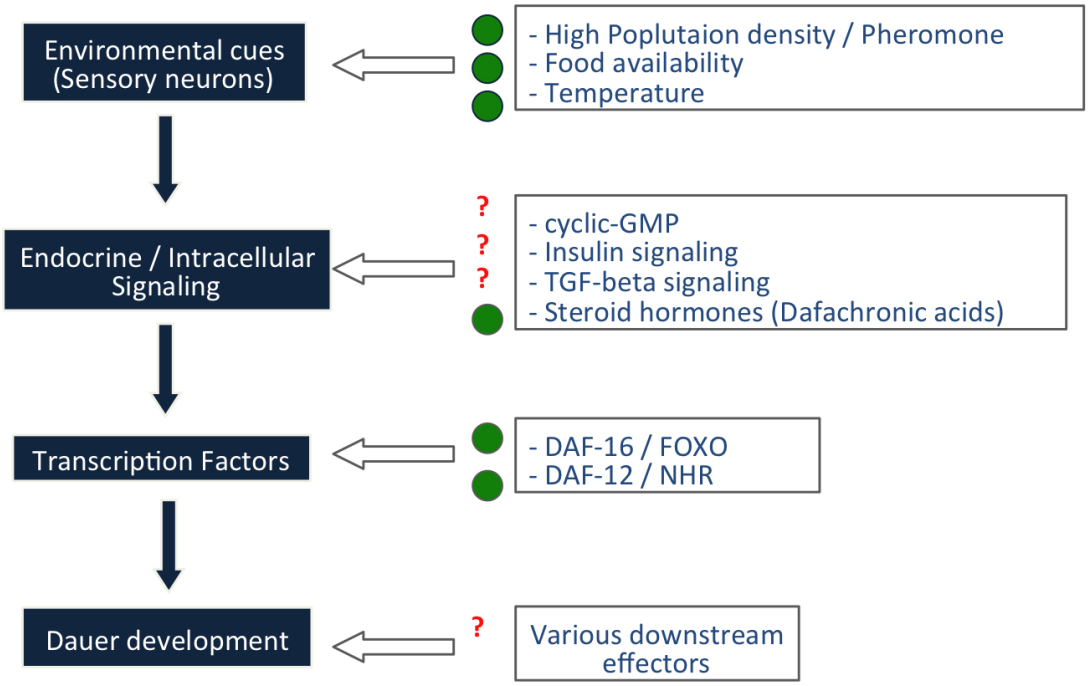


Figure 1.5: The regulatory system governing dauer formation in *C. elegans*, and their comparison to known components in *P. pacificus*. - Conserved elements are indicated by green dots, while unknown elements are marked by red question marks.

The phosphorylation of these key transcription factor prevents their nuclear translocation and hence cause suppression of target genes. Under dauer inducing conditions, the TGF-β and Insulin pathway are downregulated, thereby activating the downstream transcription factors and their respective target genes, which include stress responsive genes such as heat-shock factors, chaperones, super-oxide dismutase and other genes required for dauer development. The TGF-β pathway also regulates the expression of hormone synthesis enzymes DAF-9/CYP450, which produce the steroid hormone

Dafachronic acids in a cholesterol dependent manner. These hormones act through the nuclear hormone receptor DAF-12 (Motola et al., 2006)

Comparative studies in other model systems including parasitic nematodes have found limited conservation of pathways such as TGF-β pathway in dauer or the homologous infective larvae (Viney et al., 2005) . However, the action of Dafachronic acids has been shown to be conserved across *C. elegans*, *P. pacificus* and the mammalian parasitic nematode *Strongyloides papillosus* (Ogawa et al., 2009). The role of the NHR/DAF-12 and the FKH/DAF-16 has also shown to be conserved in *C. elegans* and *P. pacificus* (Ogawa et al., 2011, 2009), but the role of TGF-β pathway in *P. pacificus* has not been studied so far. Moreover, it is not known whether the downstream targets of the conserved transcriptional regulators DAF-16 and DAF-12 are also conserved between *C. elegans* and *P. pacificus* during dauer formation.

1.5 Innate immunity and response to pathogens in *P. pacificus* and *C. elegans*

Host pathogen interactions are considered to be important agents during evolution, where the virulence mechanisms of the pathogens and the defense mechanism of the host are constantly co-evolving (Haldane, 1949; Thompson, 1994; Woolhouse et al., 2002). Although invertebrates like *D. melanogaster* and *C. elegans* lack an adaptive immune system like the vertebrates, they have proven to represent excellent models to study the genetic basis of innate immunity, the first line of defense against pathogens (Lemaitre et al., 1996; Mahajan-Miklos et al., 1999), which is also conserved in vertebrates , (Dong et al., 2002; Medzhitov et al., 1997; Poltorak et al., 1998). Studies in *D. melanogaster* have been key in identifying the role of Toll like receptors in pathogen recognition and activation of innate immune response (Lemaitre et al., 1996), and subsequently in mammals (Medzhitov et al., 1997; Poltorak et al., 1998). Interestingly, *C. elegans* lacks the homologs of key components of the Toll signaling pathway (reviewed in Irazoqui et al. (2010)). However it does mount a pathogen-specific, inducible defense response (Mallo et al., 2002) and has been used successfully to uncover the molecular nature of virulence factors of many pathogens affecting humans Irazoqui et al. (2010). Molecular genetic studies in *C. elegans* to elucidate the genetic basis of its innate immune response have uncovered roles for various signaling pathways, including the p38

MAP Kinase pathway, the TGF-β pathway, the DAF-2/DAF-16 Insulin pathway, ERK kinase pathway, as well as a generalized stress pathway that most likely acts be recognizing the damage caused by a pathogen Irazoqui et al. (2010). These upstream regulators in turn activate the expression of various downstream effector genes with potential anti-microbial and stress-resistance activities (Mallo et al., 2002; Pujol et al., 2008; Troemel et al., 2006).

Nematodes like *C. elegans* and *P. pacificus* which feed on microbes in the wild, are likely to encounter various microbes some of which can be utilized as a food-source while others could be pathogenic and need to be defended against. Hence, the ecology and life-history of the animals in the wild is expected to influence the evolution of their innate immune systems, but the evolution of the regulators and effectors of innate immune system has not been studied in much detail. Recent studies have begun to characterize the bacteria and fungi with which *C. elegans* interacts in the wild Coolon et al. (2009); Couillault et al. (2004) and tried to identify its natural pathogens. In our lab, studies of nematode-bacteria interaction in the nematode *P. pacificus* have identified various microbes naturally associated with *P. pacificus* in the wild (Rae et al., 2008), and isolated various bacterial strains that are pathogenic to *P. pacificus*, or *C. elegans* or both (Rae et al., 2010). A screen for hyper-susceptible mutants in *P. pacificus* has identified a role for muscle development and defecation-cycle regulating genes *Ppa-unc-22* and *Ppa-unc-13* in innate immunity, and their role has also been shown to be conserved in *C. elegans* (Rae et al., 2012). However, the role of other signaling pathways known to regulate innate immunity in *C. elegans* has not been elucidated in *P. pacificus* so far. Since inducible expression of genes encoding for potential anti-microbial proteins and other defense molecules is a key aspect of pathogen response in nematodes like *C. elegans*, it would be interesting to characterize and compare the transcriptomes induced in *C. elegans* and *P. pacificus* in response to different pathogens.

1.6 The role of germline and somatic gonad in the regulation of longevity and innate-immunity in *P. pacificus* and *C. elegans*

It is a common observation that different animals differ in their lifespans. Aging was generally considered to be an unavoidable, entropic process, a natural degradation of

various cells, tissues and organ systems of an organism. However, the discovery of long-lived mutants in *C. elegans* which could extend life-span by up to two-fold (Friedman and Johnson, 1988a,b; Kenyon et al., 1993; Klass, 1983; Morris et al., 1996) have dramatically altered this point of view, and aging is now understood as a genetically regulated process, which is also impacted upon by various environmental factors such as nutrition and reproduction (Kenyon, 2010). Molecular characterization of mutants that increased *C. elegans* lifespan revealed a key role for endocrine signaling via the Insulin/Insulin Growth Factor -1 (IGF-1) signaling pathway. Loss of function mutations either in daf-2, the *C. elegans* homolog of the Insulin receptor, or, the downstream PI3-kinase age-1 lead to extension in lifespan. The increased lifespan and the associated gene expression changes require the activity of DAF-16, a Forkhead/FOXO transcription factor that is a key effector of insulin signaling. These long-lived animals have also been observed to be more resistant to a variety of stressors such as heat stress, oxidative stress and to microbial pathogens (Kenyon, 2010). It is interesting to note that the DAF-2 / DAF-16 axis also plays a key role in regulating the stress resistant dauer life-stage in *C. elegans* and *P. pacificus* (Ogawa et al., 2011). Other clues that suggest that aging is a developmentally regulated process come from experiments manipulating the reproductive system of various animals, such that removal of reproductive tissue, especially the germline, results in an extension of lifespan in worms, flies, fish and humans (Cargill et al., 2003; Flatt et al., 2008; Hamilton and Mestler, 1969; Hsin and Kenyon, 1999; Kime and Larsen, 1987). In *C. elegans*, ablation of the germline precursor cells Z2 and Z3 in the L1 larval stage produces sterile animals which have the somatic gonad but no germline, and these animals have have a more than 50% increase in lifespan than the unablated controls (Hsin and Kenyon, 1999). However, this increase in lifespan is not simply due to a trade-off between reproduction and longevity. The animals that are rendered sterile by removal of somatic gonad by ablating gonadal precursor cells Z1 and Z4 in the L1 stage, lack the germline as well as the somatic gonad and cannot reproduce, but have a normal lifespan and are not long-lived like the Z2 and Z3 ablated animals without the germline (Hsin and Kenyon, 1999). It has thus been proposed that the germline produces a life-span shortening signal that is countered by a lifespan extending signal coming from the gonad. Interestingly, this increase in lifespan in germline-ablated animals also requires the activity of the FOXO transcription factor DAF-16, but works in parallel to the Insulin receptor DAF-2, such

that germline ablation of daf-2 mutants increases lifespan even beyond that of unablated daf-2 mutants. The role of germline in lifespan regulation, its gonad-dependence and requirement of DAF-16/FOXO has subsequently been shown to be conserved in other model systems such as D. melanogaster (Flatt et al., 2008). In *C. elegans*, the gonad dependent longevity signal is most likely conveyed through a lipophilic hormone because mutations in either the cytochrome DAF-9, which produces the steroid hormone Dafachronic acid, or the corresponding nuclear hormone receptor DAF-12 leads to reduction of life-span to normal levels in germline ablated animals. Other genetic factors regulating this gonad-dependent longevity signal include an intestinally expressed ankyrin repeat protein KRI-1 Berman and Kenyon (2006) and a translational elongation factor TCER-1 (Ghazi et al., 2009). Apart from this handful of genes, the identities of other genes regulated by ablation of germline, or the genes regulating the gonad-dependent longevity signal, are not known. Consistent with observations in long-lived mutants, germline ablation in *C. elegans* also leads to an increased resistance against bacterial pathogens (Alper et al., 2010; TeKippe and Aballay, 2010). However, it is not entirely clear whether the increased survival on pathogen is just a consequence of longer lifespan or a specific immune response is activated in response to pathogens in germline-ablated animals. The gonad and germline tissues in *P. pacificus* are also derived from the homologous precursor cells Z1, Z2, Z3 and Z4, and their ablation also has the same effect on *P. pacificus* lifespan as in *C. elegans* (Hsin and Kenyon, 1999; Patel et al., 2002). As in *C. elegans*, it is not known in *P. pacificus* what genes are regulated in response to germline ablation that could potentially contribute to its increased lifespan. Further, the effect of germline ablation on pathogen resistance in *P. pacificus* is not known, neither is it known whether exposure to pathogens leads to activation of additional genes beyond those required for lifespan activation.

1.7 Evolution of genome structure: Analysis of trans-splicing and operon-like organization of genes in *P. pacificus* genome

In *C. elegans* and other nematodes, most mRNA transcripts are trans-spliced to 22-nucleotide long spice-leaders (SL) at their 5 ends. Trans-splicing is a phenomenon where an identical short leader sequence called the splice leader (SL) is spliced onto 5

ends of various mRNA transcripts that are derived from independent loci (Blumenthal, 2005). It was initially discovered in trypanosomes and was later also found to occur in euglena, flatworms, hydra, some primitive chordates and various nematodes including *C. elegans* (Blumenthal, 2005). Although the function of trans-splicing is not well understood, the process has been best studied in *C. elegans*. About 50percent of the genes in *C. elegans* genome are trans-spliced at their 5 end to the splice leader SL1, which is derived from a 100 base pair small nuclear ribonucleoprotein particle (snRNP). Around 30percent of all genes are not trans-spliced at all but the remaining genes are trans-spliced to a related family of SL2-like spliced-leaders. The SL1 splice leader is found across all nematode species surveyed so far and its trans-splicing is believed to aid in related biological functions such as the stabilization of 5 ends of mRNA via cap donation, processing of the 5 untranslated region (UTR) of pre-mRNAs, and translation optimization via specific interactions of the SL sequence and trimethylguanosine (TMG) capped transcripts with the translation machinery (Guiliano and Blaxter, 2006). The SL2 variants appear to be a unique feature of clade V nematodes (see Figure 1.1 for phylogeny), and function in resolving individual transcripts from poly-cistronic pre-messenger RNA. These poly-cistronic transcripts arise from so called "operons" in *C. elegans*, which consist of a cluster of two to eight genes with the same orientation, inter-genic distances typically hundreds of base pair, and a common promoter region upstream to the most 5 gene. The first gene in the operon is predominantly trans-spliced to SL1 splice leader, while the downstream genes are typically trans-spliced to SL2 ro SL2-like splice leaders (Blumenthal, 2005). Microarray based genome-wide analysis has identified at least 20percent genes to be in about 1,000 operons with two to eight genes per operon (Blumenthal et al., 2002). The reason for this prokaryotic operon-like gene organization is not clear, although the operonic genes are enriched in functions related to core transcriptional and translational machinery, spliceosomes, ribosomes and ubiquitin conjugating enzymes in the proteasome, well as various mitochondrial functions (Blumenthal and Gleason, 2003). Some of the operons do encode for genes with related functions (see (Blumenthal and Gleason, 2003) for examples), but for the majority of these operons, functional relation between constituent genes is not very obvious. A recent meta-analysis of expression data from *C. elegans* and *Ciona intestinalis* operons has led to the conclusion that operons are enriched for highly expressed genes and that their expression is preferentially upregulated upon recovery

from growth-arrested states such as starved L1 larvae in *C. elegans* (Zaslaver et al., 2011).

Organization of genes into operonic clusters with just one common promoter for multiple genes imposes an important constraint on genome evolution. Genes within an operon are less likely to move out to another location in the genome because they would then most likely be lacking in a functional upstream promoter. Search for conserved operons across various nematode species has found some evidence for micro-syntenies between distantly related nematodes including *C. elegans* and *P. pacificus* (Guiliano and Blaxter, 2006), while other studies find only a limited conservation (Lee and Sommer, 2003). Since these examples are based only a handful of case studies of individual operons, it is not clear what would be the conservation pattern on a genome-wide scale. Further, the trans-splicing landscape and potentially operonic clusters in *P. pacificus* have not been analyzed systematically at a whole genome level.

1.8 Tools for gene expression profiling: Microarrays

Gene expression microarrays measure expression levels of thousands of genes simultaneously by using the same principle as Northern blots - namely complementary base pairing to a gene specific probe to the target transcript. The basic advantage as compared to Northern blots is the high-throughput scale, achieved by immobilizing thousands of gene specific probes on the same substrate such as a glass-slide, which hybridize to respective target molecules in a parallel fashion. The mRNA targets to be assayed are usually fluorescently labeled and intensity of each spot corresponding to a probe serves as a quantitative read-out for the corresponding target (schematic Figure 1.6). Gene expression microarrays for long established model systems like *C. elegans* are available from many commercial vendors, however no such resources existed for *P. pacificus*. The first task during my Ph.D. research was to establish a custom microarray platform for *P. pacificus* and subsequently use it tackle functional as well evolutionary questions in *P. pacificus* biology. The design process for the custom microarrays is briefly described below.

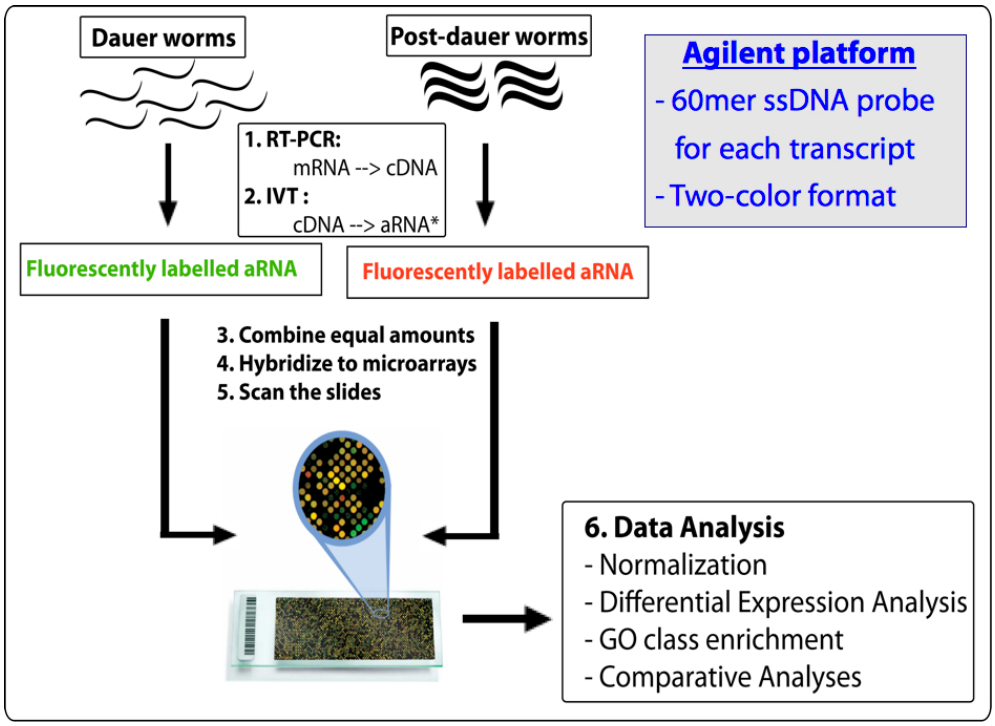


Figure 1.6: Schematic workflow of a typical microarray experiment -

1.8.1 Design of custom microarrays or *P. pacificus*

Based on the most recent transcriptome annotation (∼ 23,000 gene predictions, predicted with external evidence from 454 EST alignments), we designed one or more 60-mer oligo probes that would bind in a sequence-specific manner only to the gene of interest but not to other genes. A systematic search for such probes for all the 23,000 genes on *P. pacificus* genome was carried out using the software OligoWiz (Wernersson et al., 2007). We designed in total about ∼ 93,000 probe sequences for these 23,000 genes. The sequences were uploaded to the eArray webservice provided by Agilent Technologies, who then printed the corresponding sequences on glass-slide custom microarrays. The design of probe sequences, their location on the array and their corresponding target sequences has been submitted to NCBI GEO database as the platform accession number GPL14372.

1.9 Aims of the thesis

Studies on gene expression at a genome-wide scale provide functional insights into a particular biological process and identify candidate genes involved in an particular biological process. Further, analyzing gene expression data across comparable biological processes across species provides insights into functional constraints and extent of conservation of gene regulatory programs. My first task during my thesis was to design, develop and validate a microarray based gene expression platform, before any such studies could be undertaken in *P. pacificus*. I have subsequently used these microarrays in conjunction with commercial microarrays for *C. elegans*, to investigate different biological processes in *C. elegans* and *P. pacificus* from a functional and evolutionary point of view. The biological processes and questions that have been answered in this thesis are briefly described below.

1.9.1 Phenotypic plasticity and evolutionary novelty: comparing dauer transcriptomes

Dauer formation in *C. elegans* and *P. pacificus* is regulated by similar environmental conditions, including hormones like Dafachronic acids, and requires the activity of the FOXO/DAF-16 as well as the NHR/DAF-12 transcription factors in both the species. The aim of this project was to compare the transcriptomes of the dauers from the two species to identify the extent and nature of conservation in the downstream components regulated by these conserved factors.

1.9.2 Evolution of innate immunity: genetic basis of differential susceptibility of *P. pacificus* and *C. elegans* to various pathogens

Studies of nematode-bacteria interactions in the wild have identified various bacterial strains that show differences in their pathogenicity to *P. pacificus* and *C. elegans* Rae et al. (2010, 2008). The molecular basis of these differences in innate immunity across the two species is not known. In collaboration with Dr. Robbie Rae and Igor Iatsenko, I generated and compared the gene expression profiles of *C. elegans* and *P. pacificus* after exposing them to two gram-negative and two gram-positive bacteria. The aim was to characterize the nature and extent of conservation of innate immune responses across the two nematode species, which inhabit distinct microenvironments in the wild.

1.9.3 Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in *P. pacificus*

Ablation of germline in many animals including the model organisms *C. elegans* and *D. melanogaster* has been shown to increase the lifespan (Flatt et al., 2008; Hsin and Kenyon, 1999) as well as pathogen resistance in *C. elegans* (Alper et al., 2010; TeKippe and Aballay, 2010). It is not completely clear whether the increase in pathogen resistance is just a correlated outcome of increased lifespan or an independent phenomenon. Using the germline-ablated animals exposed to non-pathogenic E. coli and pathogenic S. marcescens by Dr. Robbie Rae, I generated and analyzed their gene expression profile to identify molecular processes responsible for extended lifespan and survival on pathogen.

1.9.4 Functional genomic investigation of pathways regulated by the somatic gonad thse contribute towards the longevity of germline-deficient *C. elegans*

The increase in lifespan of *C. elegans* upon ablation of germline tissue requires the presence of an intact somatic gonad, suggesting a role for the somatic gonad in regulating the longevity signal (Hsin and Kenyon, 1999). With Dr. Robbie Rae, I have generated and compared gene expression profiles of germline ablated (long-lived) animals versus animals with germline and gonad both ablated (short-lived). This analysis identified about a thousand candidate genes mediating the regulation of lifespan in a gonad-dependent manner. We have validated the function of some of these genes using a RNAi mediated knock-down of gene function.

1.9.5 Functional genomic investigation of novel *C. elegans* mutants *Cel-btr-1* and *Cel-btr-2*, which display enhanced resistance to the bacteria pathogen *Bacillus thuringiensis*

Bacillus thuringiensis strain DB27 is a highly pathogenic strain that can kill *C. elegans* within one day while *P. pacificus* shows no susceptibility to this pathogen (Rae et al., 2010). To identify the molecular basis of this pathogenicity in *C. elegans*, Igor Iatsenko screened for mutants that are resistant to this pathogen and identified two novel *C. elegans* genes *btr-1* and *btr-2*. To complement the functional investigations of these genes, I carried out gene expression profiling for these mutants in comparison to wild-type worms, to identify all the processes that are affected by mutations in these genes.

1.9.6 Functional genomic investigation of enhanced pathogen resistance of *C. elegans* when pre-exposed to a related but non-pathogenic bacteria

Diet can influence the innate immunity and pathogen response, especially in bacterivorous nematodes such as *C. elegans*. Igor Iatsenko has observed that feeding *C. elegans* on a non-pathogenic, naturally isolated strain of *Bacillus subtilis* increases the survival of *C. elegans* when subsequently exposed to the pathogenic *B. thuringiensis* DB27. I have generated and analyzed the differences in expression profile generated by difference in diet and identified candidate genes responsible for enhanced pathogen resistance. We have functionally validated some of these candidates

1.9.7 Characterization of the trans-splicing landscape, and operon like gene-clusters in *P. pacificus* genome using next generation sequencing

I have taken a RNA-seq based approach to characterize all potential operons in *P. pacificus* genomes, and then compare their organization and micro-synteny against those in *C. elegans* genome. By analyzing whether any of the operons are conserved, or if similar biological functions are enriched among the operonic genes across the two species, we might get insights into their potential function. Also, a comparison of genomic organization across the two species will provide insights into evolution of genome structure.

Together, the work done in this thesis has established a microarray based transcriptome-profiling resource for the nematode model *P. pacificus*, as well as made use of this tool to gain mechanistic insights into various biological processes in *P. pacificus*, and analyzed the evolution of gene expression in a comparative context with *C. elegans*.

2

Results

2.1 Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*

Amit Sinha, Ralf J Sommer and Christoph Dieterich

BMC Genomics 2012, 13:254
doi:10.1186/1471-2164-13-254.

2.1.1 Synopsis

Previous research has shown that the endocrine mechanism regulating the dauer switch is conserved in *C. elegans*, *P. pacificus* and the vertebrate parasite *Strongyloides papillosus*. The hormone Dafachronic acid acts via the nuclear hormone receptor DAF-12/NHR in all the three species (Ogawa et al., 2009). Further, the dauer switch in *P. pacificus* needs the activity of the forkhead family transcription factor Ppa-DAF-16/FOXO in a manner similar to that known in *C. elegans* (Ogawa et al., 2011). However, it was not known whether the dauer larvae of different species execute similar gene expression programs downstream to conserved regulatory factors DAF-12 and DAF-16. To comprehensively answer this question, we generated and compared the expression profile of dauers and dauer recovery stages in both *P. pacificus* and *C. elegans*. Using species-specific microarrays covering about ∼20,000 and ∼21,000 genes in *C. elegans* and *P. pacificus* respectively, we measured the relative change in gene

expression levels for the dauer exit time course. Comparing the fold-changes of 1:1 homologs across the two species, we found that most of the downstream genes activated during the dauer life-stages are surprisingly different in *P. pacificus* and *C. elegans*. This divergence either highlights the effect of ecological adaptations on expression of genes across homologous developmental stages, or indicates the activity of non-adaptive forces. Further, the non-homologous portion of the transcriptomes of the two species was a major contributor to the developmental expression profile of each of the species. We could also show developmental regulation of cellulases and diapausins, two of the gene families acquired by horizontal gene transfer, as well as differential expression of many *P. pacificus* specific pioneer genes. In summary, this study could show that the downstream effectors of homologous developmental processes can diverge considerably, especially if the developmental process is under ecological control. This study raises the issue of how much of the gene expression is adaptive versus a consequence of drift in the gene regulatory networks, and argues for extending the concept of developmental systems drift (True and Haag, 2001) from the upstream regulators to the downstream effector genes.

2.1.2 Contributions

I designed, tested and validated the custom microarray platform for *P. pacificus*. I also standardized the data-analysis pipeline for microarray data, taking into account the transcriptional repression observed in dauers (Dalley and Golomb, 1992), and making use of external positive control as spike-ins in the normalization procedure. I analyzed all the data, with help from Dr. Christoph Dieterich, and interpreted it in the context of nematode ecology and evolution. I was actively involved in experimental design and preparation of the manuscript. In total, my contribution to this work was 100%.

2.2 A system wide analysis of the evolution of innate immunity in the nematode model species *Caenorhabditis elegans* and *Pristionchus pacificus*

Amit Sinha, Robbie Rae, Igor Iatsenko and Ralf J Sommer

PLoS One 2012, 7(9): e44255.
doi:10.1371/journal.pone.0044255.

Featured in:

Fugmann S: F1000 Recommendation of [Sinha A et al., PLoS ONE 2012, 7(9):e44255].
Faculty of 1000, 13 Oct 2012; DOI: 110.3410/f.717959303.793462703.
Web URL: f1000.com/717959303#eval793462703

2.2.1 Synopsis

Host-microbe interactions are key to understanding ecological factors affecting genome and transcriptome evolution. For bacterivorous nematodes like *C. elegans* and *P. pacificus*, many of the bacteria encountered in the wild can be utilized as a food source, while others can be pathogenic and need to be defended against through an innate immune response. Previous research in the lab has isolated a large number of bacteria from the wild (Rae et al., 2008) and examined their effects on survival in *C. elegans* and *P. pacificus* (Rae et al., 2010). Some bacteria were found to pathogenic to both the nematodes, while others killed only *C. elegans* but not *P. pacificus*. The molecular basis for this difference in response of the two nematodes to the same pathogens was not well understood. To characterize this differential immune response, we carried out gene expression of both *C. elegans* and *P. pacificus*, after four hours of exposure to gram-positive pathogens *Bacillus thuringiensis* DB27 and *Staphylococcus aureus*, and gram-negative pathogens *Serratia marcescens* and *Xenorhabdus nematophila*. We could show that (1) each nematode species mounts a pathogen-specific innate immune response (2) both nematode species respond to a give pathogen in widely different ways, regulating a different set of effector genes. Apart from identifying genes that might be important for pathogen defense in the two nematode species, our systematic approach shows the enormous complexity and diversity in gene expression in ecologically relevant

traits, suggesting that ecological factors such as exposure to different bacteria in the wild, shapes the gene expression programs in different species.

2.2.2 Contributions

Dr. Robbie Rae Igor Iatsenko carried out the survival assays on pathogens, and pathogen exposure experiments and collected worms for microarray experiments. I use these collected worms for RNA extraction and subsequent microarray experiments. I processed and analyzed the microarray data and interpreted it in the context of the known biology of innate immune response in *C. elegans*. I was actively involved in experimental design and preparation of the manuscript. In total, my contribution to this work was about 60%.

2.3 Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode *Pristionchus pacificus*

Robbie Rae*, Amit Sinha* and Ralf J. Sommer

* = Authors contributed equally

PLoS Pathogens 2012, 8(8): e1002864

doi:10.1371/journal.ppat.1002864

2.3.1 Synopsis

Ablation of germline leads to an increase in lifespan of various animals, including *C. elegans* and *P. pacificus* (Hsin and Kenyon, 1999; Patel et al., 2002). The germline-less, long-lived *C. elegans* is also resistant to bacterial pathogens (Alper et al., 2010; TeKippe and Aballay, 2010). In *C. elegans*, this lifespan extension and increased pathogen resistance requires the presence of the somatic gonad, and the activities of FOXO/DAF-16 and NHR/DAF-12 transcription factors (Alper et al., 2010; Hsin and Kenyon, 1999; TeKippe and Aballay, 2010). However, it is not known whether the pathogen resistance of germline-less animals is a conserved phenomenon in species other than *C. elegans*. Further, the nature and role of genes responsible for increased lifespan and increased pathogen resistance in germline-less animals is also not well characterized. We showed that germline ablation in *P. pacificus* also increases resistance to the gram-negative pathogen *Serratia marcescens*, and that this increase in pathogen resistance works in a gonad dependent manner. We analyzed the gene expression profiles of long-lived germline-ablated animals, and compared it with the genes regulated in these long-lived animals upon exposure to *S. marcescens* for four hours. Analysis of microarray data implicated *Ppa-daf-16* and *Ppa-daf-12* in regulating gene expression and hence long-lived phenotype of long lived animals. This role of *Ppa-daf-16* and *Ppa-daf-12* was then confirmed through germline ablation experiments on the respective mutants, showing a conservation of function for these genes across the two nematode species. Microarray data analysis also suggested that the gene expression changes that occur upon germline ablation and impart longer lifespan (∼3,000 genes), are also the major contributor towards enhanced pathogen resistance, because exposure of these animals to pathogen

resulted in differential expression of a only a few genes (∼200 genes). Microarray data analysis also identified a role for genes involved in core cellular processes like translation, proteasome activity and maintenance of nuclear membrane integrity.

2.3.2 Contributions

Dr. Robbie Rae carried out the ablation experiments, pathogen exposure experiments, and collection of *P. pacificus* worms for all microarray experiments. I used these collected animals for RNA extraction and subsequent microarray experiments. I analyzed the microarray data for all comparisons and identified *Ppa-daf-16* and *Ppa-daf-12* as potential regulators of gene expression. Dr. Robbie Rae functionally validated these candidates by ablating the germlines of the respective mutants. I also analyzed the microarray data to identify other functionally relevant genes and biological processes. I was actively involved in experimental design and preparation of the manuscript. In total, my contribution to this work was about 40%.

2.4 Phosphoproteome of *Pristionchus pacificus* provides insights into architecture of signaling networks in nematode models

Nadine Borchert*, Karsten Krug*, Florian Gnad, Amit Sinha, Ralf J Sommer and Boris Macek

* = Authors contributed equally

Molecular and Cellular Proteomics 2012

doi:10.1074/mcp.M112.022103

2.4.1 Synopsis

The phosphoproteome is the set of all identified phosphorylated proteins in an organism. Protein kinases are an important class of enzymes that regulate a wide variety of cellular processes such as signal transduction pathways, by phosphorylating their target proteins. The characterization of phosphoproteome identifies the genes that are target of such regulatory processes, and hence the phosphoproteomes of major model systems, including *C. elegans* have already been analyzed (Bodenmiller et al., 2007; Ficarro et al., 2002; Olsen et al., 2006; Pan et al., 2008; Zielinska et al., 2009). For *P. pacificus*, the transcriptome and proteome have been well characterized in various studies (Borchert et al., 2010; Dieterich et al., 2008) but its phosphoproteome had not been studied so far. The current study focused on a large scale, qualitative analysis of the phosphoproteome of *P. pacificus* making use of a two-stage phosphopeptide enrichment followed by high accuracy mass spectrometry. The study identified about 7,000 phosphorylation events on about 2,100 proteins, of which only 340 sites were conserved between *C. elegans* and *P. pacificus*. The two nematode species also differ in the distributions of phosphorylation events (phosphorylated serines, threonines and tyrosines), as well as the number and distribution of different kinase families (their "kinomes"). To see if this difference in kinomes is biologically relevant, I analyzed the expression levels of different kinase classes in the dauer-exit transcriptome data available from our our previous studies (Sinha et al., 2012). Interestingly, in *C. elegans* dauers show a higher average expression of "Pkinase_Tyr" genes versus "Pkinase" genes. This difference in expression levels of different kinase families is not apparent in the dauers of *P. pacificus*.

This analysis reveals that both nematodes express all classes of kinases and point to their potentially different usage in dauer stage of the life cycle.

2.4.2 Contributions

Dr. Nadine Borchert, Karsten Krug, Florian Gnad and Prof. Boris Macek carried out all the proteomics experiments and data analysis. After the annotation and identification of different protein kinases from this analysis, I analyzed the gene expression data from the dauer exit time course in *C. elegans* and *P. pacificus* (Sinha et al., 2012), to identify the differences in use of these kinase classes during this developmental transition across the two species. In total, my contribution to this work was about 10%.

2.5 Genes and developmental pathways regulating innate immunity and lifespan in *C. elegans*

C. elegans has proved to be an excellent model to study the genetic basis of lifespan regulation (Kenyon, 2010) and to study host-pathogen interactions and innate immunity (Irazoqui et al., 2010). In this section I will briefly introduce and describe the results from unpublished projects where we have combined the transcriptomic approaches with the power of genetic tools in *C. elegans* to study the mechanistic basis of lifespan and pathogen-response regulation.

2.5.1 Reproductive signaling genes that regulate extreme longevity and innate immunity in *C. elegans*

Robbie Rae*, Amit Sinha* and Ralf J. Sommer

* = Authors contributed equally (Manuscript under preparation)

2.5.1.1 Synopsis

Germline ablation in *C. elegans* leads to a dramatic increase in lifespan (Hsin and Kenyon, 1999) and resistance to bacterial pathogens (Alper et al., 2010; TeKippe and Aballay, 2010), and both these processes are dependent on activity of FOXO/DAF-16 and NHR/DAF-12. However this increase in lifespan is not due to a simple trade-off between longevity and reproduction, as the ablation of the somatic gonad results in animals that are sterile but are not long-lived (Hsin and Kenyon, 1999). Based on these ablation experiments, it has been proposed the somatic gonad tissue is responsible for generation and secretion of lifespan enhancing signals, which counteracts the lifespan reducing signals originating from the germline (Hsin and Kenyon, 1999). Since the activity of the NHR/DAF-12 and steroid hormone synthesizing enzyme cytochrome DAF-9 are required for this increased longevity, the secreted signal has been proposed to be lipophilic in nature (Berman and Kenyon, 2006; Gerisch et al., 2001; Hsin and Kenyon, 1999). Search for other genes involved in this gonad-dependent lifespan extension has identified a role for an intestinal protein KRI-1 (Berman and Kenyon, 2006) and a transcriptional elongation factor TCER-1 (Ghazi et al., 2009). However, apart from these few genes, the natures of the signal or role of other genes mediating this phenomenon are not known. To tackle this question, we have taken a functional genomic

approach as described below. We first compared the expression profiles of germline-less long-lived animals against that of germline- and gonad-less normal-lived animals, to identify a set of candidate genes that are upregulated in the long-lived animals in a gonad dependent manner (schematic Figure 2.1). We then analyzed the functional requirement for the activity of these genes through a RNAi based screen in germline-less animals, that would shorten the survival of germline-less animals exposed to the gram negative pathogen *Xenorhabdus nematophila*.

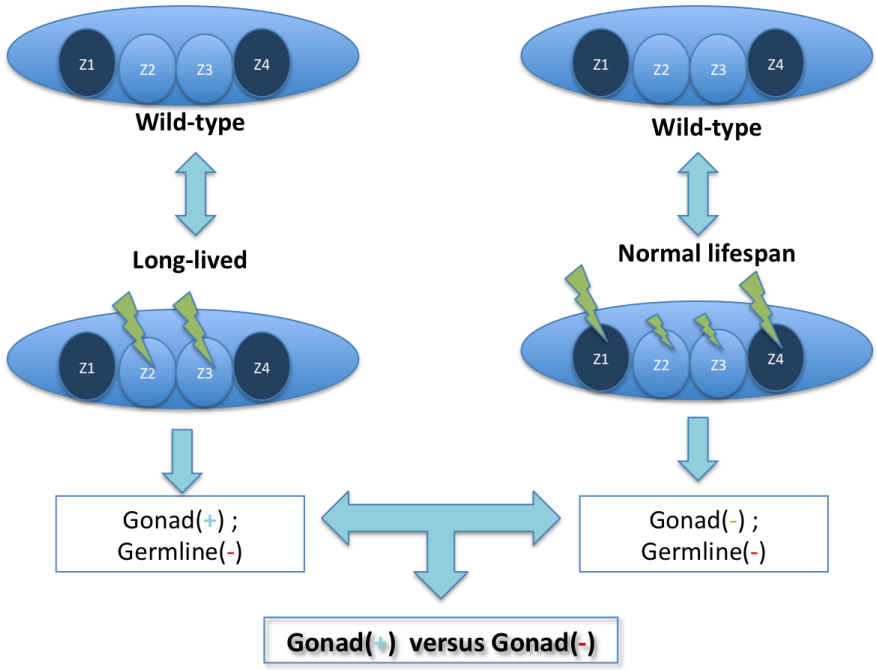


Figure 2.1: Schematic of microarray experiments of germline ablated and gonad ablated animals -

2.5.1.2 Results

Microarray data analysis identified about 3,000 genes being upregulated in the long-lived animals, while only 62 genes were downregulated. Many of these genes can be expected to be simply those that are involved in germline development. Interestingly, about 250 genes are known to have a "shortened life-span"phenotype upon RNAi, and about one-third of those genes are significantly upregulated in our experiment. This

indicates that these lifespan enhancing genes are also regulated in a gonad independent manner.

A gene ontology analysis indicates that the translational activity at the ribosome, cytochrome C oxidase activity at the mitochondrial membrane, and proteasomal activity in the cytoplasm are being regulated in gonad-dependent manner in the long-lived animals. Consistent with this, RNAi knockdown of the cytochrome C oxidase gene *cco-2* results in a significant decrease in lifespan of germline-less long-lived animals. Other candidates whose RNAi resulted in reduction of lifespan include the genes for nuclear hormone receptor *nhr-49*, the PTEN homolog *daf-18*, and a 14-3-3 protein encoding gene *par-5*. Our screen also uncovered genes with previously unknown roles in lifespan regulation (e.g. *gna-2*) and others with completely unknown biological function (e.g. T12G3.6, C14B1.2, Y45F10C.2). Interestingly, we observe a massive upregulation of *tbh-1*, a gene expressed specifically in the gonad that encodes a tyramine beta-hydroxylase enzyme involved in biosynthesis of the neurotransmitter octopamine. This raises the interesting possibility that the lifespan-promoting signal from the gonad might be a secreted neurotransmitter such as octopamine.

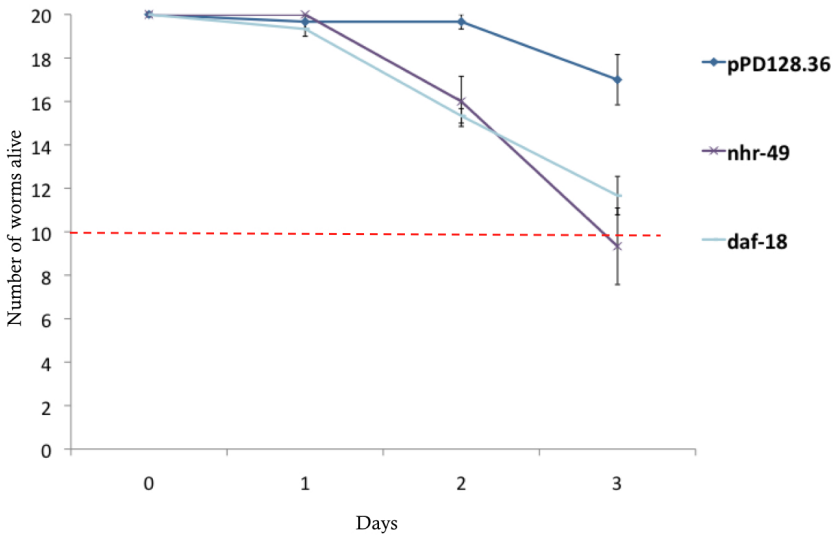


Figure 2.2: RNAi screen identifies genes regulating lifespan - Survival curves for some of the genes tested in our screen, whose reduction of function leads to reduction in survival of germline-less *glp-1* mutant animals exposed to *X. nematophila*. The curve pPD128.36 corresponds to the *E. coli* with an empty RNAi-vector, used as the control.

2.5.1.3 Contributions

Dr. Robbie Rae carried out all the ablation experiments, including collection of worms for microarray experiments. I used these collected worms for RNA extraction and subsequent microarray experiments. I analyzed the microarray data to identify the candidate genes. Both of us screened candidate genes using RNAi-by-feeding approach to functionally validate these candidates. In total, my contribution to the project is about 40%.

2.5.2 Transcriptome analysis of *C. elegans* mutants *btr-1* and *btr-2* that are resistant to the pathogenic bacteria *Bacillus thuringiensis* DB27

Igor Iatsenko, Amit Sinha and Ralf J Sommer

2.5.2.1 Synopsis

The gram negative bacteria *Bacillus thuringiensis* strain DB27 has been isolated and identified as one of the most lethal pathogens that kills *C. elegans* in less than 24 hours (Rae et al., 2010). In a screen for resistant mutants, Igor Iatsenko has isolated two complementation groups, which map to previously uncharacterized genes C09H10.6 and C09H10.7. These genes have now been named *btr-1* and *btr-2*, where *btr* stands for Bacillus thuringiensis resistant. Igor is further characterizing the biological function of these genes. To complement these studies, I have carried out transcriptome analysis of the induced pathogen response of these two resistant mutants versus wild-type controls after exposure to *B. thuringiensis* DB27 for 4 hours.

2.5.2.2 Results

Microarray analysis of *btr-1* mutants exposed to pathogens revealed upregulation of about 300 genes and downregulation of about 630 genes. Interestingly, the downregulated genes are enriched for genes that are known to be targets of the *C. elegans* Argonaute/Dicer protein DCR-1, suggesting a novel role for DCR-1 in resistance to pathogen. Consistent with this, we found that DCR-1 mutants also show resistance to *B. thuringiensis* DB27, and that *btr-1* mutants are also deficient in RNAi. In addition, *dcr-1* itself was found to be downregulated in *btr-1* mutants, and expression of *dcr-1*

in *btr-1* mutants reverted their pathogen resistance phenotype. These results together indicate that *btr-1* indeed works through *dcr-1*. Further characterization of the mutant by Igor showed that *btr-1* interacts with Synthetic MultiVulva or SynMuv B class of genes, which are involved in chromatin remodeling (Cui et al., 2006), suggesting that the *btr-1* gene regulates chromatin remodeling.

Microarray analysis of the *btr-2* mutant exposed to pathogens identified 218 upregulated genes and 24 downregulated genes. Surprisingly, there was very little overlap between these genes and the genes differentially expressed in the *btr-1* mutant, indicating that these two genes have separate biological functions and act through non-overlapping pathways. The genes regulated by *btr-2* were also enriched for targets of the p38 MAP Kinase *pmk-1* and the MAPKK *sek-1* pathways that play a key role in *C. elegans* innate immunity. Consistent with this, we verified that the pathogen resistance of *btr-2* mutants required activity of *pmk-1* and *sek-1*. Further analysis revealed that many of the genes in *btr-2* expression profile are known to be regulated in response to multiple pathogens, and many of them have a signal sequence indicative of secretion. This information together suggested that such genes might have key roles in anti-microbial defense and Igor tested if their reduction-of-function via RNAi reduced the survival of *btr-2* mutants on pathogen. Although RNAi for most of the candidate genes did not change the susceptibility of the *btr-2* mutant, RNAi of an uncharacterized gene F40F4.6 did reduce the pathogen survival of *btr-2* mutants down to wild-type levels, thereby identifying it as a novel component of *C. elegans* innate immune response. To identify upstream transcription factors that potentially regulate the expression of multiple *btr-2* targets, I analyzed the promoters of differentially expressed genes for enrichment of motifs corresponding to known transcription factor binding sites using the Allegro (Halperin et al., 2009). This analysis suggested a role for a *C. elegans* homeobox protein CEH-20 in *btr-2* function. Indeed Igor could verify that RNAi of *ceh-20* suppressed the resistance phenotype of *btr-2* mutants. Additionally *ceh-20* mutants themselves were observed to be more susceptible to the pathogens. These experiments thus identified a novel role for *ceh-20* role in innate immunity, apart from its known role in post-embryonic development (Jiang et al., 2009; Pellegrino et al., 2011). The gene *ceh-20* is known to interact with another hoemodomain protein *lin-39* in vulval development (Yang et al., 2005), which in turn interacts with the beta-catenin *bar-1* (Eisenmann et al., 1998). Consistent with these facts, Igor found that both *lin-39* and

bar-1 mediate the pathogen resistance phenotypes of *btr-2* mutants. We have thus identified novel roles for *ceh-20*, *lin-39* and *bar-1* in *btr-2* mediated pathogen response.

2.5.2.3 Contributions

The mutants *btr-1* and *btr-2* have been isolated, mapped and cloned by Igor Iatsenko. I have complemented his functional characterizations of these genes via my transcriptomic analysis. For *btr-1*, I identified a role for Dicer mediated gene regulation, which has been further validated by Igor. For *btr-2*, I identified a role for the p38 MAPK function, as well as novel role for the homeobox transcription factor *ceh-20* through promoter analysis of *btr-2* regulated genes. Igor could then confirm the role of *ceh-20* in innate immunity and could further link up to novel functions of *lin-39* and *bar-1* in innate immunity. In total, my contribution to the project is about 30%.

2.5.3 *Bacillus subtilis* increases *C. elegans* pathogen survival by activating innate immunity, stress resistance and longevity genes

Igor Iatsenko, Amit Sinha and Ralf J Sommer
Manuscript under preparation

2.5.3.1 Synopsis

C. elegans has been used successfully as a model to investigate innate immunity pathways activated in response to various pathogenic bacteria (Irazoqui et al., 2010). Some recent studies have also started using it as a model to investigate the effect of diet on subsequent immune response, especially focusing on probiotic effects of lactic acid bacteria (Cooper et al., 2009; Ikeda et al., 2007; Komura et al., 2010). It is also known now that a short pre-exposure of *C. elegans* to virulent or non-virulent strains of human enteropathogenic *E. coli* (EPEC) renders the worms resistant to subsequent exposure to pathogenic strains, a phenomenon referred to as "conditioning" (Anyanful et al., 2009).

C. elegans is cultured in the lab on *E. coli* OP50, which is not a very likely food source to be encountered by *C. elegans* in the wild. Previous studies in our lab have isolated a number of *Bacillus* strains, which are typically highly abundant in soil where nematodes like *C. elegans* can also be found (Rae et al., 2010). Most of these naturally

isolated Bacillus strains were found to be benign while some of them were highly toxic to *C. elegans* (Rae et al., 2010). However, it was not known whether using the benign bacteria as food source would affect the resistance of *C. elegans* to the pathogenic *Bacillus* strains. We observed that culturing *C. elegans* on a non-pathogenic *B. subtilis* (strain GS67) rendered them resistant to a subsequent exposure to the highly pathogenic *B. thuringiensis* strain DB27. We further observed that feeding the worms *B. subtilis* GS67 also leads to a general increase in resistance to multiple stressors like heat and oxidative stress, as well as an increase in lifespan. The focus of this study was to understand the molecular mechanism of this conditioning behavior. To identify the genes mediating this conditioning response, I have analyzed the transcriptome of worms that have been fed on *B. subtilis* GS67 versus those grown on control lab food *E. coli* OP50.

2.5.3.2 Results

Microarray data analysis identified 352 upregulated and 158 downregulated genes. Functional annotations of these genes indicates that exposure to the nonpathogenic *B. subtilis* GS67 is sufficient to induce a basal activation of innate immune response, with many genes known to have a potential role in anti-microbial defense, stress response and genes regulated in a FOXO/DAF-16 dependent manner. We also observed significant downregulation of a large number of genes encoding ribosomal proteins, suggesting a repression of translation in animals fed on *B. thuringiensis* 67. We could verify that RNAi mediated knockdown of individual ribosomal genes in animals fed on *E. coli* OP50 could indeed enhance their resistance to the pathogenic *B. thuringiensis* DB27. These results are in agreement with recent studies suggesting that the disruption of core cellular processes e.g. translation, activates the innate immune response in *C. elegans* (Dunbar et al., 2012; McEwan et al., 2012; Melo and Ruvkun, 2012). To search for upstream regulators responsible for activation of these innate immunity genes, I analyzed the promoter regions of differentially expressed genes to identify transcription factors whose binding sites are enriched in our data set. I identified enrichment for the binding of transcription factor families such as GATA (e.g. *elt-2*), STAT (*sta-1* and *sta-2*) and RUNX (*rnt-1*). Interestingly, ELT-2 and STA-2 have recently been implicated in *C. elegans* immune response (Dierking et al., 2011; Kerry et al., 2006). Igor could verify that RNAi mediated knockdown of *sta-1* ameliorated the conditioning

effect, suggesting novel role for STA-1 in *C. elegans* immune response. The gene *rnt-1* encodes for the sole homolog of Runx-family of transcription factors in *C. elegans*. Igor could also verify that RNAi mediated knockdown of *rnt-1* could abolish the benefits of conditioning, thus identifying a novel role for RNT-1 in *C. elegans* innate immunity.

2.5.3.3 Contributions

Igor has initially tested and observed the diet dependent conditioning in *C. elegans* and *Bacillus* interactions. I have complemented the search for molecular mechanisms for this phenomenon via microarray experiments and subsequent data analysis. From this analysis, I could identify a role for FOXO/DAF-16 transcription factor, translational repression, and potential roles for transcription factors STA-1 and RNT-1 in *C. elegans* innate immunity. All genetic validation experiments were carried out by Igor. In total, my contribution to this study is about 30%.

2.6 Gene organization into operon-like clusters in *Pristionchus pacificus* is remarkably different from *Caenorhabditis elegans*

Amit Sinha, Ralf J Sommer and Christoph Dieterich*

Zisong Chang*, Claudia Langnick*, Wei Chen*

* = Berlin Institute of Medical Systems Biology, Max Delbrück Center, Berlin.

2.6.1 Synopsis

Nematodes such as *C. elegans* are unique among the eukaryotes in having a operon like organization of genes, where genes from an operon are transcribed as a poly-cistronic transcript which is subsequently resolved by trans-splicing of 22 nt long splice leaders (SLs) at their 5 ends (Blumenthal, 2005). There are about 1,000 such operonic clusters in *C. elegans* genome, with 2 to 8 genes per operon and a total of about 2,600 member genes (Blumenthal et al., 2002). The phenomenon of trans-splicing is conserved across the entire nematode phylogeny, but SL1 and SL2-like splice leaders are only found in clade V nematodes, which includes *C. elegans* and *P. pacificus* (Guiliano and Blaxter, 2006). A few operons are conserved between *P. pacificus* and *C. elegans* (Guiliano and Blaxter 2006) while others are not conserved (Lee and Sommer, 2003). So far, the genome-wide trans-splicing landscape in *P. pacificus*, the number and nature of all its operons and their potential conservation as compared to *C. elegans* has not been studied systematically. In this study I have taken a RNAseq based approach to answer all these questions. I have purified and separated the SL1 spliced fraction and the SL2 spliced fraction of the transcriptome, via a pull-down strategy where biotinylated primers specific to either SL1 or SL2 splice leaders bind to respective targets, and are then purified on magnetic beads coated with streptavidin. These mRNA pools were then sent to the genome center at the BIMSB/MDC Berlin for high-throughput quantitative mRNA sequencing on the ABI SOLiD platform (see Figure 2.3 for a schematic of the experiment).

2.6.2 Results

From a bioinformatics analysis of *P. pacificus* genome, I could identify 2,522 potential operonic clusters comprised of a total 5,474 genes that are separated by less than

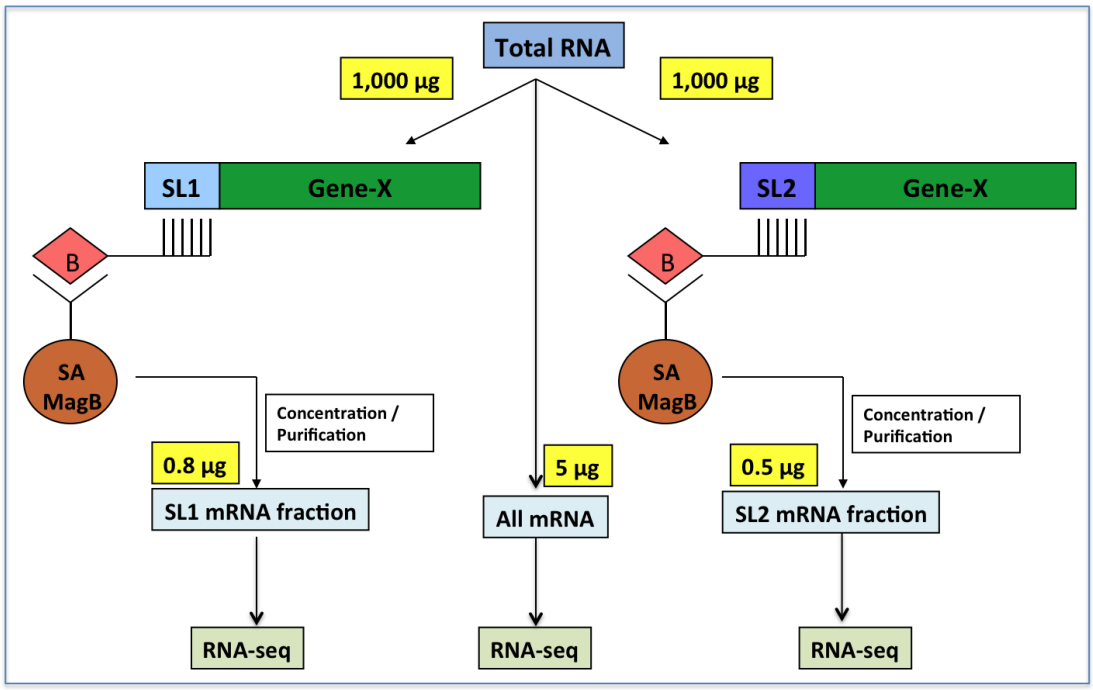


Figure 2.3: Schematic for purifying SL1 and SL2 trans-spliced fractions of the transcriptome -

500 base-pairs within each operon. We then assessed the splicing pattern for these operonic genes as well as all the other genes in the transcriptome by analyzing RNAseq data. Before the sequencing process, I first tested the efficiency and specificity of our pull-down strategy using qPCR experiments and observed about a 100 to 1,000-fold enrichment of trans-spliced transcripts in the SL1 and SL2 fractions relative to the total mRNA. From the RNAseq data of the entire poly-A tailed fraction, and the SL1 and SL2- trans-spliced fraction, we get reads mapping to 25,740 genes out of about 31,000 predicted genes in our latest genome assembly. We detected 20,575 genes in both the SL1 and SL2 spliced pools, 1,334 genes were exclusively trans-spliced to SL1 while 1,193 genes were trans-spliced exclusively to the SL2 splice leader. In addition, 1,266 genes were detected only in the poly-A pool but not in the trans-spliced pool. Thus, we find that about 70%to 80% of the *P. pacificus* mRNAs receive splice leaders via trans-splicing, a range similar to that reported in *C. elegans* (Blumenthal, 2005). I further found that out of the total 2,522 predicted operons, about 1,294 operons indeed show a preferential trans-splicing to SL2. The other predicted operons might be trans-

spliced to other variants of SL2-like splice leaders.

Next, I checked how many of these operons are at least partially syntenic and hence conserved with respect to the corresponding operon made of its homologous genes in *C. elegans*. For this analysis, I first identified regions of synteny between entire genomes of *C. elegans* and *P. pacificus*, using the tool CYNTENATOR (Rödelsperger and Dieterich, 2010) and found 2,066 syntenic blocks comprised of two more more genes. I then checked how many of these syntenic blocks contained genes predicted to be in an operon in the two respective nematode genomes. Interestingly, I found only 143 operons that are conserved between *C. elegans* and *P. pacificus* genomes, which comprise of about 3,339 and 2,522 predicted operons respectively. These results together suggest that gene ordering and organization into operon like clusters has diverged substantially in the 250-400 million years separting the two species.

2.6.3 Contributions

I purified the mRNA and the SL1 and SL2 spliced fractions from the mRNA pools, and verified that the strategy indeed efficiently separates SL1 and SL2 pools, with minimal cross-contamination. The SOLiD sequencing was carried out at the genome-center at MDC Berlin, and Dr. Christoph Dieterich mapped all the reads to the latest assembly of *P. pacificus* genome. I carried out the microsynteny analysis using CYNTENATOR tool (Rödelsperger and Dieterich, 2010) as well as all the subsequent bioinformatics analysis to (1) identify potential operons, (2) their conservation with respect to *C. elegans*, and (3) the study of SL1 versus SL2 splicing patterns across all the operons in *P. pacificus*. In total, my contribution to this study is about 70%.

Bibliography

Abouheif E and Wray GA. 2002. Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**: 249–252. 13

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al.. 2000. The genome sequence of drosophila melanogaster. *Science (New York, N.Y.)* **287**: 2185–2195. PMID: 10731132. 8

Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, and Lake JA. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**: 489–493. PMID: 9168109. 6

Alper S, McElwee MK, Apfeld J, Lackford B, Freedman JH, and Schwartz DA. 2010. The caenorhabditis elegans germ line regulates distinct signaling pathways to control lifespan and innate immunity. *Journal of Biological Chemistry* **285**: 1822 –1828. 19, 24, 31, 35

Anyanful A, Easley KA, Benian GM, and Kalman D. 2009. Conditioning protects c. elegans from lethal effects of enteropathogenic e. coli by activating genes that regulate lifespan and innate immunity. *Cell Host & Microbe* **5**: 450–462. PMID: 19454349. 40

Beldade P, Koops K, and Brakefield PM. 2002. Developmental constraints versus flexibility in morphological evolution. *Nature* **416**: 844–847. PMID: 11976682. 13

Bento G, Ogawa A, and Sommer RJ. 2010. Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution. *Nature* **466**: 494–497. 7

BIBLIOGRAPHY

Berman JR and Kenyon C. 2006. Germ-cell loss extends c. elegans life span through regulation of DAF-16 by kri-1 and lipophilic-hormone signaling. *Cell* **124**: 1055–1068. PMID: 16530050. 19, 35

Blaxter M. 2011. Nematodes: the worm and its relatives. *PLoS biology* **9**: e1001050. PMID: 21526226. 5, 6

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, et al.. 1998. A molecular evolutionary framework for the phylum nematoda. *Nature* **392**: 71–75. PMID: 9510248. 5

Blumenthal T. 2005. Trans-splicing and operons. *WormBook: The Online Review of C. Elegans Biology* pp. 1–9. PMID: 18050426. 20, 43, 44

Blumenthal T, Evans D, Link CD, Guffanti A, Lawson D, Thierry-Mieg J, Thierry-Mieg D, Chiu WL, Duke K, Kiraly M, et al.. 2002. A global analysis of caenorhabditis elegans operons. *Nature* **417**: 851–854. 20, 43

Blumenthal T and Gleason KS. 2003. Caenorhabditis elegans operons: form and function. *Nat Rev Genet* **4**: 110–118. 20

Bodenmiller B, Mueller LN, Pedrioli PGA, Pflieger D, Jünger MA, Eng JK, Aebersold R, and Tao WA. 2007. An integrated chemical, mass spectrometric and computational strategy for (quantitative) phosphoproteomics: application to drosophila melanogaster kc167 cells. *Molecular bioSystems* **3**: 275–286. PMID: 17372656. 33

Borchert N, Dieterich C, Krug K, Schütz W, Jung S, Nordheim A, Sommer RJ, and Macek B. 2010. Proteogenomics of pristionchus pacificus reveals distinct proteome structure of nematode models. *Genome Research* **20**: 837–846. PMID: 20237107. 6, 8, 10, 33

Brenner S. 1974. The genetics of caenorhabditis elegans. *Genetics* **77**: 71–94. PMID: 4366476. 4

Brown FD, D'Anna I, and Sommer RJ. 2011. Host-finding behaviour in the nematode pristionchus pacificus. *Proceedings. Biological Sciences / The Royal Society* PMID: 21411455. 8, 13

Brusca RC and Brusca GJ. 2003. *Invertebrates - Second Edition [Hardcover]*. Sinauer Associates, 2nd edition. 4

Cañestro C, Yokoi H, and Postlethwait JH. 2007. Evolutionary developmental biology and genomics. *Nature reviews. Genetics* **8**: 932–942. PMID: 18007650. 8

Cargill SL, Carey JR, Müller HG, and Anderson G. 2003. Age of ovary determines remaining life expectancy in old ovariectomized mice. *Aging Cell* **2**: 185–190. PMID: 12882411. 18

Carrasco AE, McGinnis W, Gehring WJ, and De Robertis EM. 1984. Cloning of an x. laevis gene expressed during early embryogenesis coding for a peptide region homologous to drosophila homeotic genes. *Cell* **37**: 409–414. PMID: 6327066. 2

Carroll SB. 2001. *From DNA to Diversity*. Blackwell Publishing, 1st edition. 2, 3

Cassada RC and Russell RL. 1975. The dauer larva, a post-embryonic developmental variant of the nematode caenorhabditis elegans. *Developmental Biology* **46**: 326–342. 13

Cobb NA. 1914. Nematodes and their relationships. In *Yearbook of the United States Department of Agriculture. 1914 (pub. 1915)*, Yearbook of agriculture, pp. 457–490. 4

Consortium ICGS. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**: 695–716. PMID: 15592404. 8

Coolon JD, Jones KL, Todd TC, Carr BC, and Herman MA. 2009. Caenorhabditis elegans genomic response to soil bacteria predicts environment-specific genetic effects on life history traits. *PLoS Genetics* **5**: e1000503. PMID: 19503598. 17

Cooper VS, Carlson WA, and Lipuma JJ. 2009. Susceptibility of caenorhabditis elegans to burkholderia infection depends on prior diet and secreted bacterial attractants. *PloS One* **4**: e7961. PMID: 19956737. 40

Costa M, Weir M, Coulson A, Sulston J, and Kenyon C. 1988. Posterior pattern formation in c. elegans involves position-specific expression of a gene containing a homeobox. *Cell* **55**: 747–756. PMID: 2903796. 2

BIBLIOGRAPHY

Couillault C, Pujol N, Reboul J, Sabatier L, Guichou JF, Kohara Y, and Ewbank JJ. 2004. TLR-independent control of innate immunity in caenorhabditis elegans by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat Immunol* **5**: 488–494. 17

Cui M, Kim EB, and Han M. 2006. Diverse chromatin remodeling genes antagonize the rb-involved SynMuv pathways in c. elegans. *PLoS genetics* **2**: e74. PMID: 16710447. 39

Dalley BK and Golomb M. 1992. Gene expression in the caenorhabditis elegans dauer larva: developmental regulation of hsp90 and other genes. *Developmental Biology* **151**: 80–90. PMID: 1577199. 28

Davidson EH. 2006. *The Regulatory Genome: Gene Regulatory Networks in Development And Evolution*. Academic Press, Waltham, Massachusetts. 3, 12

Davidson EH and Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. *Science (New York, N.Y.)* **311**: 796–800. PMID: 16469913. 12

Denver DR, Clark KA, and Raboin MJ. 2011. Reproductive mode evolution in nematodes: insights from molecular phylogenies and recently discovered species. *Molecular phylogenetics and evolution* **61**: 584–592. PMID: 21787872. 4

Dierking K, Polanowska J, Omi S, Engelmann I, Gut M, Lembo F, Ewbank JJ, and Pujol N. 2011. Unusual regulation of a STAT protein by an SLC6 family transporter in c. elegans epidermal innate immunity. *Cell Host & Microbe* **9**: 425–435. PMID: 21575913. 41

Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, et al.. 2008. The pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet* **40**: 1193–1198. 6, 8, 9, 10, 33

Dieterich C and Sommer RJ. 2009. How to become a parasite - lessons from the genomes of nematodes. *Trends in Genetics: TIG* **25**: 203–209. PMID: 19361881. 9, 14

Dobzhansky TG. 1955. *Evolution, genetics, and man*. Wiley. 2

Domazet-Lošo T and Tautz D. 2010. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* **468**: 815–818. PMID: 21150997. 12

Dong C, Davis RJ, and Flavell RA. 2002. MAP kinases in the immune response. *Annual review of immunology* **20**: 55–72. PMID: 11861597. 16

Dunbar TL, Yan Z, Balla KM, Smelkinson MG, and Troemel ER. 2012. C.elegans detects pathogen-induced translational inhibition to activate immune signaling. *Cell Host & Microbe* **11**: 375–386. PMID: 22520465. 41

Eisenmann DM, Maloof JN, Simske JS, Kenyon C, and Kim SK. 1998. The beta-catenin homolog BAR-1 and LET-60 ras coordinately regulate the hox gene lin-39 during caenorhabditis elegans vulval development. *Development (Cambridge, England)* **125**: 3667–3680. PMID: 9716532. 39

Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, and White FM. 2002. Phosphoproteome analysis by mass spectrometry and its application to saccharomyces cerevisiae. *Nature biotechnology* **20**: 301–305. PMID: 11875433. 33

Fielenbach N and Antebi A. 2008. C. elegans dauer formation and the molecular basis of plasticity. http://genesdev.cshlp.org/content/22/16/2149. 14, 15

Fitch DHA. 2005. Introduction to nematode evolution and ecology. *WormBook: the online review of C. elegans biology* pp. 1–8. PMID: 18050395. 5

Flatt T, Min KJ, D'Alterio C, Villa-Cuesta E, Cumbers J, Lehmann R, Jones DL, and Tatar M. 2008. Drosophila germ-line modulation of insulin signaling and lifespan. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 6368–6373. PMID: 18434551. 18, 19, 24

Floyd R, Abebe E, Papert A, and Blaxter M. 2002. Molecular barcodes for soil nematode identification. *Molecular ecology* **11**: 839–850. PMID: 11972769. 4

Friedman DB and Johnson TE. 1988a. A mutation in the age-1 gene in caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. *Genetics* **118**: 75–86. PMID: 8608934. 18

Friedman DB and Johnson TE. 1988b. Three mutants that extend both mean and maximum life span of the nematode, caenorhabditis elegans, define the age-1 gene. *Journal of Gerontology* **43**: B102–109. PMID: 3385139. 18

Gerhart J and Kirschner M. 1997. *Cells, Embryos and Evolution; toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability.* Blackwell Science, Malden, Massachusetts. 12

Gerisch B, Weitzel C, Kober-Eisermann C, Rottiers V, and Antebi A. 2001. A hormonal signaling pathway influencing c. elegans metabolism, reproductive development, and life span. *Developmental cell* **1**: 841–851. PMID: 11740945. 35

Ghazi A, Henis-Korenblit S, and Kenyon C. 2009. A transcription elongation factor that links signals from the reproductive system to lifespan extension in caenorhabditis elegans. *PLoS Genetics* **5**: e1000639. PMID: 19749979. 19, 35

Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, Scott G, Steffen D, Worley KC, Burch PE, et al.. 2004. Genome sequence of the brown norway rat yields insights into mammalian evolution. *Nature* **428**: 493–521. PMID: 15057822. 8

Gilad Y, Oshlack A, and Rifkin SA. 2006. Natural selection on gene expression. *Trends in Genetics* **22**: 456–461. 12

Gogarten JP. 2003. Gene transfer: gene swapping craze reaches eukaryotes. *Current biology: CB* **13**: R53–54. PMID: 12546803. 9

Gould SJ. 1977. *Ontogeny and Phylogeny.* Harvard University Press. 1

Guiliano DB and Blaxter ML. 2006. Operon conservation and the evolution of trans-splicing in the phylum nematoda. *PLoS Genetics* **2**: e198. PMC1657053. 20, 21, 43

Gupta BP and Sternberg PW. 2003. The draft genome sequence of the nematode caenorhabditis briggsae, a companion to c. elegans. *Genome biology* **4**: 238. PMID: 14659008. 4

Haegeman A, Jones JT, and Danchin EGJ. 2011. Horizontal gene transfer in nematodes: a catalyst for plant parasitism? *Molecular plant-microbe interactions: MPMI* **24**: 879–887. PMID: 21539433. 9

Haldane JBS. 1949. Disease and evolution. *La Ricerca Scientifica* **19**: 68–76. 16

Halperin Y, Linhart C, Ulitsky I, and Shamir R. 2009. Allegro: Analyzing expression and sequence in concert to discover regulatory programs. *Nucl. Acids Res.* **37**: 1566–1579. 39

Hamilton JB and Mestler GE. 1969. Mortality and survival: comparison of eunuchs with intact men and women in a mentally retarded population. *Journal of Gerontology* **24**: 395–411. PMID: 5362349. 18

Hashimshony T and Yanai I. 2010. Revealing developmental networks by comparative transcriptomics. *Transcription* **1**: 154–158. PMID: 21326891. 11, 12

Herrmann M, Mayer WE, Hong RL, Kienle S, Minasaki R, and Sommer RJ. 2007. The nematode pristionchus pacificus (Nematoda: diplogastridae) is associated with the oriental beetle exomala orientalis (Coleoptera: scarabaeidae) in japan. *Zoological Science* **24**: 883–9. PMID: 17960992. 6, 7, 8

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, and Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science (New York, N.Y.)* **313**: 101–104. PMID: 16825572. 3

Hong RL and Sommer RJ. 2006. Pristionchus pacificus: a well-rounded nematode. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **28**: 651–9. PMID: 16700067. 6

Hong RL, Witte H, and Sommer RJ. 2008. Natural variation in pristionchus pacificus insect pheromone attraction involves the protein kinase EGL-4. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 7779–84. PMID: 18509055. 7, 8

Hsin H and Kenyon C. 1999. Signals from the reproductive system regulate the lifespan of c. elegans. *Nature* **399**: 362–366. PMID: 10360574. 18, 19, 24, 31, 35

BIBLIOGRAPHY

Hu PJ. 2007. Dauer. *WormBook* . 14

Ikeda T, Yasui C, Hoshino K, Arikawa K, and Nishikawa Y. 2007. Influence of lactic acid bacteria on longevity of caenorhabditis elegans and host defense against salmonella enterica serovar enteritidis. *Applied and Environmental Microbiology* **73**: 6404–6409. PMID: 17704266. 40

Initiative AG. 2000. Analysis of the genome sequence of the flowering plant arabidopsis thaliana. *Nature* **408**: 796–815. PMID: 11130711. 8

Irazoqui JE, Urbach JM, and Ausubel FM. 2010. Evolution of host innate defence: insights from caenorhabditis elegans and primitive invertebrates. *Nature Reviews. Immunology* **10**: 47–58. PMID: 20029447. 16, 17, 35, 40

Irie N and Kuratani S. 2011. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nature communications* **2**: 248. PMID: 21427719. 12

Jiang Y, Shi H, and Liu J. 2009. Two hox cofactors, the Meis/Hth homolog UNC-62 and the Pbx/Exd homolog CEH-20, function together during c. elegans postembryonic mesodermal development. *Developmental biology* **334**: 535–546. PMID: 19643105. 39

Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, Ohler U, Bergman CM, and Tomancak P. 2010. Gene expression divergence recapitulates the developmental hourglass model. *Nature* **468**: 811–814. PMID: 21150996. 12

Kanzaki N, Ragsdale EJ, Herrmann M, Mayer WE, and Sommer RJ. 2012. Description of three pristionchus species (Nematoda: diplogastridae) from japan that form a cryptic species complex with the model organism p. pacificus. *Zoological Science* **29**: 403–417. PMID: 22639812. 6

Keeling PJ. 2009. Functional and ecological impacts of horizontal gene transfer in eukaryotes. *Current opinion in genetics & development* **19**: 613–619. PMID: 19897356. 9

Kenyon C. 1986. A gene involved in the development of the posterior body region of c. elegans. *Cell* **46**: 477–487. PMID: 3731276. 2

Kenyon C, Chang J, Gensch E, Rudner A, and Tabtiang R. 1993. A c. elegans mutant that lives twice as long as wild type. *Nature* **366**: 461–464. PMID: 8247153. 18

Kenyon CJ. 2010. The genetics of ageing. *Nature* **464**: 504–512. 18, 35

Kerry S, TeKippe M, Gaddis NC, and Aballay A. 2006. GATA transcription factor required for immunity to bacterial and fungal pathogens. *PLoS ONE* **1**: e77. PMC1762309. 41

Kienle S. 2012. *Microevolution in Pristionchus pacificus: Genetic Diversity and Natural Variation in Vulva Induction*. Ph.D. thesis, Eberhard Karls Universität Tubingen, Tubingen. 8

Kime DE and Larsen LO. 1987. Effect of gonadectomy and hypophysectomy on plasma steroid levels in male and female lampreys (Lampetra fluviatilis, l.). *General and Comparative Endocrinology* **68**: 189–196. PMID: 2962897. 18

Kimura M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press. 2

Kiontke K and Sudhaus W. 2006. Ecology of caenorhabditis species. *WormBook* . 4

Klass M and Hirsh D. 1976. Non-ageing developmental variant of caenorhabditis elegans. *Nature* **260**: 523–525. PMID: 1264206. 13, 14

Klass MR. 1983. A method for the isolation of longevity mutants in the nematode caenorhabditis elegans and initial results. *Mechanisms of ageing and development* **22**: 279–286. PMID: 6632998. 18

Komura T, Yasui C, Miyamoto H, and Nishikawa Y. 2010. Caenorhabditis elegans as an alternative model host for legionella pneumophila, and protective effects of bifidobacterium infantis. *Applied and Environmental Microbiology* **76**: 4105–4108. PMID: 20418445. 40

Kopp A, Barmina O, Hamilton AM, Higgins L, McIntyre LM, and Jones CD. 2008. Evolution of gene expression in the drosophila olfactory system. *Molecular Biology and Evolution* **25**: 1081–1092. 12

Lee DL, ed. 2002. *The Biology of Nematodes*. CRC Press, 1st edition. 4

Lee KZ and Sommer RJ. 2003. Operon structure and trans-splicing in the nematode pristionchus pacificus. *Molecular Biology and Evolution* **20**: 2097–103. PMID: 12949121. 21, 43

Lemaitre B, Nicolas E, Michaut L, Reichhart JM, and Hoffmann JA. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in drosophila adults. *Cell* **86**: 973–983. PMID: 8808632. 16

Levin M, Hashimshony T, Wagner F, and Yanai I. 2012. Developmental milestones punctuate gene expression in the caenorhabditis embryo. *Developmental Cell* **22**: 1101–1108. PMID: 22560298. 12

Levine M, Rubin GM, and Tjian R. 1984. Human DNA sequences homologous to a protein coding region conserved between homeotic genes of drosophila. *Cell* **38**: 667–673. PMID: 6091895. 2

Lewis EB. 1978. A gene complex controlling segmentation in drosophila. *Nature* **276**: 565–570. PMID: 103000. 1, 2

Mahajan-Miklos S, Tan MW, Rahme LG, and Ausubel FM. 1999. Molecular mechanisms of bacterial virulence elucidated using a pseudomonas aeruginosa-caenorhabditis elegans pathogenesis model. *Cell* **96**: 47–56. PMID: 9989496. 16

Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, Kohara Y, and Ewbank JJ. 2002. Inducible antibacterial defense system in c. elegans. *Current Biology* **12**: 1209–1214. 16, 17

Mayer MG and Sommer RJ. 2011. Natural variation in pristionchus pacificus dauer formation reveals cross-preference rather than self-preference of nematode dauer pheromones. *Proceedings of the Royal Society B: Biological Sciences* **278**: 2784 – 2790. 7

Mayer W, Schuster L, Bartelmes G, Dieterich C, and Sommer R. 2011. Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evolutionary Biology* **11**: 13. 9, 10

Mayr E. 1963. *ANIMALS SPECIES AND EVOLUTION*. THE BELKNAP PRESS/HARVARD UNIVERSITY. 2

McEwan DL, Kirienko NV, and Ausubel FM. 2012. Host translational inhibition by pseudomonas aeruginosa exotoxin a triggers an immune response in caenorhabditis elegans. *Cell Host & Microbe* **11**: 364–374. PMID: 22520464. 41

McGinnis W, Hart CP, Gehring WJ, and Ruddle FH. 1984. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of drosophila. *Cell* **38**: 675–680. PMID: 6091896. 2

McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, Wylie K, Mardis ER, Wilson RK, et al.. 2001. A physical map of the human genome. *Nature* **409**: 934–941. PMID: 11237014. 8

Medzhitov R, Preston-Hurlburt P, and Janeway C A J. 1997. A human homologue of the drosophila toll protein signals activation of adaptive immunity. *Nature* **388**: 394–397. PMID: 9237759. 16

Meiklejohn CD, Parsch J, Ranz JM, and Hartl DL. 2003. Rapid evolution of male-biased gene expression in drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 9894–9899. PMID: 12907700. 12

Melo JA and Ruvkun G. 2012. Inactivation of conserved c.elegans genes engages pathogen- and xenobiotic-associated defenses. *Cell* **149**: 452–466. PMID: 22500807. 41

Moczek AP and Nijhout HF. 2004. Trade-offs during the development of primary and secondary sexual traits in a horned beetle. *The American naturalist* **163**: 184–191. PMID: 14970921. 13

Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B. 2011. How many species are there on earth and in the ocean? *PLoS Biol* **9**: e1001127. 1

Morgan K, McGaughran A, Villate L, Herrmann M, Witte H, Bartelmes G, Rochat J, and Sommer RJ. 2012. Multi locus analysis of pristionchus pacificus on la runion island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. *Molecular Ecology* **21**: 250–266. PMID: 22126624. 6, 7, 8

Morris JZ, Tissenbaum HA, and Ruvkun G. 1996. A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in caenorhabditis elegans. *Nature* **382**: 536–539. PMID: 8700226. 18

Motola DL, Cummins CL, Rottiers V, Sharma KK, Li T, Li Y, Suino-Powell K, Xu HE, Auchus RJ, Antebi A, et al.. 2006. Identification of ligands for DAF-12 that govern dauer formation and reproduction in c. elegans. *Cell* **124**: 1209–1223. 16

Nüsslein-Volhard C and Wieschaus E. 1980. Mutations affecting segment number and polarity in drosophila. *Nature* **287**: 795–801. PMID: 6776413. 1, 2

Ogawa A, Bento G, Bartelmes G, Dieterich C, and Sommer RJ. 2011. Pristionchus pacificus daf-16 is essential for dauer formation but dispensable for mouth form dimorphism. *Development* **138**: 1281 –1284. 7, 16, 18, 27

Ogawa A, Streit A, Antebi A, and Sommer RJ. 2009. A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes. *Current Biology* **19**: 67–71. 7, 16, 27

Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, and Mann M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**: 635–648. PMID: 17081983. 33

O'Riordan VB and Burnell AM. 1990. Intermediary metabolism in the dauer larva of the nematode caenorhabditis elegans–II. the glyoxylate cycle and fatty-acid oxidation. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **95**: 125–130. 14

Pan C, Gnad F, Olsen JV, and Mann M. 2008. Quantitative phosphoproteome analysis of a mouse liver cell line reveals specificity of phosphatase inhibitors. *Proteomics* **8**: 4534–4546. PMID: 18846507. 33

Parikh A, Miranda ER, Katoh-Kurasawa M, Fuller D, Rot G, Zagar L, Curk T, Sucgang R, Chen R, Zupan B, et al.. 2010. Conserved developmental transcriptomes in evolutionarily divergent species. *Genome Biology* **11**: R35. PMID: 20236529. 12

Patel MN, Knight CG, Karageorgi C, and Leroi AM. 2002. Evolution of germ-line signals that regulate growth and aging in nematodes. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 769–774. PMID: 11805331. 19, 31

Pellegrino M, Steinbach N, Stensmyr MC, Hansson BS, and Vosshall LB. 2011. A natural polymorphism alters odour and DEET sensitivity in an insect odorant receptor. *Nature* **478**: 511–514. PMID: 21937991. 3, 39

Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, et al.. 2007. Diet and the evolution of human amylase gene copy number variation. *Nature genetics* **39**: 1256–1260. PMID: 17828263. 3

Pires-daSilva A and Sommer RJ. 2003. The evolution of signalling pathways in animal development. *Nature Reviews. Genetics* **4**: 39–49. PMID: 12509752. 2

Pires-daSilva A and Sommer RJ. 2004. Conservation of the global sex determination gene tra-1 in distantly related nematodes. *Genes & Development* **18**: 1198–208. PMID: 15155582. 7

Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, et al.. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in tlr4 gene. *Science (New York, N.Y.)* **282**: 2085–2088. PMID: 9851930. 16

Prud'homme B and Gompel N. 2010. Evolutionary biology: Genomic hourglass. *Nature* **468**: 768–769. 12

Pujol N, Zugasti O, Wong D, Couillault C, Kurz CL, Schulenburg H, and Ewbank JJ. 2008. Anti-fungal innate immunity in c. elegans is enhanced by evolutionary diversification of antimicrobial peptides. *PLoS Pathogens* **4**: e1000105. PMC2453101. 17

Rae R, Iatsenko I, Witte H, and Sommer RJ. 2010. A subset of naturally isolated bacillus strains show extreme virulence to the free-living nematodes caenorhabditis elegans and pristionchus pacificus. *Environmental Microbiology* **12**: 3007–3021. PMID: 20626457. 8, 17, 23, 24, 29, 38, 40, 41

Rae R, Riebesell M, Dinkelacker I, Wang Q, Herrmann M, Weller AM, Dieterich C, and Sommer RJ. 2008. Isolation of naturally associated bacteria of necromenic pristionchus nematodes and fitness consequences. *The Journal of Experimental Biology* **211**: 1927–36. PMID: 18515723. 8, 17, 23, 29

Rae R, Witte H, Rödelsperger C, and Sommer RJ. 2012. The importance of being regular: Caenorhabditis elegans and pristionchus pacificus defecation mutants are hypersusceptibile to bacterial pathogens. *International Journal for Parasitology* PMID: 22705203. 17

Ramsey M and Crews D. 2007. Steroid signaling system responds differently to temperature and hormone manipulation in the red-eared slider turtle (Trachemys scripta elegans), a reptile with temperature-dependent sex determination. *Sexual development: genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation* **1**: 181–196. PMID: 18391529. 13

Rangini Z, Frumkin A, Shani G, Guttmann M, Eyal-Giladi H, Gruenbaum Y, and Fainsod A. 1989. The chicken homeo box genes CHox1 and CHox3: cloning, sequencing and expression during embryogenesis. *Gene* **76**: 61–74. PMID: 2568317. 2

Ranz JM, Castillo-Davis CI, Meiklejohn CD, and Hartl DL. 2003. Sex-dependent gene expression and evolution of the drosophila transcriptome. *Science* **300**: 1742–1745. 12

Ranz JM and Machado CA. 2006. Uncovering evolutionary patterns of gene expression using microarrays. *Trends in Ecology & Evolution* **21**: 29–37. 11

Rifkin SA, Kim J, and White KP. 2003. Evolution of gene expression in the drosophila melanogaster subgroup. *Nat Genet* **33**: 138–144. 11

Rödelsperger C and Dieterich C. 2010. CYNTENATOR: progressive gene order alignment of 17 vertebrate genomes. *PLoS ONE* **5**: e8861. 45

Rödelsperger C and Sommer RJ. 2011. Computational archaeology of the pristionchus pacificus genome reveals evidence of horizontal gene transfers from insects. *BMC Evolutionary Biology* **11**: 239. PMID: 21843315. 11

Rudel D, Riebesell M, and Sommer RJ. 2005. Gonadogenesis in pristionchus pacificus and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad. *Developmental Biology* **277**: 200–21. PMID: 15572150. 7

Rudel D, Tian H, and Sommer RJ. 2008. Wnt signaling in pristionchus pacificus gonadal arm extension and the evolution of organ shape. *Proceedings of the National Academy of Sciences* **105**: 10826–10831. 7

Schena M, Shalon D, Davis RW, and Brown PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)* **270**: 467–470. PMID: 7569999. 11

Schlager B, Wang X, Braach G, and Sommer RJ. 2009. Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode pristionchus pacificus. *Genesis* **47**: 300–304. 6

Schuster LN and Sommer RJ. 2012. Expressional and functional variation of horizontally acquired cellulases in the nematode pristionchus pacificus. *Gene* PMID: 22814175. 9, 10

elegans Sequencing Consortium C. 1998. Genome sequence of the nematode c. elegans: a platform for investigating biology. *Science* **282**: 2012–2018. PMID: 9851916. 8

Sinha A, Sommer RJ, and Dieterich C. 2012. Divergent gene expression in the conserved dauer stage of the nematodes pristionchus pacificus and caenorhabditis elegans. *BMC Genomics* **13**: 254. PMID: 22712530. 33, 34

Sommer R and Ogawa A. 2011. Hormone signaling and phenotypic plasticity in nematode development and evolution. *Current Biology* **21**: R758–R766. 3, 4, 13, 14

Sommer RJ. 2009. The future of evo-devo: model systems and evolutionary theory. *Nature Reviews. Genetics* **10**: 416–422. PMID: 19369972. 2, 4

Sommer RJ, Carta LK, Kim SY, and Sternberg PW. 1996. Morphological, genetic and molecular description of pristionchus pacificus sp. n. (Nematoda: neodiplogastridae). *Fundamental and Applied Nematology* **19**: 511–522. 4, 6

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, et al.. 2003. The genome sequence of caenorhabditis briggsae: a platform for comparative genomics. *PLoS biology* **1**: E45. PMID: 14624247. 4

Steiner CC, Weber JN, and Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS biology* **5**: e219. PMID: 17696646. 3

TeKippe M and Aballay A. 2010. C. elegans germline-deficient mutants respond to pathogen infection using shared and distinct mechanisms. *PLoS ONE* **5**: e11777. 19, 24, 31, 35

Thompson JN. 1994. *The coevolutionary process*. University of Chicago Press. 16

Tian H, Schlager B, Xiao H, and Sommer RJ. 2008. Wnt signaling induces vulva development in the nematode pristionchus pacificus. *Current Biology: CB* **18**: 142–6. PMID: 18207741. 7

Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, and Kim DH. 2006. p38 MAPK regulates expression of immune response genes and contributes to longevity in c. elegans. *PLoS Genetics* **2**: e183. PMC1635533. 17

True JR and Haag ES. 2001. Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development* **3**: 109–119. PMID: 11341673. 28

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al.. 2001. The sequence of the human genome. *Science (New York, N.Y.)* **291**: 1304–1351. PMID: 11181995. 8

Viney ME, Thompson FJ, and Crook M. 2005. TGF-beta and the evolution of nematode parasitism. *International Journal for Parasitology* **35**: 1473–1475. PMID: 16139836. 16

Wagner A. 2005. *Robustness and Evolvability in Living Systems*. Princeton University Press, Princeton (New Jersey). 3

Wang X and Sommer RJ. 2011. Antagonism of LIN-17/Frizzled and LIN-18/Ryk in nematode vulva induction reveals evolutionary alterations in core developmental pathways. *PLoS Biol* **9**: e1001110. 7

Wang Z, Gerstein M, and Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* **10**: 57–63. PMID: 19015660. 11

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al.. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520–562. PMID: 12466850. 8

Weller AM, Mayer WE, Rae R, and Sommer RJ. 2010. Quantitative assessment of the nematode fauna present on geotrupes dung beetles reveals species-rich communities with a heterogeneous distribution. *The Journal of Parasitology* **96**: 525–531. PMID: 20557197. 8

Wernersson R, Juncker AS, and Nielsen HB. 2007. Probe selection for DNA microarrays using OligoWiz. *Nat. Protocols* **2**: 2677–2691. 22

West-Eberhard MJ. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, USA, 1st edition. 13

Wolpert. 2001. *Principles of Development*. Oxford Higher Education, 2nd edition. 3

Woolhouse MEJ, Webster JP, Domingo E, Charlesworth B, and Levin BR. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics* **32**: 569–577. PMID: 12457190. 16

Yanai I and Hunter CP. 2009. Comparison of diverse developmental transcriptomes reveals that coexpression of gene neighbors is not evolutionarily conserved. *Genome Research* **19**: 2214–2220. PMID: 19745112. 12

Yanai I, Peshkin L, Jorgensen P, and Kirschner MW. 2011. Mapping gene expression in two xenopus species: evolutionary constraints and developmental flexibility. *Developmental Cell* **20**: 483–496. PMID: 21497761. 12

Yang L, Sym M, and Kenyon C. 2005. The roles of two c. elegans HOX co-factor orthologs in cell migration and vulva development. *Development (Cambridge, England)* **132**: 1413–1428. PMID: 15750187. 39

Zaslaver A, Baugh LR, and Sternberg PW. 2011. Metazoan operons accelerate recovery from growth-arrested states. *Cell* **145**: 981–992. PMID: 21663799. 21

Zauner H, Mayer WE, Herrmann M, Weller A, Erwig M, and Sommer RJ. 2007. Distinct patterns of genetic variation in pristionchus pacificus and caenorhabditis elegans, two partially selfing nematodes with cosmopolitan distribution. *Molecular Ecology* **16**: 1267–80. PMID: 17391412. 6, 7

Zhang Y, Sturgill D, Parisi M, Kumar S, and Oliver B. 2007. Constraint and turnover in sex-biased gene expression in the genus drosophila. *Nature* **450**: 233–237. 12

Zheng M, Messerschmidt D, Jungblut B, and Sommer RJ. 2005. Conservation and diversification of wnt signaling function during the evolution of nematode vulva development. *Nature Genetics* **37**: 300–4. PMID: 15696167. 7

Zheng-Bradley X, Rung J, Parkinson H, and Brazma A. 2010. Large scale comparison of global gene expression patterns in human and mouse. *Genome Biology* **11**: R124. PMID: 21182765. 12

Zielinska DF, Gnad F, Jedrusik-Bode M, Winiewski JR, and Mann M. 2009. Caenorhabditis elegans has a phosphoproteome atypical for metazoans that is enriched in developmental and sex determination proteins. *Journal of proteome research* **8**: 4039–4049. PMID: 19530675. 33

Lebenslauf / Curriculum vitae

Amit Sinha

Born on **25th March, 1979**

In **Dehra Dun, India**

Education and Research Experience

July 2008 - Present **Ph. D. Thesis Dissertation**

Thesis supervisor : Prof. Dr. Ralf J Sommer

Department of Evolutionary Biology,

Max Planck Institute for Developmental Biology, Tübingen.

Thesis title: **Gene Expression Studies on the Evolution of Development and Innate Immunity in the Nematode *Pristionchus pacificus***

July 2004 - May 2007 **Master of Technology**

Thesis supervisor : Prof. Dr. Pradip Sinha

Biological Sciences and Bioengineering Department,

Indian Institute of Technology, Kanpur.

Thesis title: **Microarray Profiling of Hyperplastic Tumors in *Drosophila* to Understand the Genetic Network Regulating Organ Size**

July 1998 - May 2002 **Bachelor of Technology**

Materials Science and Metallurgical Engineering,

Indian Institute of Technology, Kanpur.

Publications

Sinha et al. BMC Genomics 2012, 13:254
http://www.biomedcentral.com/1471-2164/13/254

BMC Genomics

RESEARCH ARTICLE Open Access

Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*

Amit Sinha[1], Ralf J Sommer[1*] and Christoph Dieterich[2*]

Abstract

Background: An organism can respond to changing environmental conditions by adjusting gene regulation and by forming alternative phenotypes. In nematodes, these mechanisms are coupled because many species will form dauer larvae, a stress-resistant and non-aging developmental stage, when exposed to unfavorable environmental conditions, and execute gene expression programs that have been selected for the survival of the animal in the wild. These dauer larvae represent an environmentally induced, homologous developmental stage across many nematode species, sharing conserved morphological and physiological properties. Hence it can be expected that some core components of the associated transcriptional program would be conserved across species, while others might diverge over the course of evolution. However, transcriptional and metabolic analysis of dauer development has been largely restricted to *Caenorhabditis elegans*. Here, we use a transcriptomic approach to compare the dauer stage in the evolutionary model system *Pristionchus pacificus* with the dauer stage in *C. elegans*.

Results: We have employed Agilent microarrays, which represent 20,446 *P. pacificus* and 20,143 *C. elegans* genes to show an unexpected divergence in the expression profiles of these two nematodes in dauer and dauer exit samples. *P. pacificus* and *C. elegans* differ in the dynamics and function of genes that are differentially expressed. We find that only a small number of orthologous gene pairs show similar expression pattern in the dauers of the two species, while the non-orthologous fraction of genes is a major contributor to the active transcriptome in dauers. Interestingly, many of the genes acquired by horizontal gene transfer and orphan genes in *P. pacificus*, are differentially expressed suggesting that these genes are of evolutionary and functional importance.

Conclusion: Our data set provides a catalog for future functional investigations and indicates novel insight into evolutionary mechanisms. We discuss the limited conservation of core developmental and transcriptional programs as a common aspect of animal evolution.

Keywords: Dauer larvae, Developmental systems drift, Transcriptomics, Evolution of gene regulation, Horizontal gene transfer

Background

The primary mechanisms of an organism to respond to changing environmental conditions are alterations of gene expression profiles and the formation of dormant or dormant-like developmental stages. Nematodes are found in great numbers and species richness in most ecosystems on earth. This omnipresence is attributed to the evolution of a potent life history strategy to respond to changing environments [1]. Under favorable conditions nematodes such as *Caenorhabditis elegans* and *Pristionchus pacificus* undergo direct development, which can be completed within 3–4 days. In contrast, under harsh environmental conditions such as food scarcity, high temperature and high population density, these nematodes form an arrested developmental stage, the so-called dauer larvae [2].

* Correspondence: ralf.sommer@tuebingen.mpg.de; christoph.dieterich@mdc-berlin.de
[1]Max-Planck Institute for Developmental Biology, Department for Evolutionary Biology, Spemannstrasse 37, 72076, Tübingen, Germany
[2]Berlin Institute for Medical Systems Biology (BIMSB) at the Max Delbrück Center for Molecular Medicine Berlin, Robert-Rößle-Str. 10, 13125, Berlin, Germany

Dauer larvae are stress-resistant and "non-aging" and they facilitate the survival and dispersion of the organism [3]. Entry into the dauer stage represents the major life history response of nematodes to escape unfavorable environmental conditions [1]. This developmental switch is accompanied with remarkable transcriptional and metabolic changes as shown in several studies in *C. elegans* [4-8]. Since the dauer larva is an ecologically induced, conserved developmental stage found in many species of free-living nematodes, it can be expected that some genetic components involved in development might be conserved, while those facilitating ecological adaptation might diverge across species.

However, the similarity and divergence patterns of nematode transcriptomes have never been systematically investigated in the context of dauer larvae. While several transcriptional studies in *C. elegans* provide a platform for comparative studies [5,6,8-14], little is known in other free-living nematodes. Here we present a comparative transcriptomic approach to characterize the expression profiles of the dauer and dauer exit stages in *C. elegans* and *P. pacificus*. We measured transcriptome-wide expression changes with the help of the Agilent microarray platform, and use this data to identify the extent and nature of similarity or differences in the dauer-associated transcriptomes in the two nematode species.

C. elegans was the first metazoan to have its complete genome sequenced [15]. Ever since, *C. elegans* has been at the forefront of embracing novel "omics" technologies, including transcriptomics [16,17], proteomics [18,19] and RNAseq [20]. Comparative genomics is a powerful tool to address elementary questions in evolutionary developmental biology, such as what components of the regulatory "tool-kit" are conserved [21]. By now, the genome of six additional nematodes and many more transcriptomic studies have been reported (for review see [22]). Comparative transcriptomics opens up new vistas on how regulatory "tool-kit" components are employed to generate different morphologies across species [21,23]. In nematodes, comparative transcriptomic approaches provide a powerful way to analyze whole body responses to changes in the environment or the exposure to pathogens. The dauer stage represents a whole body response of the nematode to changing environmental conditions and is ideally suited for comparative transcriptomics studies.

We compare *C. elegans* to *P. pacificus*, an established model for comparative developmental biology, evolutionary biology and ecology [24]. Forward and reverse genetic as well as transgenic techniques have been established in *P. pacificus,* its genome has been sequenced [25] and its proteome has been analyzed by tandem mass spectrometry [26]. The sequencing of the *P. pacificus* genome revealed many important and unexpected features, such as a substantially larger size and a higher number of predicted protein-coding genes when compared to *C. elegans*, the presence of horizontal gene transfer and the duplication of genes encoding enzymes involved in the detoxification of xenobiotics [25]. Many of these features have been discussed in the context of a specific association with scarab beetles, in which *P. pacificus* is found in the wild [27]. Specifically, *P. pacificus* exists exclusively as dauer larvae on the living beetle and only reproduces after the beetle´s death by feeding on microbes that develop on the carcass [27-29]. This specific, so-called necromenic lifestyle shows the unique importance of the dauer stage and asks for a detailed functional investigation of the associated transcriptional programs.

Two developmental processes that have been studied in great detail at the genetic and molecular level in *P. pacificus* are the formation of the vulva, the egg-laying structure of nematode females and hermaphrodites [30-32] and the regulation of the dauer development [33,34]. Surprisingly, the regulation of vulva development in *C. elegans* and *P. pacificus* employs different signaling pathways to control the formation of this homologous morphological structure. In dauer formation, the two transcriptional regulators DAF-12, a nuclear hormone receptor, and DAF-16, a forkhead transcription factor, are well conserved between *C. elegans* and *P. pacificus* [33,34]. *In P. pacificus,* upstream factors consisting of insulin and TGF-beta signaling pathways in *C. elegans* await future analysis. Similarly, it is unclear whether downstream targets that are activated by DAF-12 and DAF-16 are also evolutionarily conserved. To address this question, we have taken a transcriptomic approach to directly compare expression profiles of the dauer stage (0 h) and dauer-exit stage (12 h post recovery-induction) in *C. elegans* and *P. pacificus*.

Here we show that *P. pacificus* has a more dynamic transcriptome during the dauer to dauer-exit transition as compared to that in *C. elegans*. The expression profiles from the two species look surprisingly different with limited overlap and weak correlation between orthologous genes that are differentially expressed during the dauer stage and the dauer-recovery process. Within the conserved genes, functionally different GO classes and protein domains are enriched in each profile, most striking differences being observed in regulation of metabolism related genes. These results highlight the importance of transcriptomic studies in revealing functional divergence in downstream effectors of homologous developmental processes, despite conservation of upstream regulatory factors.

Results

Significant drop in RNA abundance in dauer larvae of *C. Elegans* and *P. Pacificus*

The dormant dauer stage in *C. elegans* is associated with a global repression of Pol-II based transcription,

down to 11 - 17% of other stages [4]. We quantified the amount of total RNA per worm in the dauer and mix-stage samples and observed that dauers of *C. elegans* and *P. pacificus* contain approximately 20-fold less total RNA per worm as compared to the respective mix-stage sample. On average, ~100,000 dauer larvae yielded the same amount of total RNA as ~5,000 mixed stage worms. On top of this global transcriptional repression, we detected less mRNA in the total RNA of dauer larvae. The mRNA proportion is reduced to about half of the mix-stage levels in both species as measured by *in vitro* transcription (Additional file 1: Figure S1a). This global repression is expected to result in most of the genes being down regulated in a dauer versus mix-stage comparison. Our normalization strategy is based on differential weighing of spiked-in RNA probes (see Methods, Additional file 1: Figure S1b) and confirms this trend for both the species (Additional file 1: Figure S1c).

The *P. Pacificus* transcriptome is more dynamic in the dauer to dauer-exit transition

The Agilent microarray technology enabled simultaneous expression profiling of 20,143 genes in *C. elegans* and 20,446 genes in *P. pacificus*. Our experimental setup contrasted gene expression in the dauer stage (time point 0 hour) with the dauer-exit samples (time point 12 hour) via a common reference sample (mixed stage) for each species. This so called "common reference" experimental design for microarray makes use of a species-specific pool of RNA as a common technical reference, and was chosen as it potentially facilitates an extension of this study to other time-points and conditions, if needed (see [35], Methods and Additional file 1: Figure S2 for details). We have chosen the 12-hour post induction time-point to enable comparisons with published expression profiling experiments [5]. We have chosen the same 12-hour time-point for the *P. pacificus* dauer exit samples after verifying that the post-dauer development follows similar kinetics in both species. We base this conclusion on the following morphological and developmental observations: First, most of the population (~90%) of the recovering worms from *C. elegans* as well as *P. pacificus* resume pharyngeal pumping within 3 hours after inducing dauer exit. Second, at the 12-hour time-point (the stage when we collect the dauer-exit samples), no discernible morphological differences are found between the recovered dauers from both species. Third, the recovering animals from both species enter a lethargus stage between 13 to 14 hours post-recovery [3]. Fourth, the worms that are allowed to develop further at 20^0 C reach the next moult (L4 for *C. elegans,* J4 for *P. pacificus*) between 22 to 23 hours after induction of dauer exit. Finally, recovered worms from both species started laying eggs between 42 to 45 hours post-induction. Based on all these criteria, we consider the 12-hour post recovery induction time-point to be developmentally equivalent and comparable across the two species.

The number of genes, which are differentially expressed in the dauer to dauer exit transition (Table 1), is much larger in *P. pacificus* (4942 genes, Additional file 2: Table S1) than that seen in *C. elegans* (917 genes, Additional file 3: Table S2), suggesting more dynamic transcriptional changes accompanying dauer recovery in *P. pacificus*. Also, the number of up-regulated versus down-regulated genes in the dauers to dauer-exit comparison is much larger in *P. pacificus* than that in *C. elegans*, where the number of up- and down- regulated genes is essentially the same (Table 1). Thus *P. pacificus* dauers potentially require the activity of more genes to survive in their ecological niche than the dauer recovery stage.

Signatures of conservation and divergence in the transcriptomes of the two species

Using pairwise best BLAST mapping, we identified 6,126 1:1 orthologous gene pairs in the two species, which are represented on both the arrays, while the remaining 14,212 *P. pacificus* genes and 13,099 *C. elegans* genes represent unresolved homology relations and lineage-specific genes. Within the 6,126 orthologs, we find that only 184 gene pairs are expressed at significantly different levels in the dauer versus dauer-exit comparisons of both species (FDR corrected p-value < 0.05) (Figure 1a). This small overlap is nonetheless statistically significant (Fisher's exact test p-value = 0.029) indicating some "conservation" between dauer related genes in the two species.

The Pearson's correlation coefficient (r) between the fold changes of the 1:1 orthologs is r = 0.12 when calculated over all 6,126 orthologs, and increases to r = 0.29 when calculated on the 184 genes that are significantly

Table 1 Differentially expressed genes in the dauer versus dauer-exit comparison

***P. pacificus* transcriptome is relatively more dynamic in dauer to dauer-exit transition (FDR corrected p-value < = 0.05)**

	Up	Down	Total	Total genes on array
C. elegans	476	441	917	20143
P. pacificus	3545	1394	4939	20446

The Agilent arrays interrogate 20,446 genes for *P. pacificus* and 20,143 genes for *C. elegans*. We measure expression changes in the dauer to dauer-exit (12-hour time-point) transition using species-specific mix-stage sample as a common technical reference. Genes were called significantly differentially expressed (up- or down- regulated) based on a FDR corrected p-value cut-off of 0.05. *P. pacificus* transcriptome appears to be more dynamic during this transition.

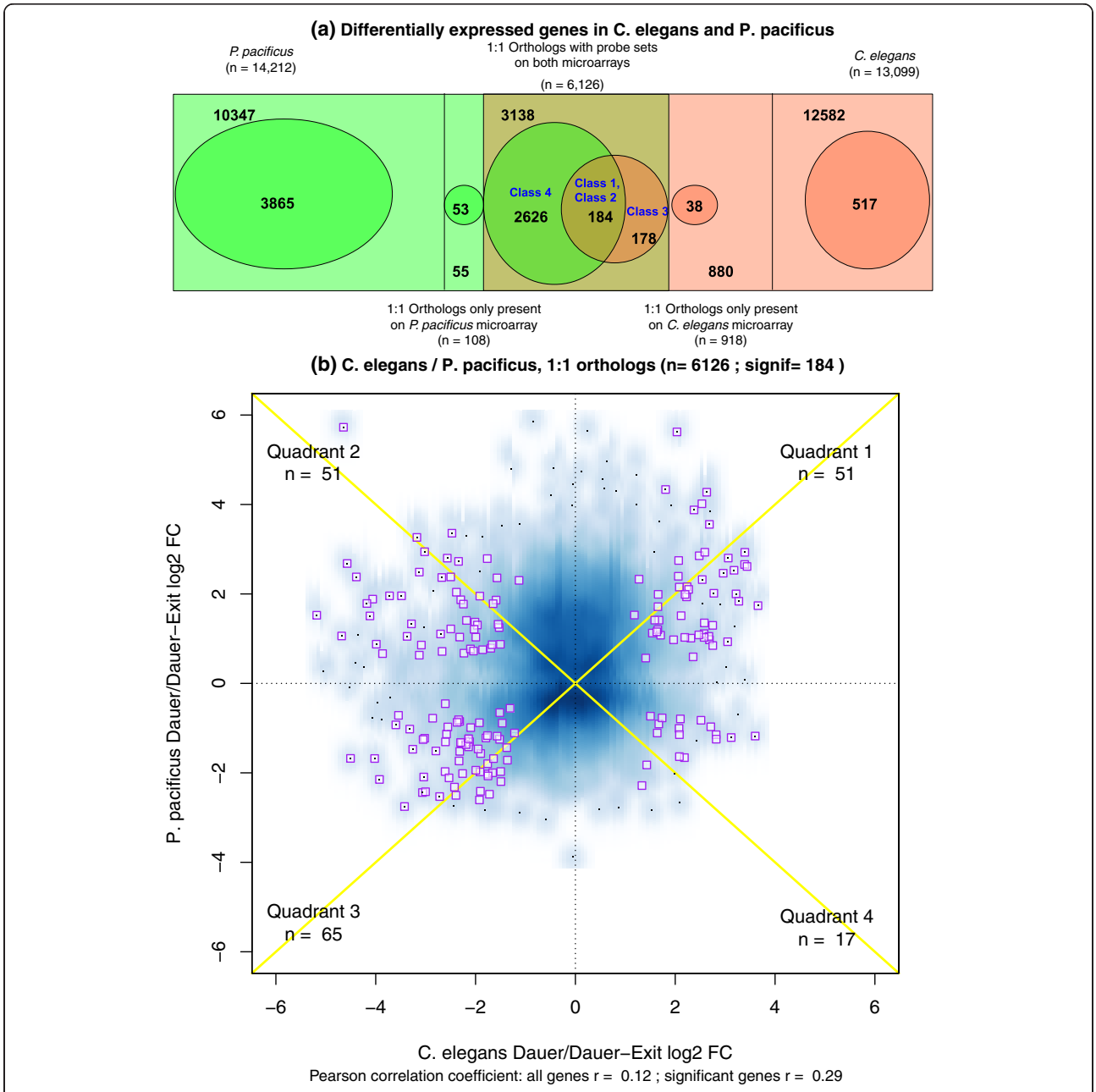


Figure 1 Limited conservation in the dauer related transcriptomes of the two species. (**a**) Overlap between genes differentially expressed in *P. pacificus* (green) and *C. elegans* dauers versus dauer-exit (12 hours) samples. The rectangular boxes represent the entire transcriptome on the array and their area of overlap represents the 1:1 orthologs between the two species, which are represented on both microarrays. The ovals represent the set of differentially expressed genes in each species. 184 orthologous gene pairs are called significantly differentially expressed in both the species, indicating limited conservation (Fisher's exact test p-value = 0.029). A substantial number of non-orthologous, species-specific genes are also differentially expressed in both the species. These gene sets are further analyzed by dividing them into distinct classes which are indicated in blue text (see main text for details) (**b**) Comparison of log2 expression fold changes for the set of 6,126 1:1 orthologs for the dauer versus dauer-exit comparison. Pearson's correlation coefficient over the entire set of 6,126 orthologs is r = 0.12, and increases to r = 0.29 for the 184 orthologs, which are significantly differentially expressed in both species (FDR corrected p-value < =0 0.05; purple boxes mark the significant orthologs, number per quadrant I - IV: 51 + 51 + 65 + 17 = 184).

differential in both species (Figure 1b). To further characterize the cross-species similarities and differences in transcriptomes, we group the set of 1:1 orthologs into the following classes:

(1) Class 1 = common orthologs that show a concordant pattern of differential regulation between the two species (up-regulated in both species, n = 51, quadrant 1, and down-regulated in

both species n = 65, quadrant 3, Figure 1b. Total = 116 out of 184 common orthologs). These are expected to be the most conserved part of the transcriptomes across species (r = 0.88)

(2) Class 2 = orthologous genes that show a negative correlation in their direction of fold-change (n = 51, quadrant 2, and n = 17, quadrant 4, Figure 1b. Total = 68 out of total 184 common orthologs). This class of genes exhibits the most divergent expression pattern in the two transcriptomes (r = –0.80) and is comparable in size to class 1.

(3) Class 3 = Orthologs differentially expressed only in *C. elegans* dauers (n = 178 in *C. elegans*, Figure 1a)

(4) Class 4 = Orthologs differentially expressed only in *P. pacificus* expression profile (n = 2626 in *P. pacificus*, Figure 1a).

Gene ontology term enrichment

To better understand which functions are enriched in these gene sets, we made use of Gene Ontology (GO) term annotations from WormBase [36]. The set of 6,126 1:1 orthologs was used as the background set, and GO annotations were assigned to *P. pacificus* genes by directly mapping them from the corresponding *C. elegans* 1:1 orthologs. Enrichment statistics were calculated with the Bioconductor package topGO [37] (see Methods). We found the biological process of "neuropeptide-signaling pathway" to be the most significantly enriched (Table *2*) in the class 1 gene set, which subsumes the most conserved expression pattern across the two species. We find six out of the 16 genes belonging to this GO class to be upregulated in dauers of both *C. elegans* and *P. pacificus*, and four of them are members of "FMRF-Like Peptide" family of neuropeptides, which also have a known role in regulating pharyngeal pumping [38]. GO enrichment analysis of class 2 genes that show an opposite fold change in their expression pattern between species yields "molting cycle, collagen and cuticulin-based cuticle" as the most significantly enriched term (Table *3*). This suggests that the dauers of both species possibly use different genetic components for synthesis and/or shedding of the cuticle. This potential difference in molting related cuticle processing in the two species is further supported by the observation that this GO term is also the most significantly enriched in Class 3 genes - the set of orthologous genes differentially regulated exclusively in *C. elegans* (Table *4*). For class 4 genes, the most significantly enriched GO terms

Table 2 GO terms enriched in orthologs with same direction of fold change in both species

GO:BP terms enriched in orthologs regulated in same direction in both species					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0007218	neuropeptide signaling pathway	16	6	0.29	2.00E-07
GO:0055114	oxidation-reduction process	219	10	3.97	0.0056
GO:0006576	cellular biogenic amine metabolism	7	2	0.13	0.0064
GO:0006508	proteolysis	197	9	3.57	0.0086
GO:0046496	nicotinamide nucleotide metabolism	10	2	0.18	0.0133
GO:0009069	serine family amino acid metabolism	10	2	0.18	0.0133
GO:0019752	carboxylic acid metabolism	128	8	2.32	0.0172
GO:0009309	amine biosynthetic process	33	3	0.6	0.0212
GO:0040019	positive regulation of embryonic development	13	2	0.24	0.0222
GO:0008202	steroid metabolic process	14	2	0.25	0.0256
GO:0008033	tRNA processing	19	2	0.34	0.0455
GO:0006006	glucose metabolic process	20	2	0.36	0.0499
GO:MF terms enriched in orthologs regulated in same direction in both species					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0016616	oxidoreductase activity; acting on the CH-OH group of donors, NAD or NADP as acceptor	34	4	0.64	0.0034
GO:0016491	oxidoreductase activity	283	14	5.29	0.0107
GO:0016747	transferase activity; transferring acyl groups other than amino-acyl groups	53	4	0.99	0.0164
GO:0016831	carboxy-lyase activity	11	2	0.21	0.017
GO:0004175	endopeptidase activity	115	6	2.15	0.0193
GO:0008233	peptidase activity	180	10	3.37	0.0258
GO:0051287	NAD binding	14	2	0.26	0.0271

Table 3 GO terms enriched in orthologs with opposite direction of fold change in both species

GO:BP terms enriched in orthologs regulated in opposite direction in both species					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0018996	molting cycle; collagen and cuticulin-based cuticle	157	7	1.77	0.0016
GO:0010171	body morphogenesis	322	10	3.63	0.0025
GO:0040011	locomotion	789	17	8.89	0.0038
GO:0040010	positive regulation of growth rate	1024	18	11.54	0.0244
GO:0040018	positive regulation of multicellular organism growth	167	5	1.88	0.0382
GO:0019318	hexose metabolic process	28	2	0.32	0.039
GO:MF terms enriched in orthologs regulated in opposite direction in both species					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0042302	structural constituent of cuticle	21	4	0.2	3.80E-05
GO:0015078	hydrogen ion transmembrane transporter activity	30	2	0.29	0.033
GO:0050662	coenzyme binding	83	3	0.8	0.045
GO:0046912	transferase activity; transferring acylgroups, acyl groups converted into alkyl on transfer	5	1	0.05	0.047

are mostly development related but are too broad to point to any specific function (Table *5*).

Expression cluster analysis

Apart from analyzing enrichment for GO annotations, we also performed a functional analysis based on overlap of our data with gene expression clusters in *C. elegans*, which originate from other published microarray data sets. To this end, we retrieved predefined expression clusters from WormBase [36] and used them for a meta-level analysis. These expression clusters represent sets of genes reported to be co-expressed under various conditions. We restricted the analysis to the set of 1:1 orthologs, and assigned expression clusters to *P. pacificus* genes by

Table 4 GO terms enriched in orthologs, which are only differentially expressed in *C. elegans*

GO:BP terms enriched in orthologs differentially expressed exclusively in C. elegans					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0018996	molting cycle; collagen and cuticulin-based cuticle.	157	17	4.19	5.70E-07
GO:0040011	locomotion	789	34	21.07	0.0019
GO:0040018	positive regulation of multicellular organism growth	167	11	4.46	0.0046
GO:0006560	proline metabolic process	5	2	0.13	0.0067
GO:0010171	body morphogenesis	322	16	8.6	0.0106
GO:0009084	glutamine family amino acid biosynthesis	8	2	0.21	0.0178
GO:0006694	steroid biosynthetic process	10	2	0.27	0.0277
GO:0055114	oxidation-reduction process	219	11	5.85	0.0306
GO:0006508	proteolysis	197	10	5.26	0.0362
GO:0018991	oviposition	145	8	3.87	0.0387
GO:MF terms enriched in orthologs differentially expressed exclusively in C. elegans					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0042302	structural constituent of cuticle	21	10	0.53	1.60E-11
GO:0004222	metalloendopeptidase activity	50	6	1.25	0.0014
GO:0016776	phosphotransferase activity; phosphate group as acceptor	5	2	0.13	0.0059
GO:0016903	oxidoreductase activity; acting on the aldehyde or oxo group of donors	8	2	0.2	0.0157
GO:0003854	3-beta-hydroxy-delta5-steroid dehydrogenase activity	8	2	0.2	0.0157
GO:0019205	nucleobase; nucleoside; nucleotide kinase activity	8	2	0.2	0.0157
GO:0020037	heme binding	51	4	1.28	0.0374
GO:0004601	peroxidase activity	14	2	0.35	0.0464

Table 5 GO terms enriched in orthologs, which are only differentially expressed in *P. pacificus*

GO:BP terms enriched in orthologs differentially expressed exclusively in P. pacificus					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0006898	receptor-mediated endocytosis	433	231	192.89	6.30E-05
GO:0009792	embryo development ending in birth or egg hatching	1535	740	683.81	0.00015
GO:0002119	nematode larval development	1123	549	500.27	0.00034
GO:0040020	regulation of meiosis	45	30	20.05	0.00219
GO:0042127	regulation of cell proliferation	45	30	20.05	0.00219
GO:0006412	translation	165	93	73.5	0.00691
GO:0016246	RNA interference	63	38	28.07	0.00812
GO:0055114	oxidation-reduction process	219	115	97.56	0.00912
GO:0000003	reproduction	1235	585	550.17	0.00934
GO:0040007	growth	1481	694	659.75	0.01358
GO:0009396	folic acid-containing compound biosynthetic process	5	5	2.23	0.01749
GO:0009067	aspartate family amino acid biosynthesis	10	8	4.45	0.02546
GO:0009072	aromatic amino acid family metabolic process	10	8	4.45	0.02546
GO:0006732	coenzyme metabolic process	41	27	18.26	0.0326
GO:0051603	proteolysis involved in cellularprotein catabolic process	36	24	16.04	0.03319
GO:0006396	RNA processing	72	40	32.07	0.03822
GO:0006511	ubiquitin-dependent protein catabolic process	29	18	12.92	0.04326
GO:0006418	tRNA aminoacylation for protein translation	33	20	14.7	0.04611
GO:0015684	ferrous iron transport	9	7	4.01	0.04679
GO:0042026	protein refolding	9	7	4.01	0.04679
GO:0006399	tRNA metabolic process	51	32	22.72	0.04805
GO:MF terms enriched in orthologs differentially expressed exclusively in P. pacificus					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**p-Value**
GO:0004298	threonine-type endopeptidase activity	13	12	5.76	4.30E-04
GO:0003899	DNA-directed RNA polymerase activity	22	17	9.74	0.00168
GO:0016491	oxidoreductase activity	283	149	125.32	0.00192
GO:0008026	ATP-dependent helicase activity	49	32	21.7	0.00229
GO:0050662	coenzyme binding	83	48	36.76	0.00828
GO:0003735	structural constituent of ribosome	93	53	41.18	0.00847
GO:0008168	methyltransferase activity	67	39	29.67	0.01438
GO:0016884	carbon-nitrogen ligase activity; with glutamine as amido-N-donor	10	8	4.43	0.02443
GO:0051082	unfolded protein binding	28	18	12.4	0.02591
GO:0030170	pyridoxal phosphate binding	34	21	15.06	0.02977
GO:0004812	aminoacyl-tRNA ligase activity	34	21	15.06	0.02977
GO:0003993	acid phosphatase activity	12	9	5.31	0.03154
GO:0031072	heat shock protein binding	23	15	10.19	0.03483
GO:0004190	aspartic-type endopeptidase activity	9	7	3.99	0.04517
GO:0015093	ferrous iron transmembrane transporter activity	9	7	3.99	0.04517
GO:0008483	transaminase activity	9	7	3.99	0.04517

mapping annotation from the corresponding *C. elegans* orthologs. The significance scores for enrichment of each pre-defined cluster was calculated separately for the significantly up- and down-regulated genes from the dauer versus dauer exit comparison in our data (see Methods for details), and are summarized in Table 6.

These significantly enriched clusters enable us not only to identify system-wide trends in our data but

also validate them against existing data-sets. For example, we see a highly significant overlap between our *C. elegans* data and previously reported expression profiles of dauers and dauer-exit stages [5,9]. Specifically, the cluster "Dauer enriched" identified by Wang and Kim in their time-course study [5] is enriched in genes called up-regulated in our data (Table 6, row 1, column "cel_dauer"). Furthermore, other clusters from later dauer-exit time points in the same study show a significant overlap with genes called over-expressed in our dauer-exit samples (Table 6, rows 2 to 4, column "cel_exit"), thus indicating good agreement between the two data-sets. Interestingly, the clusters corresponding to "early" and "climbing" genes in *C. elegans* (Table 6, rows 2 and 3) also show a significant overlap with genes over-expressed in our *P. pacificus* dauer-exit samples (column "ppa_exit" in Table 6), pointing towards a conservation of a part of the active transcriptome during dauer recovery in the two species. A similar trend is also seen for clusters obtained from a microarray study of dauers from daf-c/TGF-beta mutants [9] (Table 6, rows 5 and 6). We also see a significant overlap between genes regulated in response to heat shock, oxidative stress, osmotic stress and response to pathogens, highlighting the fact that activation of stress response pathways is a common feature of dauer stage in both *C. elegans* and *P. pacificus*. Further, genes activated in response to starvation and on resumption of feeding are also found to be enriched in dauer and dauer-exit stages respectively for both the species, in good agreement with non-feeding status of dauers and resumption of a feeding program during dauer-exit. We observe a significant overlap between the expression clusters that are enriched in the dauers of both the species (Fisher's exact test p-value = 0.0185), indicating a conserved signature of gene expression. This overlap of expression clusters is stronger for the dauer-exit stages of both the species (Fisher's exact test p-value = 6.833E-06), indicating that similar transcriptional programs are activated during dauer recovery in both the species. In summary, in spite of the low overlap and correlation between differentially expressed genes in the two species, functional analysis based on GO and expression cluster enrichment could identify conserved aspects of the transcriptional program associated with the dauer recovery process in the two species.

Apart from the differentially expressed genes in the limited set of 1:1 orthologous genes, there is a comparable number of differentially expressed genes for which either no orthologs exist in the other species or the orthology cannot be uniquely resolved (2,129 *P. pacificus* specific genes versus 2626 + 184 = 2810 1:1 orthologs in *P. pacificus*, 555 *C. elegans* specific genes versus 178 +184 = 362 1:1 orthologs in *C. elegans* data, Figure 1a). This indicates that the non-conserved part of the active transcriptome show a similar dynamic like the conserved part and might play a significant role in the biology of dauer larvae of the two species. Further functional studies would be required to understand the adaptive significance of these species-specific genes.

PFAM protein domain based analysis identifies potentially conserved and diverged functional gene classes

We further investigated the functional differences in the two transcriptomes by stratifying gene expression based on their protein domain annotations. This approach has the benefit of not being dependent on orthology relationships (like some of the analyses performed above), and can identify conserved functional signatures if the same gene function is carried out by different paralogs in the two species. For this analysis, we inferred possible gene functions by annotating 13,344 gene loci in *C. elegans* and 14,018 gene loci in *P. pacificus* with PFAM protein domains (p-value of domain match < 0.001). We then stratified the gene sets from the two species into gene families by PFAM domains with at least five members on each microarray. These 441 gene families may partially overlap and are represented by their median log2 fold expression changes. The Pearson's correlation coefficient of log2 fold changes is r = 0.52 for the dauer versus dauer exit comparison (Figure 2), an improvement over the aforementioned correlation values for 1:1 orthologs (Figure 1b). Some protein domains indeed show concordant expression in both the species, such as the HSP20 family of heat shock proteins that is induced in dauers of both the species (Figure 2) and potentially confers stress resistance. Similarly, the Peptidase_S28 family of proteins is repressed in the dauer stage of both species possibly because it contains proteins with lysosomal activity, which might not be required during the repressed metabolic state of dauers [7,39]. Interestingly, we also find some protein domains with totally opposite expression patterns, such as some of the proteasome subunit domains. The genes with domains "Proteasome" and "Proteasome_A_N" are induced strongly in *P. pacificus* dauers but are down regulated in *C. elegans* dauers. Since proteasome function has been implicated in regulation of longevity [40-42], up regulation of these specific proteasome domains in *P. pacificus* might be correlated with the increased longevity of *P. pacificus* dauers versus *C. elegans* dauers [43]. Thus, in summary, PFAM annotation based analysis of the two transcriptomes also identifies some conserved functional signatures as well as some divergent signatures. This

Table 6 Expression cluster enrichment analysis

Expression clusters over-represented in "dauer enriched" and "exit enriched genes" in C. elegans and P. pacificus

Expression clusters related to dauer larvae, stress response and starvation / feeding

Sl.No.	Expression Cluster	cel_dauer	cel_exit	ppa_dauer	ppa_exit
1	Wang_Kim_WBPaper00005859_DauerEnriched	10.45	0	0	0
2	Wang_Kim_WBPaper00005859_EarlyGenes	0	5.99	0	10.67
3	Wang_Kim_WBPaper00005859_ClimbingGenes	0	14.58	0	4.62
4	Wang_Kim_WBPaper00005859_LateGenes	0	63.09	0	0
5	WBPaper00024393:strongly_regulated_dauer_genes_UP	1.39	0	0	0
6	WBPaper00024393:strongly_regulated_dauer_genes_DOWN	0	8.49	0	10.85
7	WBPaper00034757:up_by_oxidative_stress	4.2	0	0	0
8	WBPaper00034757:down_by_oxidative_stress	0	6.47	0	6.82
9	WBPaper00035227:heat_shock_regulated	9.74	0	0	0
10	WBPaper00035873:dpy-10_regulated	0	25.06	0	3.89
11	WBPaper00035873:dpy-9_regulated	0	2.02	0	1.57
12	WBPaper00035873:osm-11_regulated	0	2.02	0	1.57
13	WBPaper00035873:osm-7_regulated	0	2.4	0	2.07
14	WBPaper00035873:osm-8_regulated	0	42.83	0	4.45
15	WBPaper00035873:osmotically_regulated	0	0	0	4.65
16	WBPaper00032948:StarveUp2	6.17	0	2.5	0
17	WBPaper00032948:StarveUp3	7.15	0	0	0
18	WBPaper00032948:StarveUp4	10.45	0	0	0
19	WBPaper00032948:FedUp	0	78.95	0	5.12
20	WBPaper00032948:MoltOssilate	0	74.78	0	0
21	WBPaper00032062:age_regulated_genes	0	5.37	0	4.75

Expression clusters related to to pathogen response, RNAi machinery etc.

Sl.No.	Expression Cluster	cel_dauer	cel_exit	ppa_dauer	ppa_exit
22	WBPaper00028482:PA14_upregulate	0	0	0	2.03
23	WBPaper00028789:PA14_vs_gacA_downregulated_4hr	0	5.9	0	0
24	WBPaper00028789:PA14_vs_gacA_downregulated_8hr	0	6.7	0	0
25	WBPaper00028789:PA14_vs_OP50_downregulated_8hr	0	6.85	0	1.81
26	WBPaper00028789:PA14_vs_OP50_upregulated_4hr	0	0	0	1.4
27	WBPaper00028789:PA14_vs_OP50_upregulated_8hr	0	0	0	2.02
28	WBPaper00030985:Enterococcus_faecalis_upregulated	0	0	0	1.6
29	WBPaper00028789:pmk-1_downregulated	0	0	0	1.46
30	WBPaper00029437:dcr-1_upregulated	3.13	0	0	0
31	WBPaper00029437:rde-4_upregulated	0	0	0	1.69
32	WBPaper00027111:eri-1(mg366)_downregulated	0	2.23	0	0
33	WBPaper00027111:rde-3(r459)_upregulated	0	0	0	1.52
34	WBPaper00035892:KIM5_regulated	0	0	0	4.79
35	WBPaper00035892:KIM5_vs_OP50_Up	0	0	0	2.36

Tissue specific, and other expression clusters

Sl.No.	Expression Cluster	cel_dauer	cel_exit	ppa_dauer	ppa_exit
36	WBPaper00030839:Embryo_Pan_Neuronal	0	0	6.67	0
37	WBPaper00030839:Larval_Pan_Neuronal	4.75	0	2.71	0
38	WBPaper00031003:0hr_muscle_depleted	0	2.99	0	0

Table 6 Expression cluster enrichment analysis *(Continued)*

39	WBPaper00031003:24hr_muscle_depleted	0	2.11	0	2.73
40	WBPaper00031003:total_muscle_depleted	0	2.29	0	0
41	WBPaper00031003:total_muscle_enriched	0	0	2.42	0
42	WBPaper00031532:Larva_Pan_Neuronal_Depleted	0	28.03	0	11.41
43	WBPaper00031532:Larva_Pan_Neuronal_Enriched	2.48	0	4.69	0
44	WBPaper00026980:intestine_enriched	0	0	3.46	1.5
45	WBPaper00024671:AFD_AWB_vs_unsorted_downregulated	0	3.02	0	3.59
46	WBPaper00031832:slr-2_regulated	0	0	0	2.75
47	WBPaper00033101:spr-5_regulated	0	6.44	0	0
48	WBPaper00034739:N2lessDR1350	0	39.13	0	0
49	WBPaper00034739:RIL17lessRIL14	0	18.11	0	0
50	WBPaper00035905:FBF-1_Associated	0	0	5.54	0
51	WBPaper00037611:RNP-8-associated	0	0	4.92	0
52	WBPaper00025032:PAL-1_target_genes	0	2.58	0	0

This table lists - log10 transformed p-values of tests for enrichment of dauer and dauer-exit genes from both species for 52 out of 169 expression clusters. These 52 expression clusters passed the FDR corrected p-value threshold of 0.05 for at least one condition. The overlap between dauer enriched clusters and dauer-exit enriched clusters is statistically significant (Fisher's exact test p-values = 0.0185 and 6.833E-06 respectively).

analysis also provides an important catalog of gene function, and the role of genes corresponding to the protein families identified here can be studied in more detail in *P. pacificus* dauers in the future.

Metabolic recovery during dauer exit displays different expression patterns of key enzymes in *P. Pacificus*

Dauer larvae of *C. elegans* are known to undergo a remarkable shift in their intermediary metabolism to

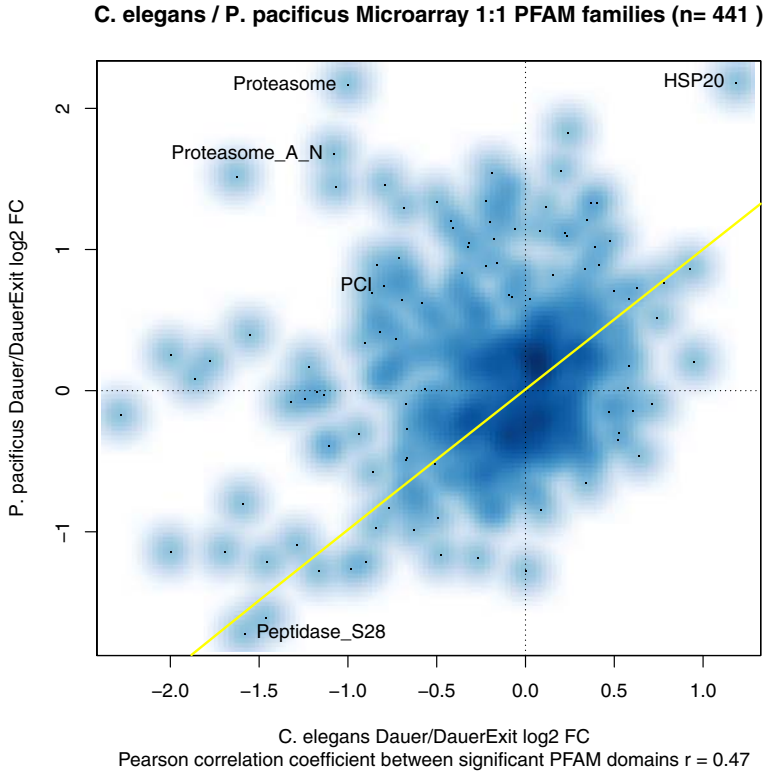


Figure 2 Log2 fold changes of median PFAM gene expression levels in *P. pacificus* and *C. elegans*. All genes on the microarrays were stratified by PFAM domain annotation into partially overlapping gene families with at least 5 members. A positive log2 fold change indicates that the corresponding genes were up-regulated in the dauer stages in comparison to the dauer exit time point.

survive long periods without actively feeding [7,39]. We therefore looked in more detail into expression changes in metabolism related genes using the KEGG database [44]. Gene-pathway assignments were obtained for *C. elegans* from KEGG and transferred to *P. pacificus* via the 6,126 orthologs identified through best pairwise BLAST mapping. For each KEGG pathway, we determined the number of genes from the 1:1 orthologous set that are upregulated, downregulated or show no differences in expression levels in the dauer versus dauer-exit comparisons in both species. We observed that the three key processes of central carbon metabolism, namely glycolysis, the Krebs (TCA) cycle and oxidative phosphorylation, show entirely different dynamics during the dauer-exit time course in the two species (see Figure 3). For *C. elegans*, mRNA abundance levels for most of the genes in the three pathways are found to be similar in the dauers and dauer-exit stage (category "no_change", Figure 3). Only a few genes are downregulated, while none of the genes is upregulated in *C.elegans*. In contrast, we see a more dynamic regulation of metabolic pathways in *P. pacificus*, with comparatively larger number of downregulated genes as well as a few upregulated genes. These differences in metabolic recovery in *P. pacificus* could have an obvious and trivial explanation in that *P. pacificus* might have a different developmental rate compared to *C. elegans*. However, as described above, several morphological and developmental traits suggest that the course of post-dauer development in *P. pacificus* and *C. elegans* is very similar. Hence, we hypothesize that metabolic regulation in *P. pacificus* dauers might itself be inherently different from that in *C. elegans*. This would also be consistent with the dauer recovery studies showing that *P. pacificus* dauers can survive much longer than *C. elegans* dauers [43], one of the possible reason being their ability to metabolize stored fats at a different rate and/or in a different way. Our genome-wide expression studies reflect this potential difference as divergence in pathway expression profiles between the two species. Since gene expression profiles are dynamic and can be sensitive to differences in developmental timing (e.g. [11]), more fine-grained and detailed functional studies will be needed to verify whether these differences are really due

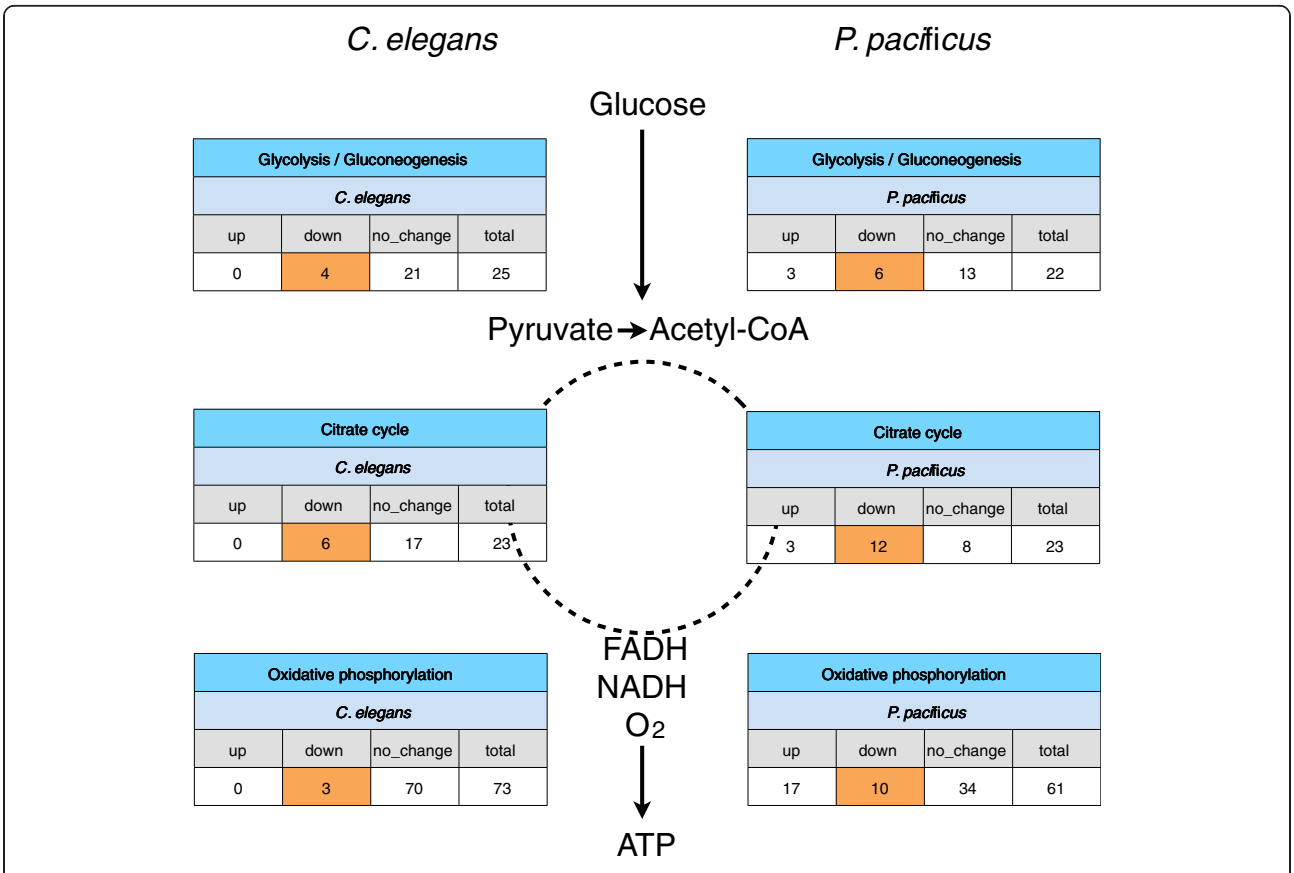


Figure 3 Differences in regulation of metabolism related genes during the dauer exit time course in *C. elegans* and *P. pacificus*. Gene-to-pathway assignments were retrieved for *C. elegans* from the KEGG database. Pathway assignments were transferred to *P. pacificus* by mapping 1:1 orthologs. All tables show genes that are up, down or at a similar expression level in a dauer vs. dauer exit (12 hours post induction) comparison.

to inherent metabolic differences or an artifact of slight variations in developmental timing.

Genes acquired by horizontal gene transfer and orphan genes in *P. Pacificus* are developmentally regulated

Previous studies have identified genes in *P. pacificus* that likely originate from lateral gene transfer events. For example, *P. pacificus* is the first nematode, which is not a plant parasitic species yet contains cellulase genes [25]. *P. pacificus* cellulase genes are most similar to genes found in the social amoeba of the Dictyostelid group [45]. These cellulases are fully functional and have been maintained over long evolutionary time periods [46]. We have identified seven cellulase genes in the current transcriptome annotation, six of which are represented on our *P. pacificus* microarrays. In a dauer versus mixed-stage comparison, four of these cellulases are significantly differentially expressed, all of them being down regulated in dauers (or equivalently, expressed predominantly in non-dauer mix-stages, Additional file 4: Table S3). This is in agreement with their hypothesized function of enabling feeding on novel food sources [46]. For example, one gene (Contig66-snapTAU.4) is already significantly up-regulated as early as 12 hours post induction (Additional file 4: Table S3, column logFC(D0/D12)).

Another interesting gene family, which was most likely acquired from beetles, is the diapausins. These genes encode small peptides, which are thought to protect the dormant beetle against microbial infections in diapause. Diapausins provide antifungal activity by acting as Ca2+ channel blockers [47]. Out of five members of this gene family in *P. pacificus* genome, three are represented on our microarrays. Two of them are significantly up-regulated in dauers compared to dauer exit (Additional file 4: Table S4, column logFC(D0/D12)) and might be important for antifungal immunity in the dauer stage on the beetle, while the third gene does not respond to the dauer exit program. This gene is in fact downregulated in a dauer versus mix-stage comparison (column logFC (D0/MixStage), Additional file 4: Table S4) suggesting potential anti-fungal role in other developmental stages. Thus, while previous studies in our lab have demonstrated that both gene families (cellulases and diapausins) originate from lateral gene transfer events, we are now able to show members of both gene families are developmentally regulated in agreement with their proposed functions: cellulase genes are implicated in feeding and are hence downregulated in dauers, and at least two diapausin gene appears to have a role in protecting hibernating stages. These hypotheses generated from our microarray data can further be investigated at a functional level.

One third of all *P. pacificus* genes are pioneer genes, which do not show any obvious sequence similarity to other organisms [26]. On average, pioneer genes are significantly higher expressed in the dauer stage in comparison to the remaining genes (p-value < 1.0E-06; Additional file 1: Figure S3). We speculate that pioneer genes could be especially important to the evolution of the dauer stage in *P. pacificus,* which is essential for its necromenic life style.

Discussion

The dauer is an ecologically regulated developmental stage observed in many free-living nematode species and hence provides an interesting model to investigate how ecological adaptations are integrated into developmental pathways during evolution [1]. Extensive studies in *C. elegans* have uncovered the genetic regulators of dauer formation and comparative studies from other free living nematodes have begun to provide insights into evolution of dauer regulatory genes and pathways [33,34,48]. *P. pacificus* presents an ideal model for comparing dauer larvae of free living nematodes. In this study we provide a comprehensive comparison of gene expression in the dauer stage and dauer exit (12 hour post induction) of the two nematode model systems *P. pacificus* and *C. elegans* using the Agilent microarray platform. While we are aware of the available *C. elegans* expression data on dauer development (e.g. [5,6,8-14], we felt that it was necessary to generate *C. elegans* data *de novo* for two reasons. First, we wanted to use the same platform as that used for the analysis of *P. pacificus* in order to enhance the power of a direct comparison. Second, we wanted to benefit from technical advances in custom and long nucleotide microarrays. Based on cross-species comparison of transcriptomes of the ecologically important dauer stage, we draw four major conclusions from our studies.

First, we provide a list of similarities and differences between *P. pacificus* and *C. elegans*, which shows an unexpected level of divergence at transcriptome level, even though the dauer stage and recovery process appear to be developmentally conserved. While this comparison allows several evolutionary conclusions (see below), the *P. pacificus* data set on its own can be used as a starting point for a functional analysis of the dauer stage and dauer exit. This data set represents an invaluable resource given the importance of the dauer stage for the ecology of this nematode. The association of *P. pacificus* with scarab beetles is restricted to the dauer stage as long as the beetle is alive [27,28]. Only after the beetle´s death, the nematode exits from the dauer stage to feed on the microbes, which develop on the carcass of the insect. Thus, the *P. pacificus* dauer stage has a well-defined ecological niche and the expression profiles described in this study will serve as an entry point to future functional studies. For example, our transcriptomic data identifies many *P. pacificus* specific genes as

upregulated in the dauer stage, which implies a potential function in adaptations enabling survival on beetles.

Second, we show that metabolic differences exist between both species, with different patterns of regulation of genes involved in the central carbon metabolism pathways (glycolysis, tricarboxylic acid cycle and oxidative phosphorylation) of *P. pacificus* upon dauer exit. This difference does not result in obvious phenotypic changes during dauer exit, but may be linked to the differences in life history traits. This is especially relevant given that *P. pacificus* is adapted for longevity [43]. Under experimental conditions, *P. pacificus* survives for up to one year in the dauer stage, whereas *C. elegans* N2 dauer larvae die after approximately 22 weeks. Further studies will reveal how much of the observed metabolic differences are explained by different life-history traits versus differences in rate of development.

Third, this study provides the first nematode evo-devo comparison looking at the downstream consequences of homologous developmental processes between species belonging to different nematode genera. While detailed studies between nematodes as distinct as *C. elegans* and *P. pacificus* have investigated the regulation of vulva and gonad development, sex determination and dauer formation [30,34,49,50], most of these studies are concerned with the regulatory mechanisms rather than the "executional programs" of the corresponding developmental processes. The divergence in the expression profiles of *C. elegans* and *P. pacificus* adds an important new finding to the growing literature of evo-devo. Previous studies have indicated the limited conservation in the genetic and molecular control of developmental processes in *P. pacificus* and *C. elegans*. For example, vulva induction relies on different signaling pathways, requires a novel regulatory linkage and the acquisition of novel protein domains in *P. pacificus* Wnt signaling [32]. This type of result has been discussed as an example for the theory of developmental systems drift, which proposes that conserved developmental and morphological structures can be regulated by largely diverse regulatory mechanisms [51]. Considering that gene regulatory networks are hierarchically structured, with possibly different rates of evolution at the top level regulatory genes and the most downstream level of effector genes [52], it can be argued that in principle, developmental systems may diverge due to differences/drift at any of these levels. Unfortunately, unlike the top-level regulatory network, the downstream effector programs of vulva development have largely escaped identification by developmental genetic approaches and have not been easily accessible to transcriptome studies as they are single-cell or small group of cell responses. Our study circumvents this limitation because dauer formation is a "whole body response" of the organism to harsh environmental conditions.

The dauer context is also interesting because in *P. pacificus*, the key transcription factors of the dauer regulatory network, DAF-16 and DAF-12, are conserved [33,34]. However, this in itself does not indicate the extent to which the downstream targets of the regulatory network are subject to evolutionary change. Herein, we could demonstrate for the first time that the core downstream execution program of a developmental stage can differ tremendously between *P. pacificus* and *C. elegans,* in spite of conservation of upstream regulators like DAF-16 and DAF-12. Thus, these observations make a case for extending the concept of developmental systems drift to the downstream molecular execution of specific developmental stages.

The fourth conclusion is also related to evolutionary theory. While future studies will have to reveal how much of the observed differences between *P. pacificus* and *C. elegans* is really of functional importance, it has often been assumed that such differences might simply be neutral [23,53]. Gene expression, which is neither strongly deleterious nor advantageous, previously termed "gratuitous expression" [53], will not be under selection and will be free to evolve by drift. Consequently, such expression is probably not functional. However, such arguments may not apply to the dauer stage since nematodes live off their internal limited energy resources and any random or neutral transcriptional activity would diminish these limited resources.

Comparative studies in developmental genetics have driven the studies on evolution of developmental mechanisms, and with whole genome sequencing of many animal species has now highlighted new facets of evolutionary dynamics through comparative genomic studies [21]. Since transcriptional regulation is a key building block in the genotype to phenotype translation, comparative transcriptomic studies add another dimension to the analysis of evolutionary processes [23]. Our work contributes to the growing set of results from comparative transcriptomics in diverse developmental systems (e.g. [54-58]). These studies together span a range of conclusions, from high transcriptomic conservation at one end, to relatively low conservation in others, suggesting inherent constraints as well as flexibility in the evolution of gene regulatory networks [23]. Future studies comparing transcriptomes of homologous biological processes in related species, will be important for understanding the role of transcriptome evolution in generating animal diversity. Ultimately, this will also reveal the extent to which the conserved or divergent expression changes are subject to adaptive and non-adaptive forces during evolution.

Methods

Worm strains and culture

We used wild-type strains of two distinct nematode species in all of our experiments. For *Caenorhabditis elegans*, we used the N2 (Bristol) strain. For *Pristionchus pacificus*, we used the RS2333 strain (formerly known as PS312). For mixed stage cultures, 10 to 15 early adults were spotted on 10 cm NGM plates and allowed to grow at 20°C for 5 days, and washed off with M9 for RNA extractions. The dauers for both species were obtained from liquid cultures grown at 25°C. For this, 15 to 20 mixed stage plates (see above) were washed off and suspended in a final volume of 500 ml of S-medium in a 3000 ml Erlenmayer flask. 200ul Nystatin (50 mg/ml stock in DMSO) and 500ul Kanamycin (20 mg/ml stock solution) was added to prevent fungal and bacterial contamination. On day 1, 5 and 8 of culture, 12.5 ml of OP50 (20% w/v stock in S-medium) was added as food source. The cultures were grown for 12 to 14 days in a shaker incubator @ 220 rpm, 25°C. Dauers were purified from liquid cultures between days 12 to 14. The culture was centrifuged to obtain a worm pellet, which was then incubated with 1% SDS for 30 minutes to kill any non-dauers. Live worms were separated by sucrose flotation, and a subsequent precipitation in 15% Ficoll400 resulted in a pure dauer pellet (confirmed under a stereo microscope). The washed dauers were incubated overnight in 0.1 M NaCl at 20°C, to let them recover from any possible stress induced during the harsh dauer purification process. The dauers were then collected by centrifugation, and used for RNA extraction and for starting dauer-exit cultures.

Dauer exit time course

For the dauer-exit 12-hour time-point samples, around 200 purified dauers were spotted on each 10 cm NGM plate pre-spotted with 2 ml OP50, and grown at 20°C for 12 hours. These conditions ensured abundant food supply and non-crowded conditions, favorable for inducing dauer exit [3]. Worms from 20 to 30 such plates were used for RNA extraction per sample. The 12-hour time-point was chosen for both the species to give them sufficient time to recover from dauer stage and manifest measurable transcriptional changes. The recovering worms from both species were monitored to ensure comparable developmental rates based on the following criteria: (1) Most of the population (~90%) of the recovering worms from *C. elegans* as well as *P. pacificus* resume pharyngeal pumping within 3 hours after inducing dauer exit. (2) At the 12 hour time-point (the stage of dauer-exit samples), no discernible morphological differences are found between the recovered dauers from both species. (3) The recovering animals from both the species enter a lethargus between 13 to 14 hours post-recovery [3]. (4) The worms that are allowed to develop further at 20^0 C reach the next moult (L4 for *C. elegans,* J4 for *P. pacificus*) between 22 to 23 hours after induction of dauer exit. (5) Finally, recovered worms from both species started laying eggs between 42 to 45 hours post-induction.

RNA extraction

Total RNA was isolated using TRIzol (Invitrogen) according to the manufacturer's instructions. Four biological replicates for each stage (mixed, dauers, dauer-exit at 12 hrs) were prepared from parallel cultures. The extracted RNA was purified by using phenol:chloroform: isoamyl alcohol precipitation, to remove any trace of TRIzol contamination which might interfere with subsequent reactions. The RNA pellet was suspended in RNAse free water, quantity and quality was assayed on a Nanodrop spectrophotometer, and stored at −80°C until the microarray reactions.

Microarray design and experiments

We used the Agilent *C. elegans* oligonucleotide microarrays (~ 43,000 probes for ~ 20,000 open reading frames, GEO accession: GPL10094) for all *C. elegans* experiments. For *P. pacificus*, we designed custom Agilent microarrays based on the most recent transcriptome annotation (~ 23,000 gene predictions, predicted with external evidence from 454 EST alignments). We could accommodate ~ 93,000 probe sequences on our custom *P. pacificus* microarrays (NCBI GEO platform accession GPL14372). This probe set was designed using the OligoWiz 2.0 software [59] and submitted for custom fabrication to Agilent Technologies via their eArray web tool. Out of 93,000 probes, 87,070 probes map uniquely to the latest genome assembly [26]. We restrict all subsequent analyses to a probe set of 69,916 high-confidence probes.

Microarray hybridizations were carried out in a two-color format, with 4 biological replicates per experiment including a pair of dye-flips. Two sets of hybridizations were carried out: (1) Dauers versus Mix-Stage, and (2) Dauer-Exit at 12 hr versus Mix-stage, thus making a total of 8 microarray hybridizations per species. Equal amounts of total RNA (500 ng to 1000 ng) for 4 biological replicates from each stage was used to prepare Cy5- or Cy3- labelled cRNA using Quick Amp Labelling Kit (Agilent Technologies, Inc, Santa Clara, CA, USA), as per manufacturer's instructions. Based on the starting amounts of total RNA, appropriate dilutions of positive controls (Spike Mix-A and Spike-Mix B from Agilent) had been added to the reaction-mix before the RT reaction, as per manufacturer instructions. The labelled cRNAs were hybridized either to Agilent *C. elegans* oligonucleotide microarrays or our custom *P. pacificus* microarrays from Agilent. The arrays were

scanned using GenePix 4000B Microarray Scanner, and raw data extracted using GenePix Pro sofware (version 6).

Analysis of microarray expression data

The data was analyzed using the Bioconductor [60] package limma [61]. Briefly, the array quality was assessed by checking for uniform background and foreground intensities over the entire array. The signal was background corrected using the normexp method [62]. The arrays were lowess normalized individually ("normalizeWithinArrays" option), with differential weights assigned to probes and to positive control spike-ins, which are expected to show no fold change [63]. This differential weighing of probes is particularly necessary to account for changes in the relative proportion of mRNA versus total RNA. Without this differential weighing scheme, the fold change calculations can be erroneous [64]. The weight parameters were optimized based on MA-plots such that spike-in controls show their expected fold change values. lmFit function was used to fit a linear model to probe intensities across arrays, and differential expression was calculated by empirical Bayes method using the eBayes function [65]. Control of FDR was employed as correction for multiple testing. All the data from this publication have been deposited in a MIAME compliant format [66] at NCBIs Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession numbers GSE30977 and GSE31861.

Mapping of 1:1 orthology relations

We used a pairwise best BLASTP strategy to compute 1:1 orthologs. Briefly, we ran all protein sequences from *C. elegans* as query versus the database of *P. pacificus* gene predictions and vice versa. Only hits with a BLAST score > = 50 bits were retained. We define mutually best hits as 1:1 orthologs. We identified 7,176 ortholog pairs with this methodology.

Gene ontology enrichment analysis

GO ontologies for *C. elegans* were downloaded from wormmart WS200. Gene ontologies were assigned to *P. pacificus* genes based on the previously defined 1:1 orthologs. The topGO tool [37] (version 2.4.0) was used for computing significantly enriched GO terms. We used the "GOFisher" test statistic and a p-value cutoff of 0.05. The background set was limited to the 6,126 orthologs that are represented on both *C. elegans* and *P. pacificus* microarrays. We did not apply any multiple testing correction to the reported p-values. We followed this strategy to uncover all "trends" in the data to attain a comprehensive picture of the underlying biology.

Expression cluster enrichment analysis

WormBase [36] contains information on co-expressed gene groups in *C. elegans*. The list of microarray experiments where a given *C. elegans* gene is known to be differentially expressed can be extracted from the section "Expression Cluster" from the WormBase gene summary page for each gene. We retrieved all available expression clusters for *C. elegans* genes from the WormBase web site. We inferred expression clusters for *P. pacificus* based on the set of 1:1 orthologs. Only the clusters associated with the 6,126 1:1 orthologs were used as the background set. P-values for expression cluster enrichment in dauer-enriched or dauer exit-enriched gene sets were computed with a 2x2 Fisher exact test, and a multiple testing correction to control FDR was applied.

Pfam domain annotation

We annotated the proteome of *C. elegans* and *P. pacificus* with PFAM domain matches (PFAM release V23/4 [67])). Protein HMM searches were performed with HMMer 3.0 [68] using a p-value cutoff of 0.001.

KEGG pathway analysis

We retrieved the latest gene to pathway mapping for *C. elegans* from the KEGG SOAP server (see www.genome.jp/kegg/soap for details). Again, pathway annotations were transferred to *P. pacificus* with the help of the 1:1 orthologs. Gene expression information and pathway information were overlaid to generate Figure 3.

Data access

Microarray data have been deposited in NCBIs Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession numbers GSE30977 and GSE31861. The gene set annotation for this manuscript may be obtained from http://www.pristionchus.org/download/ . The list of 1:1 orthologs and PFAM domain annotations may be also obtained from there.

Additional files

Additional file 1: Figure S1. The impact of different normalization strategies on expression fold change estimation. a) The general concept of using spike-ins to control for changes in total mRNA abundance relative to total RNA. See van de Peppel *et al.*, 2003 for a detailed explanation. b) MA-plot of a Loess based normalization assuming a net mRNA expression fold change of 0 between the two samples. The M axis depicts the log2 fold change between the red and green microarray channels. The A axis is the average log2 intensity across both microarray channels. The position of the spike-in signals (colored dots) clearly indicate a deviation from this assumption. The corresponding legend in the upper left shows the expected fold change of the spike-ins relative to total RNA levels. Consequently, the location of the spike-ins corresponds to a down-shift of mRNA abundance levels within the total RNA population. c) Differentially expressed gene counts for the dauer vs. mixed stage comparison conditional on the

normalization strategy. Figure S2. Explanation of a common reference design for microarrays (see Eisen and Brown 1999). For each species, RNA from biological replicates of mix-stage worms (M1 to M4) were combined together to generate one common reference pool. Labelled aRNA produced from independent biological replicates were then co-hybridized with labelled aRNA from the common reference pool, including two dye-swaps. Samples D1 to D4 represent Dauer samples and DE1 to DE4 represent dauer exit samples at 12 hour timepoint. The blue arrows indicate the direction of labelling in each co-hybridization (arrow head = Cy3 labelled, arrow tail = Cy5 labelled). This design allows comparison via the common reference pool, but at the same time remains flexible for adding more time-points to the study, if needed. Figure S3. Cumulative plot of log2 expression fold changes for genes with and without sequence conservation (pioneer genes). The two distributions are significantly different (two-sample Kolmogorov-Smirnov test; p-value < 10E- 16).

Additional file 2: Table S1. Differential expressed genes in the dauer versus dauer-exit comparison in *P. pacificus*.

Additional file 3: Table S2. Differential expressed genes in the dauer versus dauer-exit comparison in *C. elegans*.

Additional file 4: Table S3. Differential expression of *P. pacificus* cellulase genes that were acquired by horizontal gene transfer. Table S4. Differential expression of *P. pacificus* diapausin genes that were acquired by horizontal gene transfer.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
AS, RJS and CD designed the experiments and wrote the manuscript. AS performed the experiments. AS and CD analyzed the data. All authors read and approved the final manuscript.

Acknowledgements
The authors want to thank various members of the labs for ongoing discussion. In particular, we thank Drs. G. Markov and C. Rödelsperger for bioinformatic support and suggestions.

Received: 5 October 2011 Accepted: 12 April 2012
Published: 19 June 2012

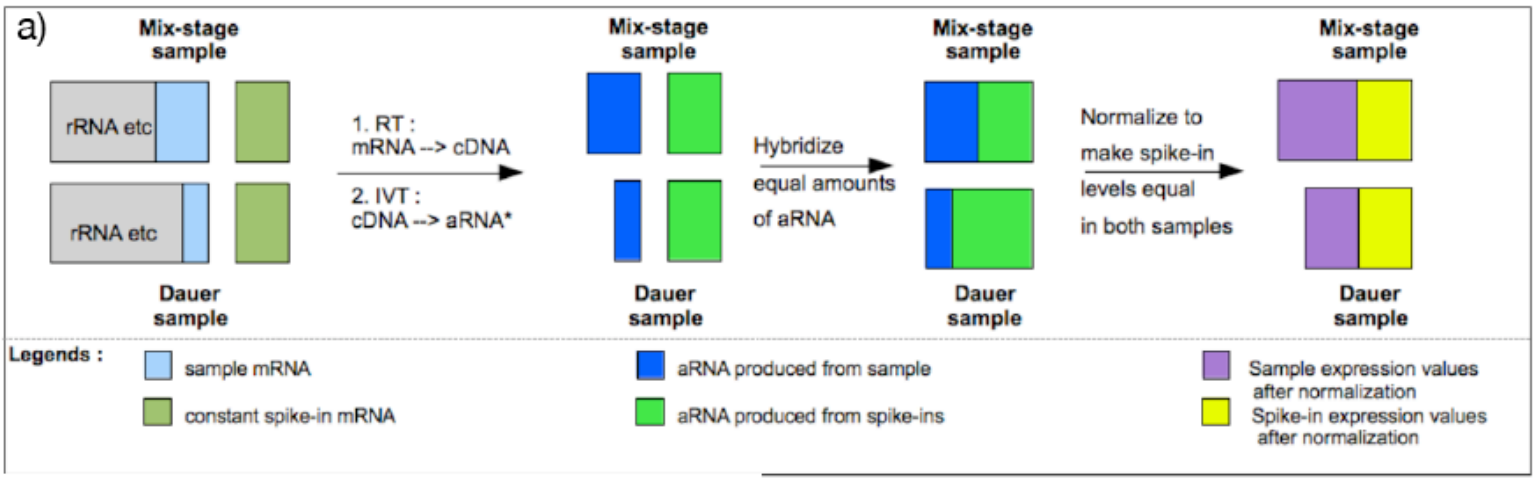
References
1. Sommer RJ, Ogawa A: **Hormone signaling and phenotypic plasticity in nematode development and evolution.** *Current Biology* 2011, **21**:R758–R766.
2. Fuchs G: **Die Naturgeschichte der Nematoden und einiger anderer Parasiten 1. des Ips typographus L. 2. des Hylobius abietis L.** *Zool Jahrb* 1915, **38**:109–222.
3. Cassada RC, Russell RL: **The dauer larva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans.** *Developmental Biology* 1975, **46**:326–342.
4. Dalley BK, Golomb M: **Gene expression in the Caenorhabditis elegans dauer larva: developmental regulation of Hsp90 and other genes.** *Dev. Biol* 1992, **151**:80–90.
5. Wang J, Kim SK: **Global analysis of dauer gene expression in Caenorhabditis elegans.** *Development* 2003, **130**:1621–34.
6. Jones SJ, Riddle DL, Pouzyrev AT, Velculescu VE, Hillier L, Eddy SR, Stricklin SL, Baillie DL, Waterston R, Marra MA: **Changes in gene expression associated with developmental arrest and longevity in Caenorhabditis elegans.** *Genome Res* 2001, **11**:1346–1352.
7. O'Riordan VB, Burnell AM: **Intermediary metabolism in the dauer larva of the nematode Caenorhabditis elegans–II. The glyoxylate cycle and fatty-acid oxidation.** *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 1990, **95**:125–130.
8. Holt SJ, Riddle DL: **SAGE surveys C. elegans carbohydrate metabolism: evidence for an anaerobic shift in the long-lived dauer larva.** *Mech. Ageing Dev* 2003, **124**:779–800.
9. Liu T, Zimmerman K, Patterson G: **Regulation of signaling genes by TGFbeta during entry into dauer diapause in C. elegans.** *BMC Developmental Biology* 2004, **4**:11.
10. Cherkasova V, Ayyadevara S, Egilmez N, Reis RS: **Diverse Caenorhabditis elegans genes that are upregulated in dauer larvae also show elevated transcript levels in long-lived, aged, or starved adults.** *Journal of Molecular Biology* 2000, **300**:433–448.
11. Harvey S, Barker G, Shorto A, Viney M: **Natural variation in gene expression in the early development of dauer larvae of Caenorhabditis elegans.** *BMC Genomics* 2009, **10**:325.
12. Ruzanov P, Riddle DL, Marra MA, McKay SJ, Jones SM: **Genes that may modulate longevity in C. elegans in both dauer larvae and long-lived daf-2 adults.** *Exp. Gerontol* 2007, **42**:825–839.
13. Jeong P-Y, Kwon M-S, Joo H-J, Paik Y-K: **Molecular time-course and the metabolic basis of entry into dauer in Caenorhabditis elegans.** *PLoS ONE* 2009, **4**:e4162.
14. McElwee JJ, Schuster E, Blanc E, Thomas JH, Gems D: **Shared transcriptional signature in Caenorhabditis elegans Dauer larvae and long-lived daf-2 mutants implicates detoxification system in longevity assurance.** *J. Biol. Chem* 2004, **279**:44533–44543.
15. C. elegans Sequencing Consortium: **Genome sequence of the nematode C. elegans: a platform for investigating biology.** *Science* 1998, **282**:2012–2018.
16. Kim SK, Lund J, Kiraly M, Duke K, Jiang M, Stuart JM, Eizinger A, Wylie BN, Davidson GS: **A gene expression map for Caenorhabditis elegans.** *Science* 2001, **293**:2087–2092.
17. Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, Jones SJM, Davis EB, Scherer S, Ward S, Kim SK: **A Global Profile of Germline Gene Expression in C. elegans.** *Molecular Cell* 2000, **6**:605–616.
18. Mawuenyega KG, Kaji H, Yamuchi Y, Shinkawa T, Saito H, Taoka M, Takahashi N, Isobe T: **Large-scale identification of Caenorhabditis elegans proteins by multidimensional liquid chromatography-tandem mass spectrometry.** *J. Proteome Res* 2003, **2**:23–35.
19. Krijgsveld J, Ketting RF, Mahmoudi T, Johansen J, Artal-Sanz M, Verrijzer CP, Plasterk RHA, Heck AJR: **Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics.** *Nat. Biotechnol* 2003, **21**:927–931.
20. Gerstein MB, Lu ZJ, Van Nostrand EL, *et al*: **Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project.** *Science* 2010, **330**:1775–1787.
21. Cañestro C, Yokoi H, Postlethwait JH: **Evolutionary developmental biology and genomics.** *Nat. Rev. Genet* 2007, **8**:932–942.
22. Sommer RJ, Streit A: **Comparative genetics of nematodes: genome structure and lifestyle.** *Annual Review of Genetics* 2011, **45**:110301100158074.
23. Hashimshony T, Yanai I: **Revealing developmental networks by comparative transcriptomics.** *Transcr* 2010, **1**:154–158.
24. Hong RL, Sommer RJ: **Pristionchus pacificus: a well-rounded nematode.** *Bioessays* 2006, **28**:651–9.
25. Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, Mitreva M, Roeseler W, Tian H, Witte H, Yang S-P, Wilson RK, Sommer RJ: **The Pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism.** *Nat Genet* 2008, **40**:1193–1198.
26. Borchert N, Dieterich C, Krug K, Schütz W, Jung S, Nordheim A, Sommer RJ, Macek B: **Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models.** *Genome Res* 2010, **20**:837–846.
27. Herrmann M, Mayer WE, Hong RL, Kienle S, Minasaki R, Sommer RJ: **The nematode Pristionchus pacificus (Nematoda: Diplogastridae) is associated with the oriental beetle Exomala orientalis (Coleoptera: Scarabaeidae) in Japan.** *Zool. Sci* 2007, **24**:883–9.
28. Weller AM, Mayer WE, Rae R, Sommer RJ: **Quantitative assessment of the nematode fauna present on Geotrupes dung beetles reveals species-rich communities with a heterogeneous distribution.** *J. Parasitol* 2010, **96**:525–531.
29. Bento G, Ogawa A, Sommer RJ: **Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution.** *Nature* 2010, **466**:494–497.
30. Zheng M, Messerschmidt D, Jungblut B, Sommer RJ: **Conservation and diversification of Wnt signaling function during the evolution of nematode vulva development.** *Nat. Genet* 2005, **37**:300–4.

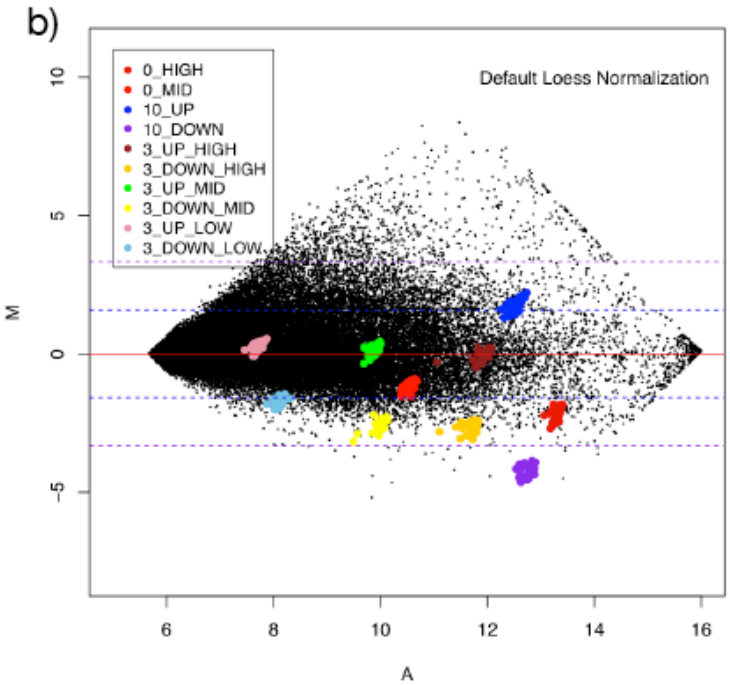
31. Tian H, Schlager B, Xiao H, Sommer RJ: **Wnt signaling induces vulva development in the nematode Pristionchus pacificus.** *Curr. Biol* 2008, **18**:142–6.
32. Wang X, Sommer RJ: **Antagonism of LIN-17/Frizzled and LIN-18/Ryk in nematode vulva induction reveals evolutionary alterations in core developmental pathways.** *PLoS Biol.* 2011, **9**:e1001110.
33. Ogawa A, Streit A, Antebi A, Sommer RJ: **A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes.** *Current Biology* 2009, **19**:67–71.
34. Ogawa A, Bento G, Bartelmes G, Dieterich C, Sommer RJ: **Pristionchus pacificus daf-16 is essential for dauer formation but dispensable for mouth form dimorphism.** *Development* 2011, **138**:1281–1284.
35. Eisen MB, Brown PO: **DNA arrays for analysis of gene expression.** *Meth. Enzymol.* 1999, **303**:179–205.
36. Harris TW, Antoshechkin I, Bieri T, Blasiar D, Chan J, Chen WJ, De La Cruz N, Davis P, Duesbury M, Fang R, Fernandes J, Han M, Kishore R, Lee R, Muller H-M, Nakamura C, Ozersky P, Petcherski A, Rangarajan A, Rogers A, Schindelman G, Schwarz EM, Tuli MA, Van Auken K, Wang D, Wang X, Williams G, Yook K, Durbin R, Stein LD, Spieth J, Sternberg PW: **WormBase: a comprehensive resource for nematode research.** *Nucleic Acids Research* 2009, **38**:D463–D467.
37. Alexa A, Rahnenfuhrer J, Lengauer T: **Improved scoring of functional groups from gene expression data by decorrelating GO graph structure.** *Bioinformatics* 2006, **22**:1600–1607.
38. Papaioannou S, Marsden D, Franks CJ, Walker RJ, Holden-Dye L: **Role of a FMRFamide-like family of neuropeptides in the pharyngeal nervous system of Caenorhabditis elegans.** *J. Neurobiol.* 2005, **65**:304–319.
39. O'Riordan VB, Burnell AM: **Intermediary metabolism in the dauer larva of the nematode Caenorhabditis elegans-- 1. Glycolysis, gluconeogenesis, oxidative phosphorylation and the tricarboxylic acid cycle.** *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 1989, **92**:233–238.
40. Ghazi A, Henis-Korenblit S, Kenyon C: **Regulation of Caenorhabditis elegans lifespan by a proteasomal E3 ligase complex.** *Proc. Natl. Acad. Sci. U.S.A.* 2007, **104**:5947–5952.
41. Li W, Gao B, Lee S-M, Bennett K, Fang D: **RLE-1, an E3 ubiquitin ligase, regulates C. elegans aging by catalyzing DAF-16 polyubiquitination.** *Dev. Cell* 2007, **12**:235–246.
42. Yun C, Stanhill A, Yang Y, Zhang Y, Haynes CM, Xu C-F, Neubert TA, Mor A, Philips MR, Ron D: **Proteasomal adaptation to environmental stress links resistance to proteotoxicity with longevity in Caenorhabditis elegans.** *Proceedings of the National Academy of Sciences* 2008, **105**:7094–7099.
43. Mayer MG, Sommer RJ: **Natural variation in Pristionchus pacificus dauer formation reveals cross-preference rather than self-preference of nematode dauer pheromones.** *Proceedings of the Royal Society B: Biological Sciences* 2011, **278**:2784–2790.
44. Kanehisa M, Goto S: **KEGG: Kyoto Encyclopedia of Genes and Genomes.** *Nucleic Acids Res.* 2000, **28**:27–30.
45. Dieterich C, Sommer RJ: **How to become a parasite - lessons from the genomes of nematodes.** *Trends Genet* 2009, **25**:203–209.
46. Mayer W, Schuster L, Bartelmes G, Dieterich C, Sommer R: **Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover.** *BMC Evolutionary Biology* 2011, **11**:13.
47. Tanaka H, Sato K, Saito Y, Yamashita T, Agoh M, Okunishi J, Tachikawa E, Suzuki K: **Insect diapause-specific peptide from the leaf beetle has consensus with a putative iridovirus peptide.** *Peptides* 2003, **24**:1327–1333.
48. Inoue T, Ailion M, Poon S, Kim HK, Thomas JH, Sternberg PW: **Genetic analysis of dauer formation in Caenorhabditis briggsae.** *Genetics* 2007, **177**:809–818.
49. Rudel D, Riebesell M, Sommer RJ: **Gonadogenesis in Pristionchus pacificus and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad.** *Dev. Biol* 2005, **277**:200–21.
50. Pires-daSilva A, Sommer RJ: **Conservation of the global sex determination gene tra-1 in distantly related nematodes.** *Genes Dev* 2004, **18**:1198–208.
51. True JR, Haag ES: **Developmental system drift and flexibility in evolutionary trajectories.** *Evol. Dev* 2001, **3**:109–119.
52. Davidson EH, Erwin DH: **Gene regulatory networks and the evolution of animal body plans.** *Science* 2006, **311**:796–800.
53. Gerhart J, Kirschner M: *Cells, Embryos and Evolution: toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability.* Malden: Blackwell Science; 1997.
54. Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B: **Constraint and turnover in sex-biased gene expression in the genus Drosophila.** *Nature* 2007, **450**:233–237.
55. Yanai I, Hunter CP: **Comparison of diverse developmental transcriptomes reveals that coexpression of gene neighbors is not evolutionarily conserved.** *Genome Res* 2009, **19**:2214–2220.
56. Yanai I, Peshkin L, Jorgensen P, Kirschner MW: **Mapping gene expression in two Xenopus species: evolutionary constraints and developmental flexibility.** *Dev. Cell* 2011, **20**:483–496.
57. Zheng-Bradley X, Rung J, Parkinson H, Brazma A: **Large scale comparison of global gene expression patterns in human and mouse.** *Genome Biol* 2010, **11**:R124.
58. Parikh A, Miranda ER, Katoh-Kurasawa M, Fuller D, Rot G, Zagar L, Curk T, Sucgang R, Chen R, Zupan B, Loomis WF, Kuspa A, Shaulsky G: **Conserved developmental transcriptomes in evolutionarily divergent species.** *Genome Biol* 2010, **11**:R35.
59. Wernersson R, Juncker AS, Nielsen HB: **Probe selection for DNA microarrays using OligoWiz.** *Nat. Protocols* 2007, **2**:2677–2691.
60. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini A, Sawitzki G, Smith C, Smyth G, Tierney L, Yang J, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biology* 2004, **5**:R80.
61. Smyth GK: **Limma: linear models for microarray data**. In *In Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer; 2005:397–420.
62. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, Smyth GK: **A comparison of background correction methods for two-colour microarrays.** *Bioinformatics* 2007, **23**:2700–2707.
63. Smyth GK, Speed T: **Normalization of cDNA microarray data.** *Methods* 2003, **31**:265–273.
64. van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, Holstege FCP: **Monitoring global messenger RNA changes in externally controlled microarray experiments.** *EMBO Rep* 2003, **4**:387–393.
65. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Statistical Applications in Genetics and Molecular Biology* 2004, **3**. doi:http://dx.doi.org/10.2202/1544-6115.1027.
66. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M: **Minimum information about a microarray experiment (MIAME) - toward standards for microarray data.** *Nat Genet* 2001, **29**:365–371.
67. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A: **The Pfam protein families database.** *Nucleic Acids Research* 2009, **38**:D211–D222.
68. Eddy SR: **A new generation of homology search tools based on probabilistic inference.** *Genome Inform* 2009, **23**:205–211.

doi:10.1186/1471-2164-13-254
Cite this article as: Sinha *et al.*: Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*. *BMC Genomics* 2012 **13**:254.

Supporting Information:
Sinha et al BMC Genomics 13:254 (2012)

Figure S1.The impact of different normalization strategies on expression fold change estimation. a) The general concept of using spike-ins to control for changes in total mRNA abundance relative to total RNA. See van de Peppel *et al.*, 2003 for a detailed explanation. b) MA-plot of a Loess based normalization assuming a net mRNA expression fold change of 0 between the two samples. The M axis depicts the log2 fold change between the red and green microarray channels. The A axis is the average log2 intensity across both microarray channels. The position of the spike-in signals (colored dots) clearly indicate a deviation from this assumption. The corresponding legend in the upper left shows the expected fold change of the spike-ins relative to total RNA levels. Consequently, the location of the spike-ins corresponds to a down-shift of mRNA abundance levels within the total RNA population. c) Differentially expressed gene counts for the dauer vs. mixed stage comparison conditional on the normalization strategy.





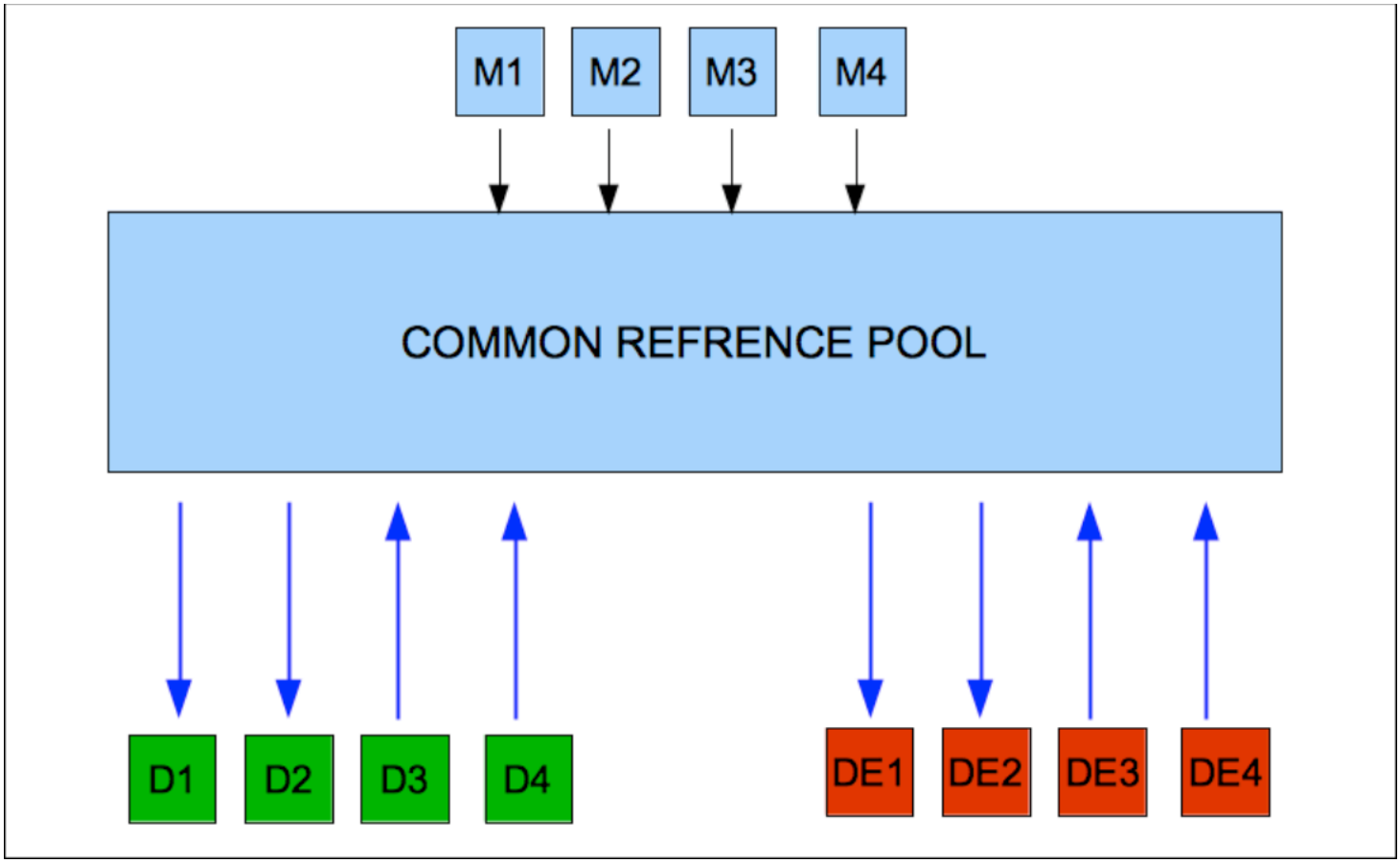
c)

C. elegans Dauer vs Mix stage comparison		
	Default Loess	Weighted loess
UP	1991	1099
DOWN	2681	3162
Total	**4672**	**4261**

P. pacificus Dauer vs Mix stage comparison		
	Default Loess	Weighted loess
UP	5166	3910
DOWN	5748	6121
Total	**10914**	**10031**

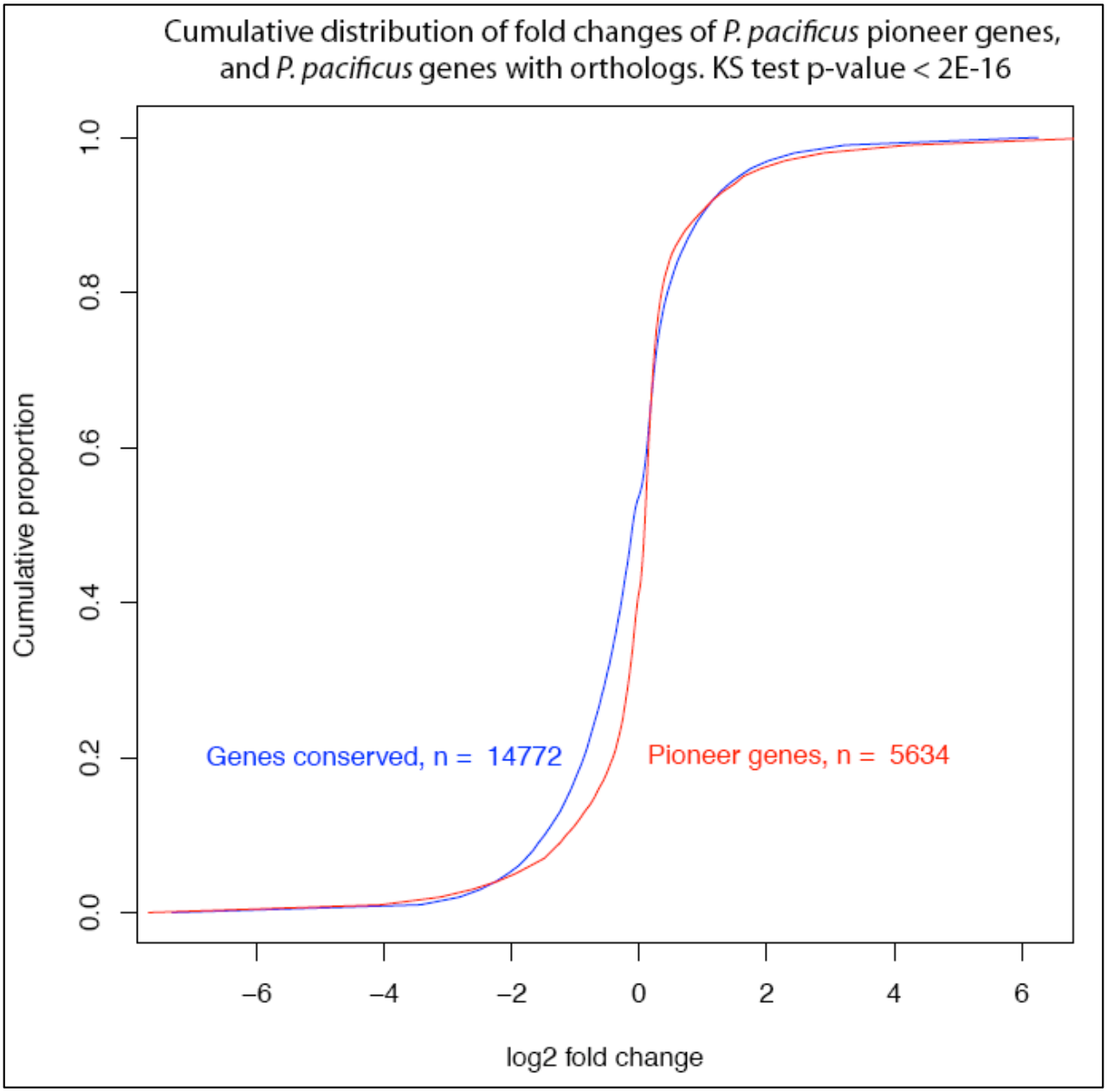
Supporting Information:
Sinha et al BMC Genomics 13:254 (2012)

Figure S2. Explanation of a common reference design for microarrays (see Eisen and Brown 1999). For each species, RNA from biological replicates of mix-stage worms (M1 to M4) were combined together to generate one common reference pool. Labelled aRNA produced from independent biological replicates were then co-hybridized with labelled aRNA from the common reference pool, including two dye-swaps. Samples D1 to D4 represent Dauer samples and DE1 to DE4 represent dauer exit samples at 12 hour timepoint. The blue arrows indicate the direction of labelling in each co-hybridization (arrow head = Cy3 labelled, arrow tail = Cy5 labelled). This design allows comparison via the common reference pool, but at the same time remains flexible for adding more time-points to the study, if needed



Supporting Information:
Sinha et al BMC Genomics 13:254 (2012)

Figure S3. Cumulative plot of log2 expression fold changes for genes with and without sequence conservation (pioneer genes). The two distributions are significantly different (two-sample Kolmogorov-Smirnov test; p-value < 10E- 16).



OPEN ACCESS Freely available online

PLOS ONE

System Wide Analysis of the Evolution of Innate Immunity in the Nematode Model Species *Caenorhabditis elegans* and *Pristionchus pacificus*

Amit Sinha, Robbie Rae, Igor Iatsenko, Ralf J. Sommer*

Max Planck Institute for Developmental Biology, Department of Evolutionary Biology, Tübingen, Germany

Abstract

The evolution of genetic mechanisms used to combat bacterial infections is critical for the survival of animals and plants, yet how these genes evolved to produce a robust defense system is poorly understood. Studies of the nematode *Caenorhabditis elegans* have uncovered a plethora of genetic regulators and effectors responsible for surviving pathogens. However, comparative studies utilizing other free-living nematodes and therefore providing an insight into the evolution of innate immunity have been lacking. Here, we take a systems biology approach and use whole genome microarrays to profile the transcriptional response of *C. elegans* and the necromenic nematode *Pristionchus pacificus* after exposure to the four different pathogens *Serratia marcescens, Xenorhabdus nematophila, Staphylococcus aureus* and *Bacillus thuringiensis* DB27. *C. elegans* is susceptible to all four pathogens whilst *P. pacificus* is only susceptible to *S. marcescens* and *X. nematophila*. We show an unexpected level of specificity in host responses to distinct pathogens within and across species, revealing an enormous complexity of effectors of innate immunity. Functional domains enriched in the transcriptomes on different pathogens are similar within a nematode species but different across them, suggesting differences in pathogen sensing and response networks. We find translation inhibition to be a potentially conserved response to gram-negative pathogens in both the nematodes. Further computational analysis indicates that both nematodes when fed on pathogens up-regulate genes known to be involved in other stress responses like heat shock, oxidative and osmotic stress, and genes regulated by DAF-16/FOXO and TGF-beta pathways. This study presents a platform for comparative systems analysis of two nematode model species, and a catalog of genes involved in the evolution of nematode immunity and identifies both pathogen specific and pan-pathogen responses. We discuss the potential effects of ecology on evolution of downstream effectors and upstream regulators on evolution of nematode innate immunity.

Citation: Sinha A, Rae R, Iatsenko I, Sommer RJ (2012) System Wide Analysis of the Evolution of Innate Immunity in the Nematode Model Species *Caenorhabditis elegans* and *Pristionchus pacificus*. PLoS ONE 7(9): e44255. doi:10.1371/journal.pone.0044255

Editor: Ilya Ruvinsky, The University of Chicago, United States of America

Received April 9, 2012; **Accepted** July 31, 2012; **Published** September 28, 2012

Funding: This work was funded by the Max-Planck Society. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ralf.sommer@tuebingen.mpg.de

Introduction

The struggle against infectious diseases caused by bacteria, viruses, fungi, protozoa and metazoan parasites is an important evolutionary agent [1] leading to rapid evolutionary changes responsible for much of the complexity found in the immune system of animals [2–4]. However, the molecular basis for the evolution of such host-pathogen interactions is only poorly understood. Over the past ten years studies of the nematode *Caenorhabditis elegans* have given insight into genes essential for host immunity [5,6] as well as identifying bacterial virulence mechanisms used by opportunistic mammalian pathogens [7,8]. These studies (and many others) have identified various signaling pathways critical for *C. elegans* survival when fed an array of bacterial and fungal pathogens e.g. ERK MAP kinase, p38 MAP kinase, TGF β, programmed cell death, DAF–2/DAF–16 insulin-like receptor signaling and JNK-like MAP kinase [6,9–13], as well as components such as the G-protein coupled receptor FSHR-1, bZIP transcription factor *zip−2* and beta-Catenin/*bar−1* which are required for an inducible pathogen response [14–16]. It remains to be discovered however, how important these pathways are in other nematode species and how these pathways contribute to the evolutionary trajectories of bacterial pathogenicity. A comparative approach with another nematode species would provide a first entry point to enhance our understanding of the evolutionary diversity of host (nematode) response to pathogens.

One nematode that has been used extensively for comparative studies with *C. elegans* is the diplogastrid species *Pristionchus pacificus* (see [17]). In addition to having a well characterized proteome and a fully sequenced genome [18,19], forward and reverse genetics [20], and transgenic techniques [21], full genome microarray technology has also been developed [22] allowing genomic analysis of many different traits. *P. pacificus* diverged from *C. elegans* 250–400 million years ago [18] and during this time there have been remarkable changes in vulva development [23,24], gonad morphogenesis [25], sex determination [26] and chemosensory behaviour [27] allowing for evolutionary and developmental comparisons with *C. elegans*. These two nematodes also differ in their ecological niches. *C. elegans* can be isolated from compost heaps, snails and rotten fruits [28], whereas *P. pacificus* is usually isolated from a range of scarab beetles [29–32]. *P. pacificus*, as well as other *Pristionchus* species live in a necromenic lifestyle,

that is feeding on microorganisms growing on the carcass of beetles once they die [29].

C. elegans and *P. pacificus* not only live in different ecological niches, but also differ in their susceptibility to bacterial pathogens. For example, *C. elegans* dies when fed the human opportunistic bacteria *Pseudomonas aeruginosa*, *Staphylococcus aureus* and insecticidal *Bacillus thuringiensis* Cry 5B toxin whereas *P. pacificus* is resistant [33,34]. More recently, a screen of about 1,400 naturally strains of *Bacillus* yielded three strains of *Bacillus thuringiensis* DB27 that are extremely toxic to *C. elegans* but *P. pacificus* remain resistant [35]. Anatomically, *C. elegans* and *P. pacificus* differ in that *C. elegans* contains a grinder in the posterior bulb of the pharynx that is involved in the physical lysis of bacterial food [36]. While the grinder is a typical structure of nematodes of the Rhabditidae family, no grinder exists in nematodes of the Diplogastridae family, to which *P. pacificus* belongs [35–37]. Given these strong differences in the ecology and anatomy of *C. elegans* and *P. pacificus*, these two species represent ideal candidates for studying the evolution of the genetic mechanisms of pathogen response in nematodes.

Here, we used a systems level approach by testing in parallel four different bacterial pathogens that cause distinct effects on the two nematodes. We analyzed whole genome gene expression of *C. elegans* and *P. pacificus* when fed four different pathogens (*Serratia marcescens*, *Xenorhabdus nematophila*, *B. thuringiensis* DB27 and *S. aureus*) and compared each nematode pathogen response to those fed on the control bacterium (the standard nematode lab food *Escherichia coli* OP50). *S. marcescens* is a broad host pathogen present in soil and insects that kills *C. elegans* [8,38] and *P. pacificus*. *X. nematophila* is a symbiotic bacteria of the entomopathogenic nematode *Steinernema carpocapsae*, which utilizes the bacteria to kill insects and feed on the resulting mass of proliferating bacteria [39], and also kills both the nematodes [34]. Our *B. thuringiensis* DB27 strain was isolated from a dung beetle (*Geotrupes* spp.) and seems to be one of the most pathogenic bacteria of *C. elegans* reported in the literature so far, which kills *C. elegans* in less than sixteen hours while *P. pacificus* is resistant [35]. We show an unexpected level of specificity in host responses to distinct pathogens within and across species, revealing an enormous complexity of effectors of innate immunity. This study presents (i) a platform for comparative systems biology of two nematode models, (ii) a catalog of genes involved in the evolution of nematode immunity and (iii) pathogen specific and pan-pathogen responses from both *C. elegans* and *P. pacificus*.

Results and Discussion

Survival of *C. elegans* and *P. pacificus* differs when fed gram-positive pathogens

To study the evolution of the genetic mechanisms involved in nematode resistance against bacteria, we fed the four bacterial pathogens *S. aureus*, *B. thuringiensis* DB27, *S. marcescens* and *X. nematophila* to the two nematode model species *C. elegans* and *P. pacificus* and assessed their effect on survival. When *C. elegans* is fed monoxenic lawns of each of the four pathogens, it dies within 2–5 days (Median survival time <5 days, Figure 1A). This is in stark contrast to *P. pacificus*, which is more resistant to the gram-positive pathogens *B. thuringiensis* DB27 and *S. aureus* and can survive for more than 7 days (Figure 1B and Figure S1, Median survival times $\sim$ 8 days). However, *P. pacificus*, like *C. elegans*, is susceptible to both *X. nematophila* and *S. marcescens* and 50% of the population dies within 2–3 days exposure (Figure 1B). We would like to note here that the difference in susceptibility of *P. pacificus* to gram-positive bacteria tested in this study is not simply an artifact of a longer life-span, as its wild-type life-span and developmental rate is comparable to that of *C. elegans* ([17,40] and our unpublished observations). Also, *P. pacificus* is highly susceptible (median survival $\sim$ 3.5 days) to a gram-positive *Bacillus* strain DB35 isolated from *Geotrupes sp*. beetles [35], indicating that it is not more resistant to gram-positive bacteria in general. We also note that *P. pacificus* is able to survive and reproduce on both *B. thuringiensis* and *S. aureus* indicating that its reduced susceptibility should not be due to reduced bacterial intake.

Significant transcriptional changes after exposure to different bacteria

To investigate the transcriptional response of the two nematodes upon exposure to these different bacterial pathogens, we identified differentially expressed genes using whole genome microarrays containing $\sim$43,000 probes for 20,149 *C. elegans* genes and $\sim$90,000 probes for 20,987 genes in *P. pacificus* respectively [22]. For each condition, total RNA was collected from four separate biological replicates of about 200 synchronized young adult worms exposed either to the pathogen or to the control *E. coli* (OP50) for 4 hours. The labeled cRNA produced from this total RNA was hybridized to species-specific microarrays according to manufacturer's protocols, and the raw data from scanned images was analyzed using the "limma" package in Bioconductor (see Methods for details). We observed that the exposure to pathogen resulted in a decrease of amount of total RNA produced per worm when compared to the relatively non-pathogenic *E. coli* strain (Figure S2). This global decrease in transcription is most likely a common feature of a core stress response, as it is also seen in case of dauer formation in both the species [22,41] or might be an effect of the various bacteria on efficieny of RNA extraction. Nonetheless, such global transcriptional changes call for optimization of parameters used in normalization of microarray data, without which the calculated fold-changes can be erroneous [22,42]. Our fold-change calculations take these factors into account (see Methods for details).

Although previous studies of pathogen response in *C. elegans* [9,43–47] have looked at the transcriptome at different time-points such as 4-hours, 8-hours or 24-hours after exposure, we chose to investigate one of the earliest time-point of 4 hours because we were interested in earliest transcriptional responses manifested in response to each of these pathogens. Pathogenesis related necrosis of host-tissue at later time-points is a common effect of many pathogens and such organism-wide necrosis might dominate the expression profile at later time points [46], masking the pathogen specific signatures. Also, pathogens like *Bacillus thuringiensis* DB27 kill *C. elegans* in less than 24 hours, making analysis of late time-points unfeasible [35].

Widely different numbers of genes are found to be up- or down-regulated in *C. elegans* or *P. pacificus* upon exposure to different pathogens (Table 1A and 1B), indicating both a pathogen-specific and a nematode-specific component to these responses. In this context it is interesting to note that just a 4-hour exposure to pathogen is sufficient to cause large transcriptional changes in both the species, suggesting rapid activation of innate immune response.

Changes in nematode gene expression depend on lethality and rate of killing

Based on the absolute number of differentially expressed genes under different conditions of survival, lethality and slower or faster killing rates of pathogens, some patterns can be discerned in our microarray data. First, the number of differentially expressed genes seems to be inversely correlated with the survival characteristics of the nematodes. For example, upon exposure to

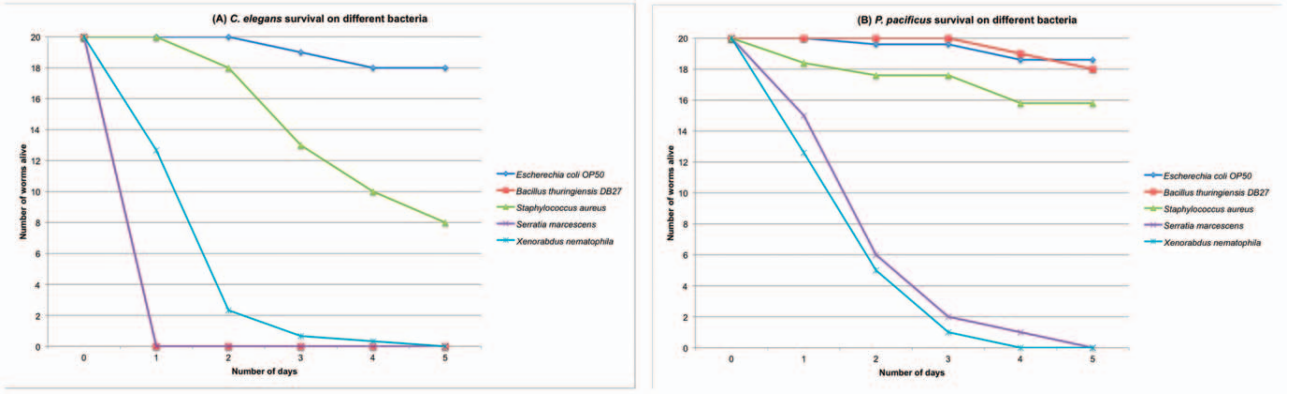


Figure 1. Differences in survival of *C. elegans* and *P. pacificus* upon exposure to different pathogenic bacteria. The survival of *C. elegans* and *P. pacificus* is different upon exposure to different bacteria. The survival curves for (A) *C. elegans* and (B) *P. pacificus* were obtained after exposure to the gram-positive bacteria *Bacillus thuringiensis* DB27 and *Staphylococcus aureus*, and the gram-negative bacteria *Serratia marcescens* and *Xenorhabdus nematophila*. Standard lab food *Escherichia coli* OP50 was used as a control for both nematodes. Both nematodes show reduced survival on *S. marcescens* and *X. nematophila*. *C. elegans* is also susceptible to *B. thuringiensis* DB27 and *S. aureus*, while *P. pacificus* shows higher resistance to these two bacteria.
doi:10.1371/journal.pone.0044255.g001

B. thuringiensis DB27, the pathogen most lethal to *C. elegans*, a remarkably large number of genes are affected in *C. elegans* (n = 5868, Table 1A), whereas much fewer genes (n = 217, Table 1B) are induced in *P. pacificus*, which is resistant to this pathogen. This can be attributed to the fact that while *P. pacificus* can use *B. thuringiensis* DB27 for food, *C. elegans* has to mount a robust response against a lethal pathogen.

For *C. elegans*, the number of differentially expressed genes is greater when exposed to faster-killing pathogens *B. thuringiensis* DB27, *X. nematophila* and *S. marcescens* as compared to that on *S. aureus*, where worms survive longer (Table 1A). Similarly, in *P. pacificus*, greater number of genes is differentially expressed on more lethal pathogens *X. nematophila* and *S. marcescens* (Table 1B) as compared to that on *B. thuringiensis* DB27 and *S. aureus* (Table 1B), to which *P. pacificus* is more resistant (Figure 1B).

Further, in *P. pacificus*, which is either more resistant to pathogens or shows slower mortality kinetics as compared to *C. elegans*, the expression profiles are observed to be usually smaller or just as large as that in *C. elegans*. We can rule out that these differences in profile size are due to potential differences in sensitivity of the two microarray platforms used, because we know from previous studies that our *P. pacificus* microarrays could detect differential expression of larger number of genes under different conditions such as dauer formation [22]. We also checked if changing the p-value cutoffs on microarray data abolishes the difference in profile sizes, but we find that the trend still holds (data not shown). Also, in *C. elegans* we have observed that when it is exposed to a non-pathogenic *Bacillus subtilis* strain for 4 hours, the number of genes differentially expressed is relatively low (~510 genes) [II, AS and RJS, unpublished observations]. Hence, the differences in profile size between *C. elegans* and *P. pacificus* are most likely biologically relevant and not just a technical artifact.

Gram-positive bacteria predominantly induce over-expression of genes while gram-negative bacteria cause transcriptional suppression

For both nematode species, the gram-positive bacteria tested induce up-regulation of relatively more genes as compared to down-regulation, while the reverse seems to be true for the gram-

Table 1. Widely different numbers of genes are differentially expressed in (A) *C. elegans* and (B) *P. pacificus* in response to the four bacteria.

(A) C. elegans	Up	Down	TOTAL	%Up	%Down
B. thuringiensis	5532	156	5688	97%	3%
S. aureus	181	68	249	73%	27%
S. marcescens	1465	4931	6396	23%	77%
X. nematophila	732	7884	8616	15%	85%
(B) P. pacificus	**Up**	**Down**	**TOTAL**	**%Up**	**%Down**
B. thuringiensis	156	61	217	72%	28%
S. aureus	178	140	318	56%	44%
S. marcescens	192	1007	1199	16%	84%
X. nematophila	848	4293	5141	16%	84%

The genes were called differentially expressed on microarrays if the FDR corrected p-value was less than 0.05 and the absolute value of fold changes was greater than 1.42 (corresponding to log2 fold change of 0.5 where log2(1.42) = 0.5).
doi:10.1371/journal.pone.0044255.t001

negative bacteria (Table 1). For example, in both the nematodes, exposure to the gram-positive pathogens *B. thuringiensis* DB27 and *S. aureus* causes induction of relatively more genes as compared to suppression (Figure 2, Table 1). On the other hand, exposure to the gram-negative pathogens *S. marcescens* and *X. nematophila* predominantly causes down-regulation of comparatively more genes than up-regulation (Figure 2, Table 1). Thus, the relative proportion of up-regulated versus down-regulated genes appears to depend upon some common factor(s) shared by either the gram-negative or the gram-positive bacteria, although more bacteria from both groups need to be tested to confirm this trend.

A bacterium-specific transcriptional response is mounted by both *C. elegans* and *P. pacificus*

Next, we investigated the intra-specific response of the two nematodes after exposure to different pathogens. Although *C. elegans* is unable to survive on any of the four pathogens, we find that the expression profiles on each of the pathogens are qualitatively quite different from each other and only a small fraction of genes are common between expression profiles obtained on different pathogens (Figure 3A). Specifically, only 102 genes change their expression upon exposure to each of the four pathogens (Figure 3A, genes with nCommon = 4 in Table S1) but the number slightly increases to about 687 genes when the criterion is relaxed to significant differential expression in more than one expression profile (genes with nCommon >1 in Table S1). A small number of overlap between multiple pathogen response profiles is a signature for highly specific pathogen response, and has also been observed before e.g. only 22 genes were reported to be common between profiles after 24 hour exposure to *Erwinia carotovara*, *Enterococcus faecalis*, and *Photorhabdus luminescens* [46]. Interestingly, the genes induced in *C. elegans* across all four pathogens include the transcription factors *pqm−1* and *zip−2* (Table S1). The stress responsive transcription factor *pqm−1* is also induced and required for defense in response to *P. aeruginosa* infection [43]. The bZIP transcription factor *zip−2* is a known to regulate a subset of *pmk-1* independent pathogen response genes on *P. aeruginosa* [15]. Similarly, the *P. pacificus* expression profiles also show a bacterium-specific signature, with only 18 genes being common across all the four profiles (n = 18, Figure 3B, genes with nCommon = 4 in Table S2) while 206 genes are common between the expression profiles on more than one pathogen (genes with nCommon >1 in Table S2).

This pathogen specific nature of expression profiles within a nematode species is further highlighted in an expression cluster analysis (see Methods), where we compare our data-sets with various published microarray studies of pathogen-response [12,43–50]. Based on the significance of overlaps between different microarray data sets (Table S3A and Table S4A for *C. elegans* and *P. pacificus* respectively), it is evident that only a small proportion of genes in each expression profile show an overlap with expression profiles on other pathogens.

We further evaluated the extent of similarities between different intra-specific expression profiles, by carrying out a two-dimensional hierarchical clustering [51] on log-fold change values of genes that were significantly differentially expressed on at least two bacterium (n = 687 genes in *C. elegans*, and n = 206 genes in *P. pacificus*, number of common 1:1 orthologs = 15, Table S1 and Table S2). Interestingly, in both *C. elegans* and *P. pacificus* heat-maps (Figure 4), the expression profile of *X. nematophila* response clusters separately from the profiles obtained in response to the other three bacteria, suggesting some differences in its pathogenicity mechanism(s) compared to other three bacteria. In summary, both *C. elegans* and *P. pacificus* show a bacterium-specific transcriptional response, with relatively few common genes being regulated across multiple bacteria in a given nematode.

Pfam domain enrichment analysis identifies a role for lipid metabolism and the detoxification machinery in pathogen response in *P. pacificus*

Identification of functional components in large gene-sets such as the pathogen response expression profiles can be better achieved by meta-analysis based on functional annotations rather than by searching for a gene-to-gene correspondence. Therefore, we tested for enrichment of various Pfam domains [52] in the proteins corresponding to the differentially expressed genes (see Methods for details) to see if common functional themes emerge despite only partially overlapping gene lists.

For *C. elegans*, we find the domain enrichment profile to be similar for all bacteria except *S. aureus*. The common domains include those related to Proteasome function, ATPase activity (AAA domain), DNA helicases with DEAD box and Helicase_C

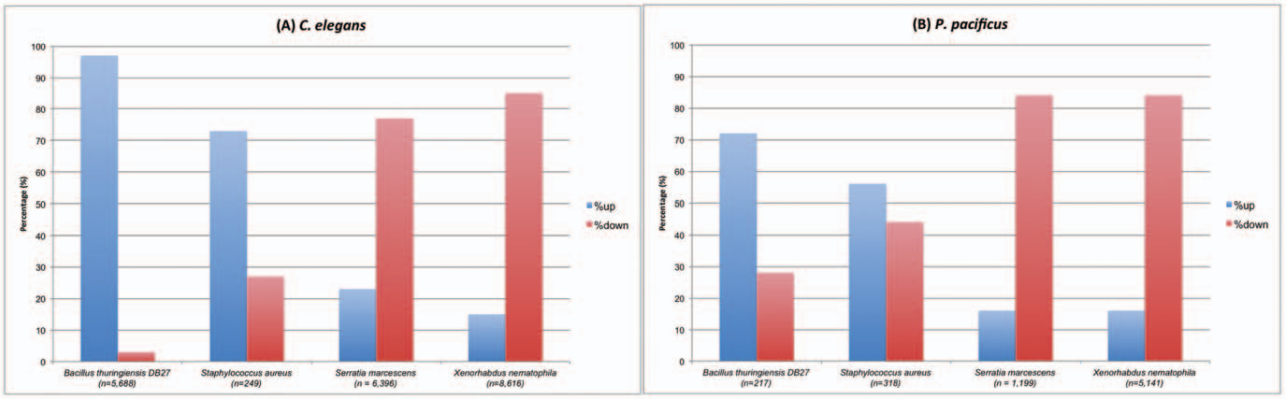


Figure 2. Gram-positive and gram-negative bacteria induce different proportions of up- versus down-regulated genes in *C. elegans* and *P. pacificus*. Despite the differences in number of genes differentially expressed on each of the pathogens in the two nematodes, the relative proportion of up-regulated genes is higher than that of down-regulated genes on gram-positive bacteria *B. thuringiensis* and *S. aureus* in both (A) *C. elegans* and (B) *P. pacificus*. On the other hand, exposure to the gram-negative bacteria *S. marcescens* and *X. nematophila* results in down-regulation of a greater fraction of genes as compared to the up-regulated genes in both (A) *C. elegans* and (B) *P. pacificus*.
doi:10.1371/journal.pone.0044255.g002

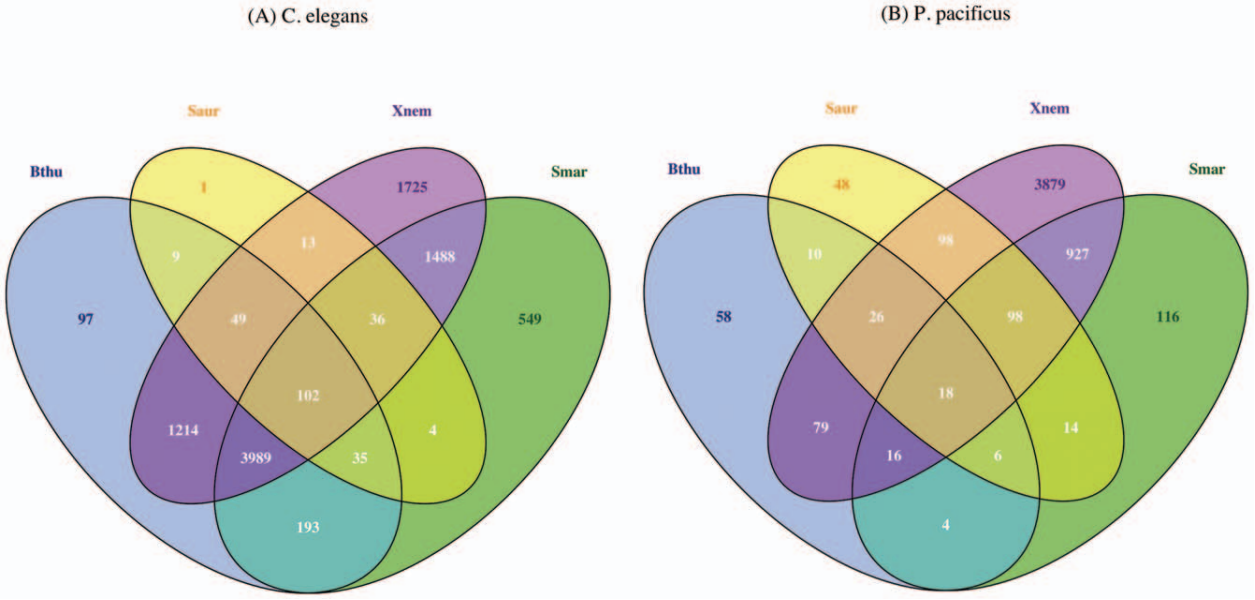


Figure 3. A pathogen-specific transcriptional response is mounted by both *C. elegans* and *P. pacificus* upon exposure to different bacteria. Overlap between the different genes differentially regulated in response to the four pathogens tested is represented as a Venn diagram for *C. elegans* and *P. pacificus*. Only 102 genes are found to be common between the expression profiles on all four pathogens in (A) *C. elegans*, while only 18 genes are common between the expression profiles corresponding to the four pathogens in (B) *P. pacificus*. This minimal overlap indicates the existence of a highly pathogen-specific immune response in both the nematodes. The abbreviations Bthu, Saur, Smar and Xnem refer to the bacteria *Bacillus thuringiensis*, *Staphylococcus aureus*, *Serratia marcescens* and *Xenorhabdus nematophila* respectively.
doi:10.1371/journal.pone.0044255.g003

domains, and the RRM_1 motif that is indicative of RNA binding protein activity (Table S5). Thus even though the gene-by-gene similarity is low between these expression profiles, we observe proteins with similar functional domains to be enriched in all of them. In *P. pacificus*, expression profiles for all four bacteria are enriched for various lipid metabolism related domains such as Lipase_GDSL, FA_desaturase, Acyl-CoA_dh_1 and Abhydro_lipase (Table S6), suggesting a role for lipid metabolism in *P. pacificus* immune response. Studies on *C. elegans* immune response have shown that the poly-unsaturated fatty acids gamma-linolenic acid and stearidonic acid are integral for immune response, acting via the p38 MAP Kinase pathway [53]. Lipases can potentially

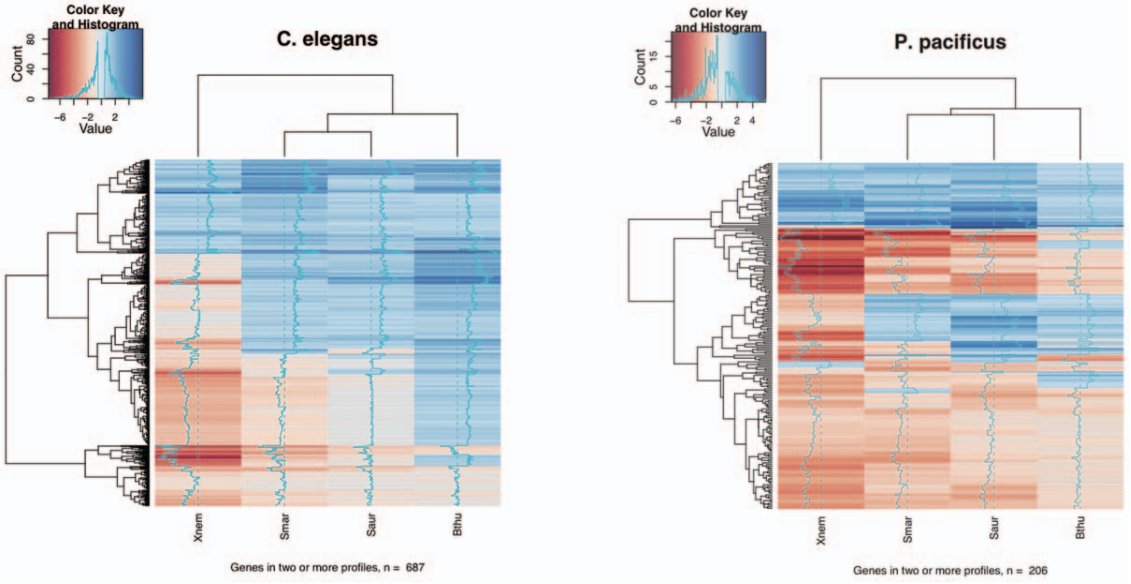


Figure 4. Hierarchical clustering of genes differentially expressed on more than one pathogen identifies clusters of co-regulated genes. The pathogen-response expression profiles for each nematode were clustered based on log2 of fold-changes for genes that were differentially expressed on at least two pathogens. In both (A) *C. elegans* and (B) *P. pacificus*, the expression profile in response to *X. nematophila* clusters separately from that in response to other pathogens, suggesting a difference in its mode of virulence from other pathogens.
doi:10.1371/journal.pone.0044255.g004

function as antagonists of invading pathogens [54] and are known to be induced in response to pathogens in both *C. elegans* [9] and *Drosophila melanogaster* [55,56]. We also see induction of lipase-like genes *lipl-1* and *lipl-3* across all pathogens in *C. elegans* (Table S1). Thus the enrichment of proteins containing lipase and related domains might contribute towards enhanced resistance of *P. pacificus* on some of the pathogens.

C- type lectins have been implicated in the *C. elegans* innate immune response [57] and in transcriptomic studies of exposure to *P. aeruginosa*, *M. nematophila* and *S. marcescens* [9,43–46,58]. Although C-type lectin domain encoding genes were differentially expressed when either *C. elegans or P. pacificus* was fed our four pathogens, the enrichment for the corresponding Pfam domain "Lectin_C" achieved statistical significance only in *P. pacificus* profiles on all pathogens except for that on *S. aureus* (Table S6).

The Pfam domains enriched in *P. pacificus* upon response to the relatively less pathogenic bacteria *S. aureus* as well as on exposure to the highly pathogenic *X. nematophila* also include various domains involved in detoxification and xenobiotic defense, such as Glucuronosyltransferase (UDPGT), Glutathione S-transferase (GST_C) and Cytochrome P450 domain (Table S6), which have been previously identified in expression studies of *C. elegans* exposed to xenobiotic compounds [59]. Interestingly, these domains have undergone an expansion in *P. pacificus* genome relative to the *C. elegans* genome and have been hypothesized to have adaptive significance in context of its necromenic lifestyle [18]. Here for the first time we show that the gene activity for the proteins containing these domains is enriched in a potentially pathogenic scenario and possibly confers an adaptive advantage.

Taken together, the Pfam domain analysis provides further insights into the pathogen response of the two nematodes. We observe similar Pfam domains to be enriched within a given nematode in response to different pathogens, but the set of enriched domains differs between *C. elegans* and *P. pacificus*, such that apart from the DNA helicase domain Helicase_C, we hardly find any other domains common between *C. elegans* and *P. pacificus*, even in response to the same bacteria.

Inhibition of translation machinery is a conserved effect of exposure to gram-negative pathogens in both *C. elegans* and *P. pacificus*

Since *C. elegans* and *P. pacificus* have very different survival behavior on the bacteria tested, we wanted to identify genes whose expression might be responsible for these differences. For comparing the expression profiles across the two nematode species, we restricted our analysis to the 6,126 1:1 orthologous pairs defined by the stringent best reverse BLAST method, for which the probes were present on both the microarrays (see Methods). Interestingly, we see different patterns of overlap between the expression profiles of the two nematodes depending on the bacteria tested. For the gram-positive bacteria *B. thuringiensis* DB27 and *S. aureus*, which kill *C. elegans* at a much higher rate than *P. pacificus*, we observe a very limited overlap in the expression profiles of the two species (Figure 5A and Figure 5B). It was *a priori* not clear whether the *P. pacificus* resistance to *B. thuringiensis* and *S. aureus* is due to induction of similar genes as in *C. elegans*, albeit at higher expression levels, or, if the activation of a totally different set of genes causes the resistance phenotype. The surprisingly small extent of overlap observed in our comparisons supports the second model. Since most of the genes induced in *P. pacificus* on gram-positive bacteria do not have a characterized function yet, future studies will shed light on their role in innate immunity.

In contrast, we observe a highly significant overlap between the *P. pacificus* and *C. elegans* expression profiles in response to the gram-negative bacteria, *S. marcescens* and *X. nematophila*, which are lethal to both the nematodes (Figure 5C and Figure 5D). This suggests that either these gram-negative bacteria induce a similar immune response in both the nematodes, or that late and secondary markers of pathogenesis dominate the expression profile related damages in both the nematode species.

Exposure to pathogens can be expected to affect germline development and reproduction, which might contribute to the set of differentially expressed genes. Consistent with this expectation, we see enrichment of oocyte and germline related expression clusters [60,61] in both *C. elegans* and *P. pacificus* profiles on various pathogens (see Tables S3 and S4, clusters "cgc6390:oogenesis-enriched", "WBPaper00037611:RNP-8-associated" and "WBPaper00037611:GLD-2-associated"). The overlap is strongest with the genes downregulated in response to the most lethal pathogens *S. marcescens* and *X. nematophila*.

To identify the conserved elements of the pathogen response in the two species, we focused on the gram-negative pathogens *S. marcescens* and *X. nematophila*, which are lethal to both the nematodes. Almost all the genes common between the two nematodes (Figure 5C and Figure 5D) show a downregulation in response to each of these bacteria (Tables S7A and S7B). We further found 410 genes to be common across both two nematodes on both the gram-negative pathogens, most of them being downregulated (Table S7C). Interestingly, the stress responsive transcription factor encoded by *pqm-1* was differentially expressed across all the four expression profiles (Table S7C), suggesting a potential and conserved role in innate immunity [43] across the two nematode species. Gene Ontology based analysis (Table S8) on any of these three lists shows an enrichment for biological processes related to "determination of adult lifespan", as well as processes related to protein translation such as "translational elongation", "translational initiation" and "ribosome biogenesis". The corresponding terms under the ontology molecular function include "structural constituent of ribosome", "translation elongation factor activity", and "translation initiation factor activity", and the enriched "cellular component" terms include "ribosome", "ribonucleoprotein complex" and "small ribosomal unit". Interestingly, it has recently been shown that inhibition of essential cellular processes like translation activates pathogen defense in *C. elegans* [62] and the gram-negative pathogen *Pseudomonas aeruginosa* response in *C. elegans* is activated by detecting its inhibitory effects on translation machinery [63,64]. Downregulation of components of translational machinery and ribosomes was also observed upon germline-ablation of *P. pacificus* that exhibited an increase in lifespan as well as resistance to the pathogen *S. marcescens* [66]. Hence our data together with these recent studies suggest that that downregulation of translation machinery could be a conserved response across the two nematodes, at least when exposed to gram-negative pathogens. We also find enrichment for genes involved in "proteasome complex" and "nuclear pore" complex, and these cellular components are known to have a potential role in immune response in *C. elegans* and *P. pacificus* longevity and immunity [62,65,66]. Other enriched processes commonly affected across the two nematodes include various processes related to metabolism, such as "glycolysis", "tricarboxylic acid cycle" and "fatty acid metabolic process" (Table S8) and cellular compartment GO terms such as "mitochondrial membrane" and "mitochondrial proton-transporting ATP synthase complex" (Table S8). These results suggest that exposure to a pathogens leads to similar changes in the metabolic activity of the two nematodes.

It should also be noted that apart from the overlap between *C. elegans* and *P. pacificus*, differential expression of a substantial number of 1:1 orthologs is specific to each of the nematode species.

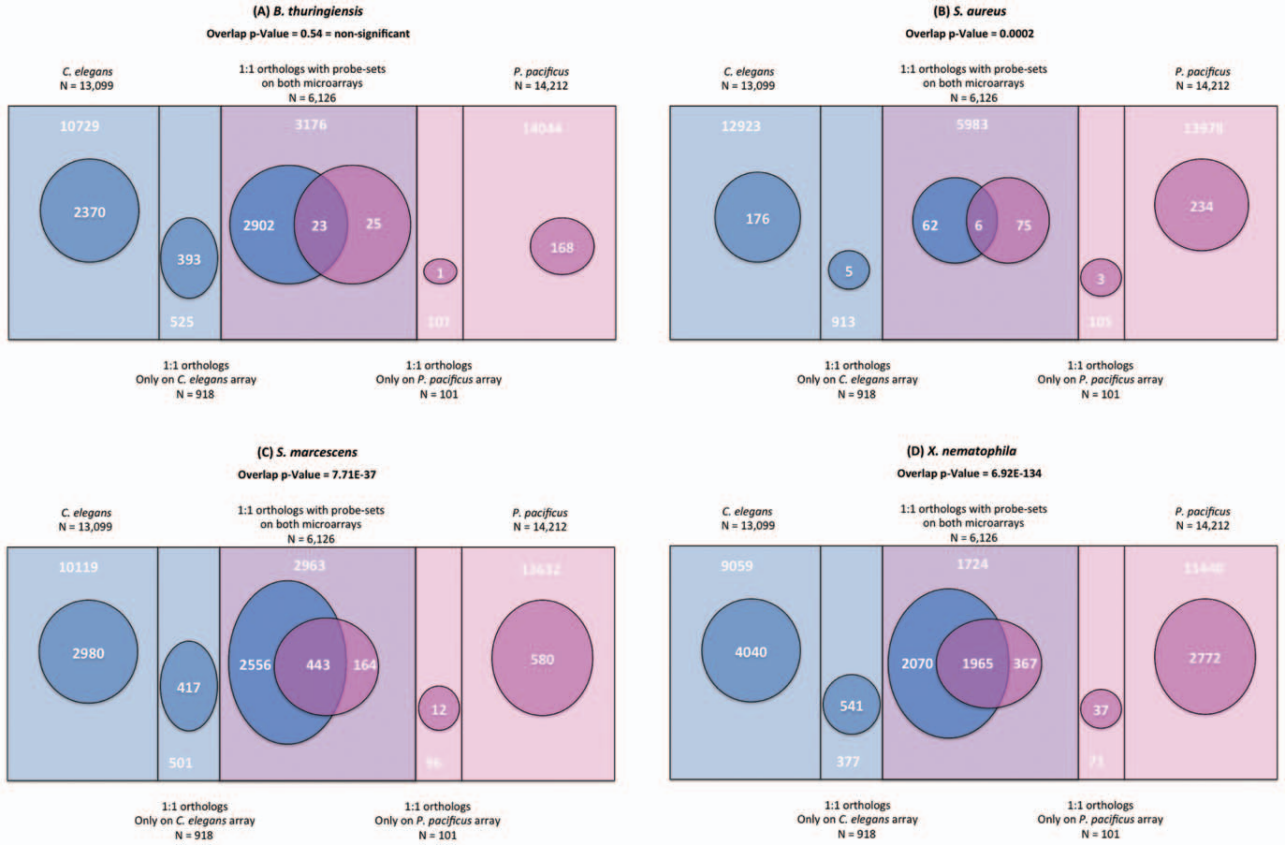


Figure 5. Overlap between *C. elegans* and *P. pacificus* expression profiles in response to different bacteria. The rectangular boxes represent the entire transcriptomes of *C. elegans* and *P. pacificus* genes assayed on our microarrays, and their area of overlap represents the set of 6,126 1:1 orthologs present on microarrays of both the nematodes. The ovals represent the fraction of differentially expressed genes in each of the subsets. For the 1:1 orthologs, we assessed the significance of overlap between the genes differentially expressed in response to a given pathogen using a 2×2 Fisher's exact test. Differences or similarities in survival characteristics of the two nematodes when exposed to the same bacteria are reflected in their respective transcriptional responses. (A) On *B. thuringiensis* DB27, which is highly lethal to *C. elegans* but not to *P. pacificus*, only 23 genes are common between the respective expression profiles of the two nematodes, this overlap being statistically not significant. (B) Similarly, *S. aureus* is more lethal to *C. elegans* than *P. pacificus*, and the overlap between the corresponding expression profiles is limited to just 6 orthologs. Although this overlap is statistically significant (p-value = 0.0002), the extent of overlap is too small to be biologically significant. (C) *S. marcescens* is lethal to both the nematodes and the extent and significance overlap between the orthologs differentially expressed in the corresponding expression profiles is also relatively high (443 common orthologs, p-value = 7.71E–37). (D) On the pathogen *X. nematophila*, the observed overlap between the expression profiles in the two nematodes is even higher, with 2,093 orthologs regulated in both nematodes (p-value = 6.92E–134).
doi:10.1371/journal.pone.0044255.g005

Additionally, the portion of the transcriptome with unresolved or no sequence similarity across the two species (Figure 5, rectangular areas specific to either *C. elegans* or *P. pacificus*) is also a significant contributor to the transcriptional response to the pathogens. Hence, based on these analysis of genes in the two nematodes exposed to the same gram-negative bacteria, it seems reasonable to conclude that some effectors of innate immunity are conserved across the two nematodes while some have diverged considerably during the last 250–300Mya separating the two nematode lineages.

Expression cluster based analysis identifies role for DAF-16, TGF-beta and p38 MAP Kinase pathways in pathogen response

To identify potential upstream regulators of immune response in the two nematodes, we tried using the existing knowledge from *C. elegans* to investigate what pathways appear to be mis-regulated. We therefore assessed the significance and the extent of overlap of our gene-sets with published microarray data sets available as "Expression Clusters" from WormBase [67] as well as with manually curated gene expression data from published microarray studies that were not available in WormBase (see Methods for details). These annotations were transferred to *P. pacificus* genes via the 1:1 orthology relations.

In agreement with the role of DAF-16 in innate immunity [68,69], we find DAF-16 targets to be enriched in genes up-regulated in *C. elegans* response to all pathogens except *B. thuringiensis* DB27 (cluster Murphy_etal_cgc5976_Class1 in Table S3B), while the DAF-16 repressed genes are over-represented in the set of down-regulated genes on *B. thuringiensis* and *S. aureus* (cluster Murphy_etal_cgc5976_Class2, Table S3B). Similarly, TGF-beta targets regulated by the ligand DBL-1 [70] are also enriched in many of our expression profiles (clusters "Roberts_etal_2010_DBL-1-UP" and "Roberts_etal_2010_DBL-1-DOWN", Table S3B), confirming an important role of TGF-beta pathway in response to specific pathogens [9,70].

Different MAP kinase pathways such as p38 MAPK and JNK pathways play a key role in *C. elegans* innate immunity and stress

response [12,49]. Consistent with this, we also see a significant overlap with genes regulated by the MAPKK SEK-1 and the JNK-like MAPK KGB-1, especially with down-regulated genes in all pathogen profiles (e.g cluster "Kao_etal2011_sek1_regulated", Table S3B). We also observe a robust induction of starvation response genes [71] (e.g. cluster "WBPaper00032948:StarveUp3", Table S3C) within just 4 hours of pathogen exposure, a reasonably short time not expected to induce actual starvation. This observation highlights the importance of metabolism related pathways in immune response [46,72–75]. We also observe enrichment of various dauer related gene clusters and other clusters regulated by stress such as heat shock and oxidative stress in some of the *C. elegans* profiles (Table S3C).

Somewhat similar patterns of overlaps with *C. elegans* expression clusters are also seen for *P. pacificus* pathogen response profiles (Table S4). Interestingly, unlike *C. elegans*, the clusters of genes regulated in response to Cry5B toxin and KGB-1 or SEK-1 MAP kinases show a significant overlap only with genes down-regulated upon exposure of *P. pacificus* to *S. marcescens* and *X. nematophila* (e.g. cluster "Kao_etal2011_sek1_regulated", Table S4B) but not with genes up-regulated in *P. pacificus*. This suggests potential differences either in targets of the MAPK pathways, or differences in mechanism of activation of these pathways, which might account for enhanced resistance of *P. pacificus*.

C. elegans DAF-16 targets are enriched in some *P. pacificus* profiles (cluster "Murphy_etal_cgc5976_Class2" in Table S4B), suggesting that DAF-16 might have a conserved role in innate response in both the species, at least on some pathogens. Interestingly, compared to *C. elegans* profiles, all *P. pacificus* profiles show a significant and more extensive overlap with genes involved in osmotic stress response [76] (e.g. cluster "WBPaper00035873:-osmotically_regulated", Table S4C), suggesting that osmotic regulation could be an important survival mechanism against potentially pathogenic bacteria [76].

Differential expression of *P. pacificus* pioneer genes

About 30% of the predicted transcriptome of *P. pacificus* is comprised of "pioneer genes", which do not show any detectable homology to the known protein universe [19] and whose functions are not known. We investigated their potential role in pathogen response by looking at their expression data. We indeed find 832 of these pioneer genes to be differentially expressed *P. pacificus* in a pathogen specific manner, with 160 genes being regulated on at least two pathogens (Figure S3 and Table S9). On each of the pathogens, the pioneer genes constitute 12% to 18% of the active transcriptome, significantly less than the expected proportion of about 30% (Fisher's 2×2 exact test p –values$<$0.001, Figure S4). Interestingly however, we find the differential expression levels of these pioneer genes to be significantly higher than the non-pioneer fraction of the respective transcriptomes (Figures S5A, S5B, S5C and S5D, Kolmogorov-Smirnov test p-values less than 2.00E–16 for all four pathogens), indicating specific increase in their expression levels after exposure to pathogens. These observations together suggest that some of these lineage specific genes might have been acquired for adaptation to a microenvironment populated by different set of bacteria, some of which might be pathogenic. We can thus ascribe a putative role for these pioneer genes in pathogen response, although further studies will be needed to test these predictions.

An ecological perspective on the evolution of effectors and regulators of nematode immunity

Our finding that *C. elegans* mounts a pathogen-specific transcriptional response is in agreement with the current understanding in the field [46,47,58]. We further show for the first time that the nematode *P. pacificus* can also activate a pathogen-specific response. Many evolutionary mechanisms contribute towards generating this specificity in invertebrates, which lack an adaptive immune system. These include high genetic diversity receptors and effectors involved in pathogen recognition [77], evolutionary diversification of innate immunity effectors e.g. C-type lectins [57], lysozymes [78] and *nlp*- family of antimicrobial effectors [79], natural variation in host susceptibility and virulence of the pathogen [35,38,80], and evolution of mechanisms such as recombination and sexual reproduction [81] or alternative splicing [82,83], all of which facilitate generation of genetic diversity.

The ecology of the organism is expected to be one of the key driving forces behind these changes, as the related species or even strains that occupy different ecological niches will be exposed to different set of non-pathogenic and pathogenic microbes and will need different strategies to survive. Due to these differences in selective pressures, the effectors of their immune systems can be expected to diverge rapidly and also affect the evolution of the host genomes.

The differences observed between response of *C. elegans* and *P. pacificus* can thus be best explained in the light of the distinct ecological niches occupied by both species. While *C. elegans* has recently been isolated from rotting fruit [28], *Pristionchus* nematodes and *P. pacificus* have a strong association with scarab beetles [29–32]. Once the beetle dies bacteria proliferate on the rotting carcass allowing mass growth of *Pristionchus* nematodes. Using a metagenomic approach we have previously shown that hundreds of plant and animal pathogenic bacteria occur on and in *Pristionchus* nematodes emerging from beetles [34]. Thus, *Pristionchus* is naturally exposed to a variety of bacteria and has evolved mechanisms to combat infections. Relative to *C. elegans*, the *P. pacificus* genome contains a larger set of genes encoding for cytochrome P450 and UDP-glucoronosyl/UDP-glucosyl transferases, which are required for coping with xenobiotic compounds [18], and we show here that differential activation of these gene families might contribute to its higher resistance to pathogens. The limited overlap on a gene-by-gene between expression profiles on different pathogens combined with the observation that similar Pfam domains are enriched within a nematode species, are consistent with evolutionary diversification and expansion of genes containing these functional domains.

Conclusions

This study provides a system wide analysis of the transcriptomic responses of the two nematode model species *C. elegans* and *P. pacificus* when feeding on four well-characterized bacterial pathogens. Studies on natural variation in the response of *C. elegans* to pathogens have contributed to micro-evolutionary studies of evolution of innate immunity. By adding studies on host-pathogen interactions in *P. pacificus*, we have tried to extend the evolution of innate immunity towards a macro-evolutionary perspective. We have studied nematode response as early as four hours after exposure to bacteria in order to capture initial events. While many previous studies have looked at various time points (4 h, 8 h, 24 h etc), it is known that by this time a common host necrotic response sets in [46]. The data generated for *P. pacificus* is the first of its kind, whereas our *C. elegans* dataset overlaps, in part, with previous studies. We performed these experiments *de novo*, rather than taking data from the literature, to rule out the effect of differences due to the microarray platform. Our *C. elegans* dataset is however, in strong agreement with existing datasets (e.g. [44,46,47]. This study fulfilled three major aims. First, it presents a platform for comparative systems biology analysis of two nematode model

species. Second, it generates a catalog of genes involved in the evolution of nematode immunity and finally, it identifies pathogen-specific as well as pan-pathogen, conserved responses in the two nematode species.

Research on *C. elegans* and its interactions with bacteria has lead to the identification of several pathways involved in innate immunity [84]. By using an alternative nematode model we have expanded on this knowledge and shown conservation as well as divergence in the transcripts regulated during an immune response when fed different pathogens. Our systematic comparisons of nematode survival and gene expression on multiple pathogens highlights the fact that substantial differences exist in the repertoire of genetic components deployed in response to varied pathogens between *C. elegans* and *P. pacificus*. The resulting catalogs of pathogen specific and pan-pathogen genes provide an entry point to study the mechanism and evolution of individual response genes in future studies. Using expression cluster analysis we could show that homologs of known targets of FOXO/DAF-16, TGF-beta and p38 MAP kinase pathways in *C. elegans* are also significantly enriched in *P. pacificus*, suggesting that the key signaling pathways might regulate innate immunity in both the species. Given the lack of corresponding mutants in the relatively new model system *P. pacificus*, future studies are needed to test this hypothesis.

Evolutionary studies will have to involve more closely related species and strains given the strong differences observed for *C. elegans* and *P. pacificus*. From a *P. pacificus* perspective, more careful analyses of individual genes in additional strains and closely related *Pristionchus* species will be necessary to obtain insight into the evolution of immunity-related gens. Finally, these results argue for the importance of a comparative approach towards uncovering mechanistic details of the genetic basis that accounts for the cross-species variation in susceptibility to a given pathogen.

Materials and Methods

Strains

C. elegans N2 and *P. pacificus* RS2333 were maintained on NGM plates seeded with *E. coli* OP50. *S. marcescens* was isolated from La Reunion, *S. aureus* Newman was purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, Germany (DSMZ), *X. nematophila* was a gift from Becker Underwood, U.K. *B. thuringiensis* DB27 was isolated from dung beetles [35] and was initially thought to be a strain of *B. cereus* but further sequence analysis has shown it is in fact a strain of closely related *B. thuringiensis*.

Assessing survival of *C. elegans* and *P. pacificus* exposed to pathogens

Each bacterium was grown in a shaking incubator at 30°C overnight in LB, apart from *S. aureus*, which was grown at 37°C. The following day 100 µl were spread onto previously dried 6 cm NGM plates and incubated overnight. Twenty adults of C. elegans or *P. pacificus* were separately placed onto 3 NGM plates per bacterium and stored at 25°C where survival was monitored daily for 7 days. The experiment was repeated twice. Pathogen survival was compared to worms cultured on *E. coli* control plates and differences in survival were analyzed using Kaplan Meier and logrank test.

RNA collection for microarray experiments

Synchronized populations of *C. elegans* or *P. pacificus* were obtained by hypochlorite treatment and allowed to grow to young adult stage on *E. coli* OP50 plates at 20°C. For each biological replicate, about 200 young adult hermaphrodites were picked onto pathogen plates for 4 hours of pathogen exposure, after which they were collected into 1mL of TRIzol (Invitrogen). Equal number of corresponding age-matched control worms were exposed to *E. coli* OP50 for 4 hours and transferred to 1ml TRIzol (Invitrogen). Four biological replicates were collected for each experimental and control condition. Total RNA was extracted using TRIzol reagent (Invitrogen) according to manufacturer's instructions. The isolated RNA was further purified using phenol: chloroform: isoamyl alcohol precipitation to remove trace left-overs of TRIzol etc. which might interfere with downstream reactions. The RNA pellet was suspended in RNAse free water and was assessed on a Nanodrop spectrophotometer for quantity and RNA quality. RNA samples were stored at −80°C until the microarray experiments.

Microarray experiments

A total of 32 microarray hybridizations were carried out for 8 conditions (2 nematode species x 4 pathogenic bacteria; 4 biological replicates per condition). Oligonucleotide microarrays for *C. elegans* containing ~43,000 unique probes for ~20,000 *C. elegans* genes were obtained from Agilent Technologies (NCBI GEO accession GPL10094). For *P. pacificus* experiments, we used our custom designed oligonucleotide microarrays manufactured by Agilent Technologies, which contain ~93,000 unique probes for the ~23,000 *P. pacificus* predicted genes (NCBI GEO accession GPL14372, see [22] for design details of custom microarrays). The *P. pacificus* gene sequences are available at http://www.pristionchus.org/download/ and the gene models can be seen in the genome browser at http://www.pristionchus.org/cgi-bin/genome.pl.

Equal amounts of total RNA (500 ng to 800 ng) from four biological replicates of each experimental and control samples was used to produce Cy5 or Cy3 dye labeled cRNA using Quick Amp Labelling Kit (Agilent Technologies Inc., USA) as per manufacturer's instructions. Depending upon the amount of total RNA used, appropriate amounts of positive control RNA (Spike Mix-A and Spike Mix-B, from Agilent Technologies) was added to the mix before reverse transcribing the total RNA, as per manufacturer's instructions. We used the *C. elegans* or *P. pacificus* microarrays in a two-color format where Cy5 and Cy3 dye labeled cRNA from experimental and control sample is co-hybridized on the same microarray. The four biological replicates per experiment included two dye-swaps experiments to account for differences in dye labeling. Hybridization and washing of the arrays was carried according to manufacturer-supplied protocol. The arrays were scanned on a GenePix 4000B Microarray Scanner, and raw data extracted using GenePix Pro sofware (version 6).

Microarray data analysis

We used the Bioconductor [85] package limma [86] for analysis of our microarray data. Array quality was checked for parameters such as uniform background and foreground intensities over the entire array. The raw signal was background corrected using the normexp method [87] and the arrays were then lowess normalized individually ("normalizeWithinArrays" option), with differential weights assigned to probes and to positive control spike-ins, which are expected to show no fold change [88]. This differential weighing of probes is particularly necessary to account for differences in differences in relative proportion of mRNA versus total RNA, and/or differences in the amount of RNA produced per worm under different experimental conditions. Without this differential weighing scheme, the fold change calculations can be erroneous [22,42]. The weight parameters were optimized based on MA-plots such that spike-in controls show their expected fold change values. lmFit function was used to fit a linear model to

probe intensities across arrays, differential expression was calculated by empirical Bayes method using the eBayes function [89], and control of FDR was employed as the multiple testing correction. MA-plots were also used as diagnostic to identify and remove outlier arrays before fold-change calculations, such that at least three biological replicates were used for each experiment. Genes with a FDR corrected p-value less or equal to than 0.05 and absolute log2 of fold change greater than 0.5 were called significantly differentially expressed. Further data analysis was carried out using custom scripts in Perl and the statistical package R. Venn diagrams were drawn using the R package VennDiagram [90]. Raw and processed data from all the experiments from this publication have been deposited in a MIAME compliant format [91] at NCBI's Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/). The accession numbers for *C. elegans* data are GSE36413, GSE36493, GSE36499, GSE36501, and the accession numbers for *P. pacificus* data are GSE36517, GSE36519, GSE36521 and GSE36523.

Identifying 1:1 orthologs between *C. elegans* and *P. pacificus*

We have previously used a pairwise best BLASTP strategy to identify 7,176 pairs of 1:1 orthologs between the *C. elegans* and *P. pacificus* [22]. Briefly, all protein sequences from *C. elegans* were run as query versus the database of *P. pacificus* gene predictions and vice versa. Only hits with a BLAST score $>=50$ bits were retained, and mutually best hits were identified as 1:1 orthologs. Probes for 6,126 of these gene pairs exist on microarrays of both the species.

Pfam domain annotation and enrichment analysis

Pfam domain annotations for *C. elegans* and *P. pacificus* proteomes (WS220 and predicted proteins respectively) were the same as described before [22]. Basically, hits with a p-value cut-off of 0.001 in HMM searches using HMMer 3.0 [92] on Pfam release V23/4 [52] were used as domain annotations. Only the domains, for which minimum 5 protein coding genes were represented on each microarray, were used for further enrichment analysis. Statistical significance of enrichment of Pfam domains in each expression profile determined using a 2×2 Fishers exact test, at a FDR corrected p-value cut-off of 0.05.

Expression cluster enrichment analysis

We have used "expression cluster" annotations from WormBase [67] in interpretations of other microarray expression profiles [22]. The list of microarray expression profiles in which a given *C. elegans* gene is known to be differentially expressed can be extracted from the section "Expression Cluster" from the WormBase gene summary page for each gene. We retrieved all available expression clusters for *C. elegans* genes from the WormBase web site. We also compiled data from other gene expression studies in *C. elegans* which are relevant to pathogen response but for which the corresponding expression clusters were not available in the WormBase (viz. [12,49,70]) and named these clusters with a prefix based on first author's last name and year of publication of the research article, and included them in our analysis. We inferred expression clusters for *P. pacificus* based on the set of 1:1 orthologs. P-values for expression cluster enrichment in each expression profile was computed with a 2×2 Fisher exact test. FDR corrected p-value cut-off of 0.05 was used as the significance threshold. The significance score was calculated as –log10 of the p-values and was set to zero to indicate non-significance when p-values was greater than 0.05.

Gene ontology analysis, prediction of signal peptide and antimicrobial activity

Gene ontology analysis (presented in Table S8) was done using Bioconductor tool topGO, using method "elimFisher" for calculating p-values [93]. For analyzing features of differentially expressed pioneer genes in *P. pacificus* (Table S9), SignalP tool was used to predict the presence of a signal peptide [94], and for genes coding for products smaller than 100 amino acids, CAMP tool [95] was used to predict whether they can act as potential Anti-Microbial Peptides (AMPs).

Supporting Information

Figure S1 Long-term survival curves for *P. pacificus* on various pathogens. *P. pacificus* has higher resistance than *C. elegans*, with longer median survival time of about 8 days on *B. thuringiensis* and *S. aureus*.
(PDF)

Figure S2 Global transcriptional suppression in response to pathogens. Exposure to pathogens resulted in a decrease of amount of total RNA produced per worm when compared to the non-pathogenic *E. coli* strain. This global decrease in transcription is seen in both (A) *C. elegans* and (B) *P. pacificus*.
(PDF)

Figure S3 Overlap between pioneer genes regulated in *P. pacificus* in response to the four pathogens. Of the 832 pioneer genes differentially expressed on any of the pathogens in *P. pacificus*, 160 genes are common between two or more than two expression profiles.
(PDF)

Figure S4 Relative proportions of pioneer genes versus non-pioneer genes in the active transcriptome of *P. pacificus* on each of the four pathogens. The *P. pacificus* genome contains about 30% pioneer genes. Compared to the random expectation of the same proportion of pioneer genes in different expression profiles, they are found to significantly under-represented (Fisher's test p-values$<$2E-16 for each comparison with whole-genome distribution.
(PDF)

Figure S5 Pioneer genes are expressed at higher levels than non-pioneer genes in each of the pathogen-induced expression profiles on *P. pacificus*. Cumulative distributions of fold-changes of pioneer genes (red curves) and non-pioneer genes (blue curves) for genes differentially expressed on (A) *B. thuringiensis* DB27 (B) *S. aureus* (C) *S. marcescens* and (D) *X. nematophila*. The Kolmogorov-Smirnov test p-values are less that 2E-16 in each case.
(PDF)

Table S1 Differential expression of genes in *C. elegans* exposed to different pathogens.
(XLSX)

Table S2 Differential expression of genes in *P. pacificus* exposed to different pathogens.
(XLSX)

Table S3 Expression clusters enriched in genes up- and down-regulated in *C. elegans* upon exposure to each of the four pathogens.
(XLSX)

Table S4 Expression clusters enriched in genes up- and down-regulated in *P. pacificus* upon exposure to each of the four pathogens.
(XLSX)

Table S5 Pfam domains enriched in *C. elegans* expression profiles in response to different pathogens.
(XLSX)

Table S6 Pfam domains enriched in *P. pacificus* expression profiles in response to different pathogens.
(XLSX)

Table S7 Genes common across expression profiles of *C. elegans* and *P. pacificus* upon exposure to the gram-negative bacteria (A) *S. marcescens* (B) *X. nematophila*, and (C) common to both *S. marcescens* and *X. nematophila* expression profiles.
(XLSX)

Table S8 GO enrichment analysis of genes common across both *C. elegans* and *P. pacificus* profiles obtained upon exposure to gram-negative bacteria *S. marcescens* and *X. nematophila*.
(XLSX)

Table S9 Features of pioneer genes differentially expressed on any of the four pathogens in *P. pacificus*.
(XLSX)

Acknowledgments

We thank members of the lab for discussion and Dr. Mitch Levesque for help with the microarray platform.

Author Contributions

Conceived and designed the experiments: AS RR II RJS. Performed the experiments: AS RR II. Analyzed the data: AS. Contributed reagents/materials/analysis tools: AS RR II. Wrote the paper: AS RR RJS.

References

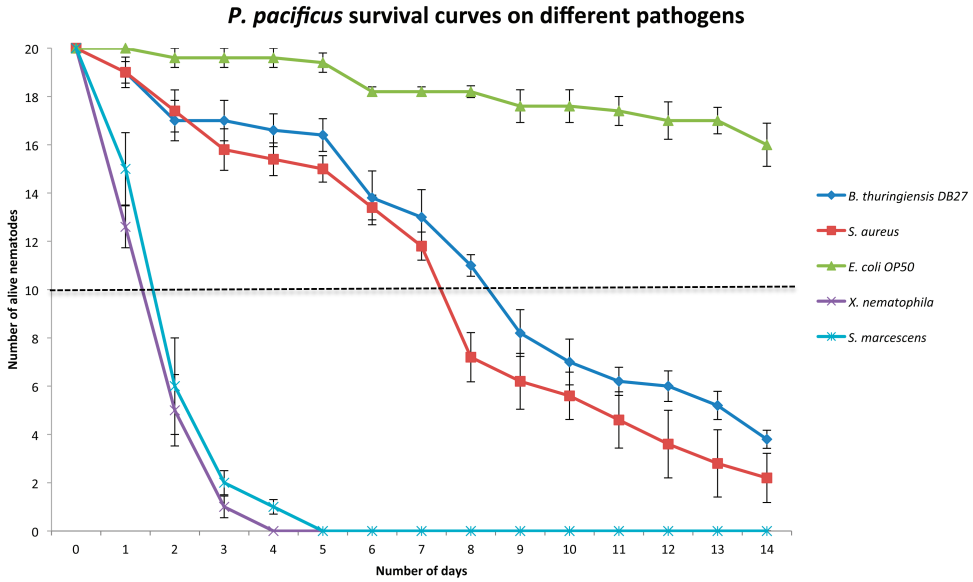
1. Haldane JBS (1949) Disease and Evolution. La Ricerca Scientifica 19: 68–76.
2. Thompson JN (1994) The coevolutionary process. University of Chicago Press. 398 p.
3. Woolhouse MEJ, Webster JP, Domingo E, Charlesworth B, Levin BR (2002) Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nat Genet 32: 569–577. doi:10.1038/ng1202-569.
4. Frank SA (2002) Immunology and Evolution of Infectious Disease. Princeton (NJ): Princeton University Press; Available: http://www.ncbi.nlm.nih.gov/books/NBK2394/.
5. Griffitts JS, Whitacre JL, Stevens DE, Aroian RV (2001) Bt toxin resistance from loss of a putative carbohydrate-modifying enzyme. Science 293: 860–864. doi:10.1126/science.1062441.
6. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, et al. (2002) A conserved p38 MAP kinase pathway in Caenorhabditis elegans innate immunity. Science 297: 623–626. doi:10.1126/science.1073759.
7. Tan MW, Rahme LG, Sternberg JA, Tompkins RG, Ausubel FM (1999) Pseudomonas aeruginosa killing of Caenorhabditis elegans used to identify P. aeruginosa virulence factors. Proc Natl Acad Sci USA 96: 2408–2413.
8. Kurz CL, Chauvet S, Andrès E, Aurouze M, Vallet I, et al. (2003) Virulence factors of the human opportunistic pathogen Serratia marcescens identified by in vivo screening. EMBO J 22: 1451–1460. doi:10.1093/emboj/cdg159.
9. Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, et al. (2002) Inducible Antibacterial Defense System in C. elegans. Current Biology 12: 1209–1214. doi:10.1016/S0960-9822(02)00928-4.
10. Aballay A, Ausubel FM (2001) Programmed cell death mediated by ced-3 and ced-4 protects Caenorhabditis elegans from Salmonella typhimurium-mediated killing. Proc Natl Acad Sci USA 98: 2735–2739. doi:10.1073/pnas.041613098.
11. Garsin DA, Villanueva JM, Begun J, Kim DH, Sifri CD, et al. (2003) Long-lived C. elegans daf-2 mutants are resistant to bacterial pathogens. Science 300: 1921. doi:10.1126/science.1080147.
12. Huffman DL, Abrami L, Sasik R, Corbeil J, van der Goot FG, et al. (2004) Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. Proc Natl Acad Sci USA 101: 10995–11000. doi:10.1073/pnas.0404073101.
13. Nicholas HR, Hodgkin J (2004) The ERK MAP kinase cascade mediates tail swelling and a protective response to rectal infection in C. elegans. Curr Biol 14: 1256–1261. doi:10.1016/j.cub.2004.07.022.
14. Powell JR, Kim DH, Ausubel FM (2009) The G protein-coupled receptor FSHR-1 is required for the Caenorhabditis elegans innate immune response. Proceedings of the National Academy of Sciences 106: 2782 –2787. doi:10.1073/pnas.0813048106.
15. Estes KA, Dunbar TL, Powell JR, Ausubel FM, Troemel ER (2010) bZIP transcription factor zip-2 mediates an early response to Pseudomonas aeruginosa infection in Caenorhabditis elegans. Proceedings of the National Academy of Sciences 107: 2153 –2158. doi:10.1073/pnas.0914643107.
16. Irazoqui JE, Ng A, Xavier RJ, Ausubel FM (2008) Role for beta-catenin and HOX transcription factors in Caenorhabditis elegans and mammalian host epithelial-pathogen interactions. Proc Natl Acad Sci USA 105: 17469–17474. doi:10.1073/pnas.0809527105.
17. Hong RL, Sommer RJ (2006) Pristionchus pacificus: a well-rounded nematode. Bioessays 28: 651–659. doi:10.1002/bies.20404.
18. Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, et al. (2008) The Pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism. Nat Genet 40: 1193–1198. doi:10.1038/ng.227.
19. Borchert N, Dieterich C, Krug K, Schütz W, Jung S, et al. (2010) Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models. Genome Res 20: 837–846. doi:10.1101/gr.103119.109.
20. Pires da Silva A (2005) Pristionchus pacificus genetic protocols. WormBook: 1–8. doi:10.1895/wormbook.1.114.1.
21. Schlager B, Wang X, Braach G, Sommer RJ (2009) Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode Pristionchus pacificus. Genesis 47: 300–304. doi:10.1002/dvg.20499.
22. Sinha A, Sommer RJ, Dieterich C (2012) Divergent gene expression in the conserved dauer stage of the nematodes Pristionchus pacificus and Caenorhabditis elegans. BMC Genomics 13: 254. doi:10.1186/1471-2164-13-254.
23. Tian H, Schlager B, Xiao H, Sommer RJ (2008) Wnt signaling induces vulva development in the nematode Pristionchus pacificus. Curr Biol 18: 142–146. doi:S0960–9822(07)02479–7.
24. Wang X, Sommer RJ (2011) Antagonism of LIN-17/Frizzled and LIN-18/Ryk in Nematode Vulva Induction Reveals Evolutionary Alterations in Core Developmental Pathways. PLoS Biol 9: e1001110. doi:10.1371/journal.-pbio.1001110.
25. Rudel D, Riebesell M, Sommer RJ (2005) Gonadogenesis in Pristionchus pacificus and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad. Dev Biol 277: 200–221. doi:S0012-1606(04)00658-X.
26. Pires-daSilva A, Sommer RJ (2004) Conservation of the global sex determination gene tra-1 in distantly related nematodes. Genes Dev 18: 1198–1208. doi:15155582.
27. Hong RL, Witte H, Sommer RJ (2008) Natural variation in Pristionchus pacificus insect pheromone attraction involves the protein kinase EGL–4. Proc Natl Acad Sci USA 105: 7779–7784. doi:0708406105.
28. Félix M-A, Braendle C (2010) The natural history of Caenorhabditis elegans. Curr Biol 20: R965–969. doi:10.1016/j.cub.2010.09.050.
29. Herrmann M, Mayer WE, Sommer RJ (2006) Nematodes of the genus Pristionchus are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. Zoology (Jena) 109: 96–108. doi:S0944-2006(06)00006-7.
30. Herrmann M, Mayer WE, Hong RL, Kienle S, Minasaki R, et al. (2007) The nematode Pristionchus pacificus (Nematoda: Diplogastridae) is associated with the oriental beetle Exomala orientalis (Coleoptera: Scarabaeidae) in Japan. Zool Sci 24: 883–889. doi:0289-0003-24-9-883.
31. Herrmann M, Kienle S, Rochat J, Mayer WE, Sommer RJ (2010) Haplotype diversity of the nematode Pristionchus pacificus on Réunion in the Indian Ocean suggests multiple independent invasions. Biological Journal of the Linnean Society 100: 170–179. doi:10.1111/j.1095-8312.2010.01410.x.
32. Morgan K, McGaughran A, Villate L, Herrmann M, Witte H, et al. (2012) Multi locus analysis of Pristionchus pacificus on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. Mol Ecol 21: 250–266. doi:10.1111/j.1365-294X.2011.05382.x.
33. Wei J-Z, Hale K, Carta L, Platzer E, Wong C, et al. (2003) Bacillus thuringiensis crystal proteins that target nematodes. Proc Natl Acad Sci USA 100: 2760–2765. doi:10.1073/pnas.0538072100.
34. Rae R, Riebesell M, Dinkelacker I, Wang Q, Herrmann M, et al. (2008) Isolation of naturally associated bacteria of necromenic Pristionchus nematodes and fitness consequences. J Exp Biol 211: 1927–1936. doi:211/12/1927.
35. Rae R, Iatsenko I, Witte H, Sommer RJ (2010) A subset of naturally isolated Bacillus strains show extreme virulence to the free-living nematodes Caenorhabditis elegans and Pristionchus pacificus. Environ Microbiol 12: 3007–3021. doi:10.1111/j.1462-2920.2010.02278.x.
36. Altun, Z.F., Herndon, L.A., Crocker, C., Lints, R. and Hall, D.H. (ed.s) (2002–2012) WormAtlas. Available: http://www.wormatlas.org.

37. Sommer RJ, Carta LK, Kim SY, Sternberg PW (1996) Morphological, genetic and molecular description of Pristionchus pacificus sp. n. (Nematoda: Neodiplogastridae). Fundamental and Applied Nematology 19: 511–522.
38. Schulenburg H, Ewbank JJ (2004) Diversity and specificity in the interaction between Caenorhabditis elegans and the pathogen Serratia marcescens. BMC Evol Biol 4: 49. doi:10.1186/1471-2148-4-49.
39. Forst S, Clarke D (2002) Bacteria-Nematode Symbiosis. In: Gaugler R, editor. Entomopathogenic nematology. CABI. 57–78.
40. Hsin H, Kenyon C (1999) Signals from the reproductive system regulate the lifespan of C. elegans. Nature 399: 362–366. doi:10.1038/20694.
41. Dalley BK, Golomb M (1992) Gene expression in the Caenorhabditis elegans dauer larva: developmental regulation of Hsp90 and other genes. Dev Biol 151: 80–90.
42. van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, et al. (2003) Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep 4: 387–393. doi:10.1038/sj.embor.embor798.
43. Shapira M, Brigham BJ, Rong J, Chen K, Ronen M, et al. (2006) A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. Proceedings of the National Academy of Sciences 103: 14086–14091. doi:10.1073/pnas.0603424103.
44. Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, et al. (2006) p38 MAPK Regulates Expression of Immune Response Genes and Contributes to Longevity in C. elegans. PLoS Genet 2: e183. doi:10.1371/journal.pgen.0020183.
45. O'Rourke D, Baban D, Demidova M, Mott R, Hodgkin J (2006) Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of C. elegans with M. nematophilum. Genome Res 16: 1005–1016. doi:10.1101/gr.50823006.
46. Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank J (2007) Genome-wide investigation reveals pathogen-specific and shared signatures in the response of Caenorhabditis elegans to infection. Genome Biology 8: R194. doi:10.1186/gb-2007-8-9-r194.
47. Irazoqui JE, Troemel ER, Feinbaum RL, Luhachack LG, Cezairliyan BO, et al. (2010) Distinct Pathogenesis and Host Responses during Infection of C. elegans by P. aeruginosa and S. aureus. PLoS Pathog 6: e1000982. doi:10.1371/journal.at.1000982.
48. Bolz DD, Tenor JL, Aballay A (2010) A conserved PMK-1/p38 MAPK is required in caenorhabditis elegans tissue-specific immune response to Yersinia pestis infection. J Biol Chem 285: 10832–10840. doi:10.1074/jbc.M109.091629.
49. Kao C-Y, Los FCO, Huffman DL, Wachi S, Kloft N, et al. (2011) Global functional analyses of cellular responses to pore-forming toxins. PLoS Pathog 7: e1001314. doi:10.1371/journal. at.1001314.
50. Pukkila-Worley R, Ausubel FM, Mylonakis E (2011) Candida albicans Infection of Caenorhabditis elegans Induces Antifungal Immune Defenses. PLoS Pathog 7: e1002074. doi:10.1371/journal. at.1002074.
51. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. PNAS 95: 14863–14868.
52. Finn RD, Mistry J, Tate J, Coggill P, Heger A, et al. (2009) The Pfam protein families database. Nucleic Acids Research 38: D211–D222. doi:10.1093/nar/gkp985.
53. Nandakumar M, Tan M-W (2008) Gamma-linolenic and stearidonic acids are required for basal immunity in Caenorhabditis elegans through their effects on p38 MAP kinase activity. PLoS Genet 4: e1000273. doi:10.1371/journal.pgen.1000273.
54. Schulenburg H, Kurz CL, Ewbank JJ (2004) Evolution of the innate immune system: the worm perspective. Immunol Rev 198: 36–58.
55. De Gregorio E, Spellman PT, Rubin GM, Lemaitre B (2001) Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc Natl Acad Sci USA 98: 12590–12595. doi:10.1073/pnas.221458698.
56. Irving P, Troxler L, Heuer TS, Belvin M, Kopczynski C, et al. (2001) A genome-wide analysis of immune responses in Drosophila. Proc Natl Acad Sci USA 98: 15119–15124. doi:10.1073/pnas.261573998.
57. Schulenburg H, Hoeppner MP, Weiner J III, Bornberg-Bauer E (2008) Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode Caenorhabditis elegans. Immunobiology 213: 237–250. doi:10.1016/j.imbio.2007.12.004.
58. Engelmann I, Griffon A, Tichit L, Montañana-Sanchis F, Wang G, et al. (2011) A Comprehensive Analysis of Gene Expression Changes Provoked by Bacterial and Fungal Infection in C. elegans. PLoS ONE 6: e19055. doi:10.1371/journal.pone.0019055.
59. Reichert K, Menzel R (2005) Expression profiling of five different xenobiotics using a Caenorhabditis elegans whole genome microarray. Chemosphere 61: 229–237. doi:10.1016/j.chemosphere.2005.01.077.
60. Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development 131: 311–323. doi:10.1242/dev.00914.
61. Kim KW, Wilson TL, Kimble J (2010) GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. Proc Natl Acad Sci USA 107: 17445–17450. doi:10.1073/pnas.1012611107.
62. Melo JA, Ruvkun G (2012) Inactivation of Conserved C. elegans Genes Engages Pathogen- and Xenobiotic-Associated Defenses. Cell 149: 452–466. doi:10.1016/j.cell.2012.02.050.
63. McEwan DL, Kirienko NV, Ausubel FM (2012) Host Translational Inhibition by Pseudomonas aeruginosa Exotoxin A Triggers an Immune Response in Caenorhabditis elegans. Cell Host Microbe 11: 364–374. doi:10.1016/j.chom.2012.02.007.
64. Dunbar TL, Yan Z, Balla KM, Smelkinson MG, Troemel ER (2012) C. elegans Detects Pathogen-Induced Translational Inhibition to Activate Immune Signaling. Cell Host Microbe 11: 375–386. doi:10.1016/j.chom.2012.02.008.
65. Alper S, Laws R, Lackford B, Boyd WA, Dunlap P, et al. (2008) Identification of innate immunity genes and pathways using a comparative genomics approach. Proc Natl Acad Sci USA 105: 7016–7021. doi:10.1073/pnas.0802405105.
66. Rae R, Sinha A, Sommer RJ (2012) Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode Pristionchus pacificus. PLoS Pathogens 8: e1002864. doi:10.1371/journal.at.1002864.
67. Harris TW, Antoshechkin I, Bieri T, Blasiar D, Chan J, et al. (2010) WormBase: a comprehensive resource for nematode research. Nucleic Acids Res 38: D463–467. doi:10.1093/nar/gkp952.
68. Murphy CT, McCarroll SA, Bargmann CI, Fraser A, Kamath RS, et al. (2003) Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature 424: 277–283. doi:10.1038/nature01789.
69. Lee SS, Kennedy S, Tolonen AC, Ruvkun G (2003) DAF-16 Target Genes That Control C. elegans Life-Span and Metabolism. Science 300: 644–647. doi:10.1126/science.1083614.
70. Roberts AF, Gumienny TL, Gleason RJ, Wang H, Padgett RW (2010) Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in Caenorhabditis elegans. BMC Dev Biol 10: 61. doi:10.1186/1471-213X-10-61
71. Baugh LR, Demodena J, Sternberg PW (2009) RNA Pol II accumulates at promoters of growth genes during developmental arrest. Science 324: 92–94. doi:10.1126/science.1169628.
72. Coolon JD, Jones KL, Todd TC, Carr BC, Herman MA (2009) Caenorhabditis elegans genomic response to soil bacteria predicts environment-specific genetic effects on life history traits. PLoS Genet 5: e1000503. doi:10.1371/journal.pgen.1000503.
73. Anagnostou SH, Shepherd PR (2008) Glucose induces an autocrine activation of the Wnt/beta-catenin pathway in macrophage cell lines. Biochem J 416: 211–218. doi:10.1042/BJ20081426.
74. Bensinger SJ, Tontonoz P (2008) Integration of metabolism and inflammation by lipid-activated nuclear receptors. Nature 454: 470–477. doi:10.1038/nature07202.
75. Pearce EL, Walsh MC, Cejas PJ, Harms GM, Shen H, et al. (2009) Enhancing CD8 T-cell memory by modulating fatty acid metabolism. Nature 460: 103–107. doi:10.1038/nature08097.
76. Rohlfing A-K, Miteva Y, Hannenhalli S, Lamitina T (2010) Genetic and Physiological Activation of Osmosensitive Gene Expression Mimics Transcriptional Signatures of Pathogen Infection in C. elegans. PLoS ONE 5: e9010. doi:10.1371/journal.pone.0009010.
77. Schulenburg H, Boehnisch C, Michiels NK (2007) How do invertebrates generate a highly specific innate immune response? Mol Immunol 44: 3338–3344. doi:10.1016/j.molimm.2007.02.019.
78. Schulenburg H, Boehnisch C (2008) Diversification and adaptive sequence evolution of Caenorhabditis lysozymes (Nematoda: Rhabditidae). BMC Evol Biol 8: 114. doi:10.1186/1471-2148-8-114.
79. Pujol N, Zugasti O, Wong D, Couillault C, Kurz CL, et al. (2008) Anti-Fungal Innate Immunity in C. elegans Is Enhanced by Evolutionary Diversification of Antimicrobial Peptides. PLoS Pathog 4: e1000105. doi:10.1371/journal.at.1000105.
80. Félix M-A, Ashe A, Piffaretti J, Wu G, Nuez I, et al. (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. PLoS Biol 9: e1000586. doi:10.1371/journal.pbio.1000586.
81. Hamilton WD, Axelrod R, Tanese R (1990) Sexual reproduction as an adaptation to resist parasites (a review). Proc Natl Acad Sci USA 87: 3566–3573.
82. Watson FL, Püttmann-Holgado R, Thomas F, Lamar DL, Hughes M, et al. (2005) Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309: 1874–1878. doi:10.1126/science.1116887.
83. Dong Y, Taylor HE, Dimopoulos G (2006) AgDscam, a hypervariable immunoglobulin domain-containing receptor of the Anopheles gambiae innate immune system. PLoS Biol 4: e229. doi:10.1371/journal.pbio.0040229.
84. Ewbank JJ (2006) Signaling in the immune response. WormBook. doi: 10.1895/wormbook.1.83.1
85. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biology 5: R80. doi:10.1186/gb-2004-5-10-r80.
86. Smyth GK (2005) Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor. Springer, New York. 397–420.
87. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, et al. (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23: 2700–2707. doi:10.1093/bioinformatics/btm412.
88. Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31: 265–273. doi:10.1016/S1046-023(03)00155-5
89. Smyth GK (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Statistical Applications in Genetics and Molecular Biology 3. doi:10.2202/1544-6115.1027.

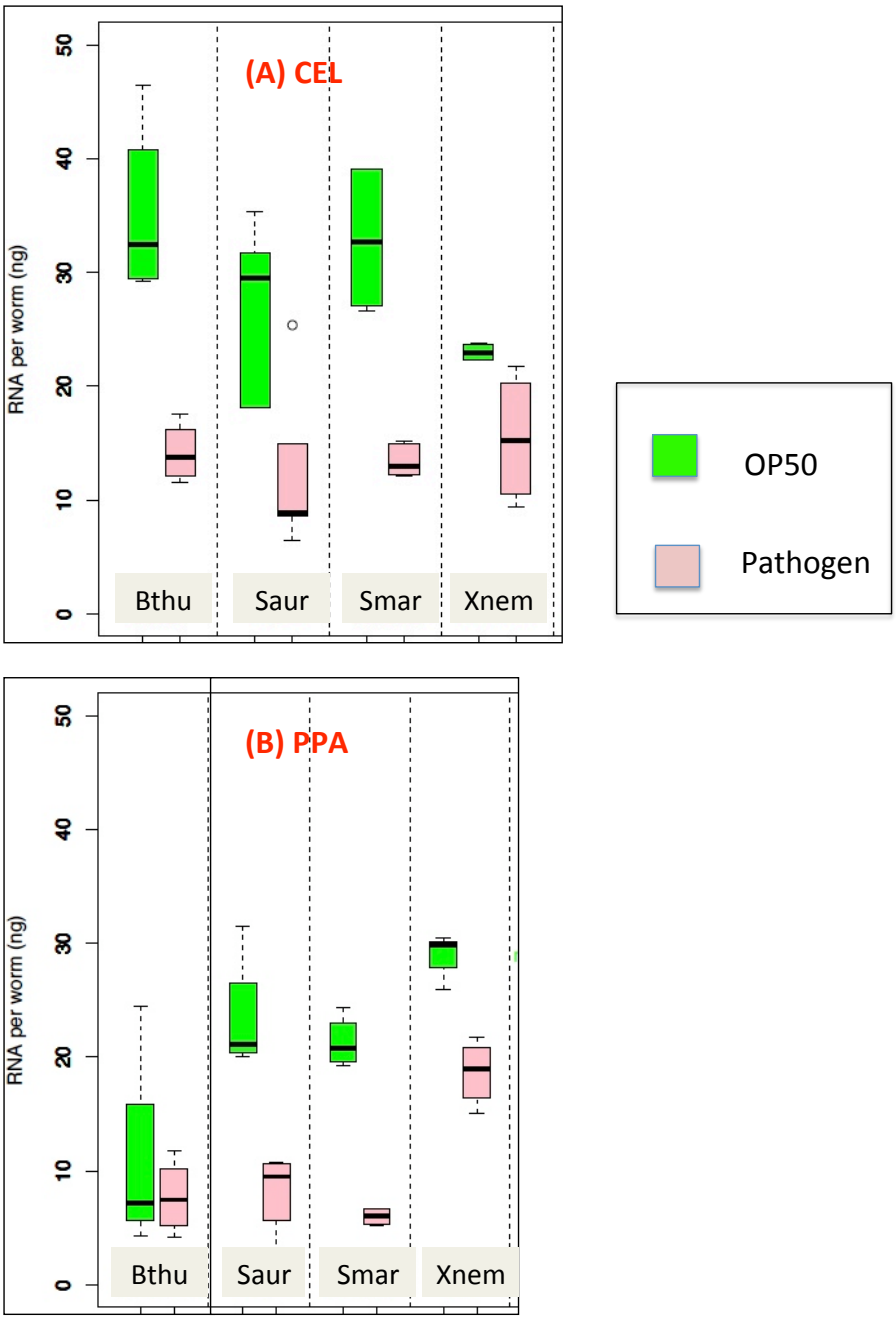
90. Chen H, Boutros P (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 12: 35. doi:10.1186/1471-2105-12-35.
91. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, et al. (2001) Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. Nat Genet 29: 365–371. doi:10.1038/ng1201-365.
92. Eddy SR (2009) A new generation of homology search tools based on probabilistic inference. Genome Inform 23: 205–211.
93. Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22: 1600–1607. doi:10.1093/bioinformatics/btl140.
94. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8: 785–786. doi:10.1038/nmeth.1701.
95. Thomas S, Karnik S, Barai RS, Jayaraman VK, Idicula-Thomas S (2010) CAMP: a useful resource for research on antimicrobial peptides. Nucleic Acids Res 38: D774–780. doi:10.1093/nar/gkp1021.

Supplementary Figure S1 : Long term survival curves for *P. pacificus* on various pathogens.

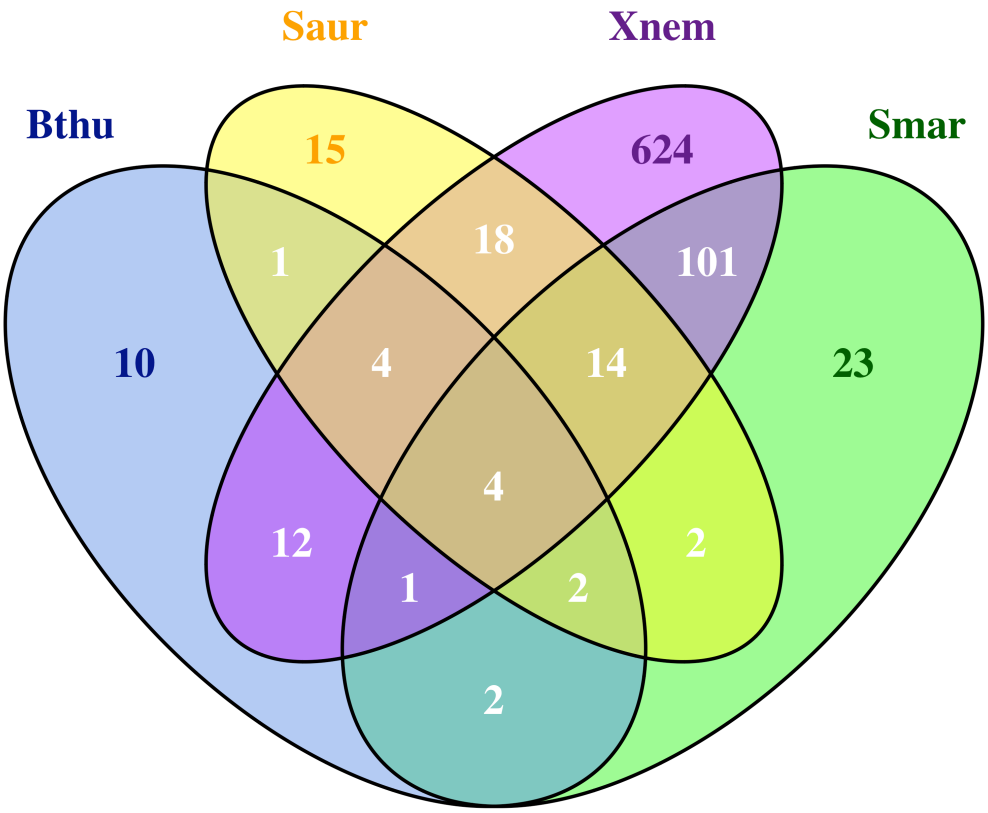
P. pacificus has higher resistance than *C. elegans* to *B. thuringiensis* and *S. aureus*, seen as longer median survival time of about 8 days on these pathogens



Supplementary Figure S2 : Global transcriptional suppression in response to pathogens in (A) *C. elegans* (B) *P. pacificus*

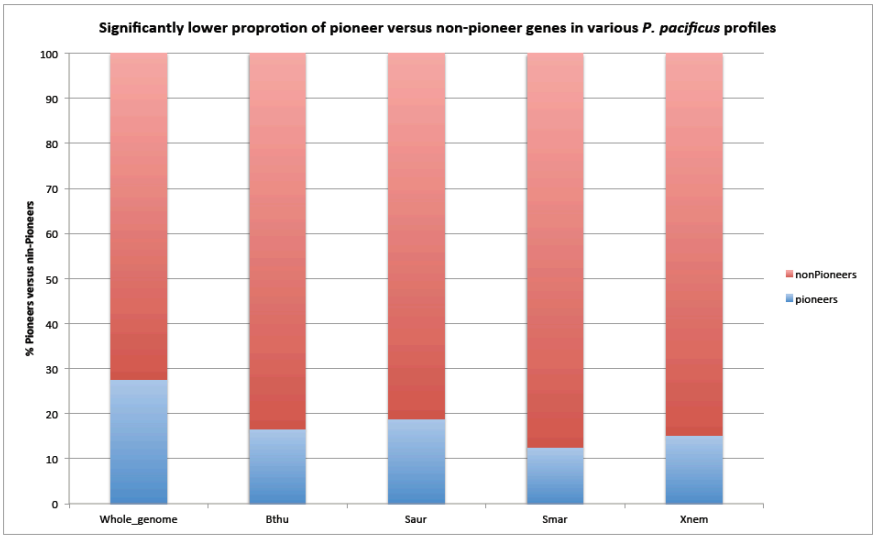


Supplementary Figure S3 : Overlap between pioneer genes regulated in *P. pacificus* in response to the four pathogens



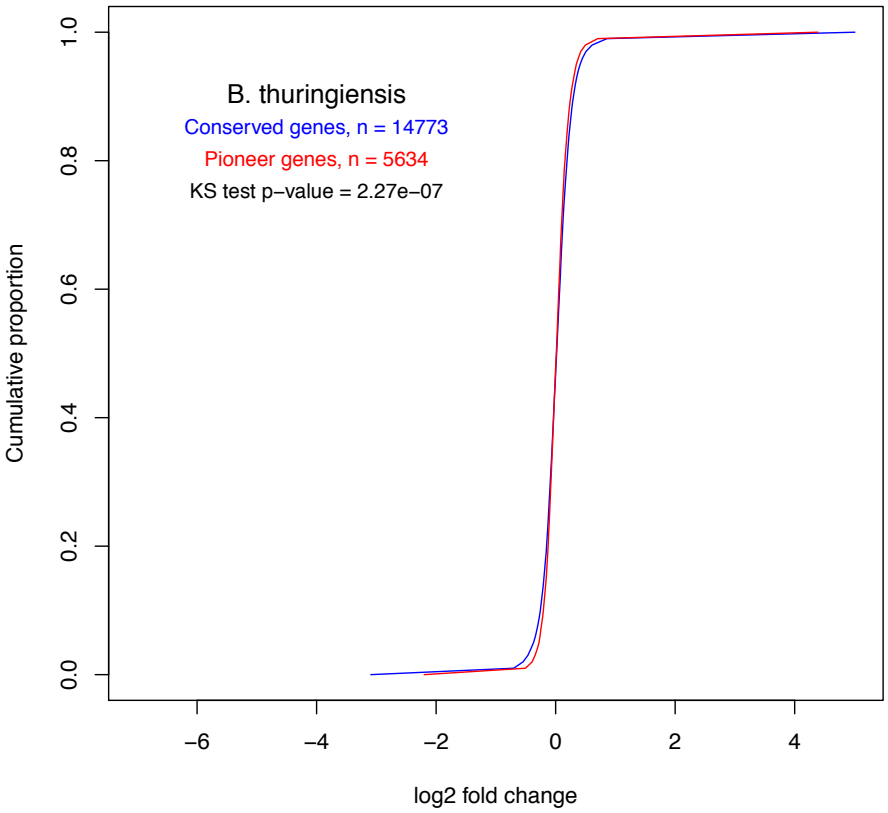
Supplementary Figure S4 : Relative proportions of pioneer genes versus non-pioneer genes in the active transcriptome of *P. pacificus* on each of the four pathogens.

On each of the pathogens, the pioneer genes constitute 12% to 18% of the active transcriptome, significantly less than the expected proportion of about 30% from the distribution in the entire transcriptome. One-sided Fisher's 2x2 exact test p –values are (a) 9.72E-05 on *B. thuringiensis* (b) 1.01E-04 on *S. aureus* (c) 8.75E-39 on *S. marcescens*, and (d) 8.01E-128 on *X. nematophila*



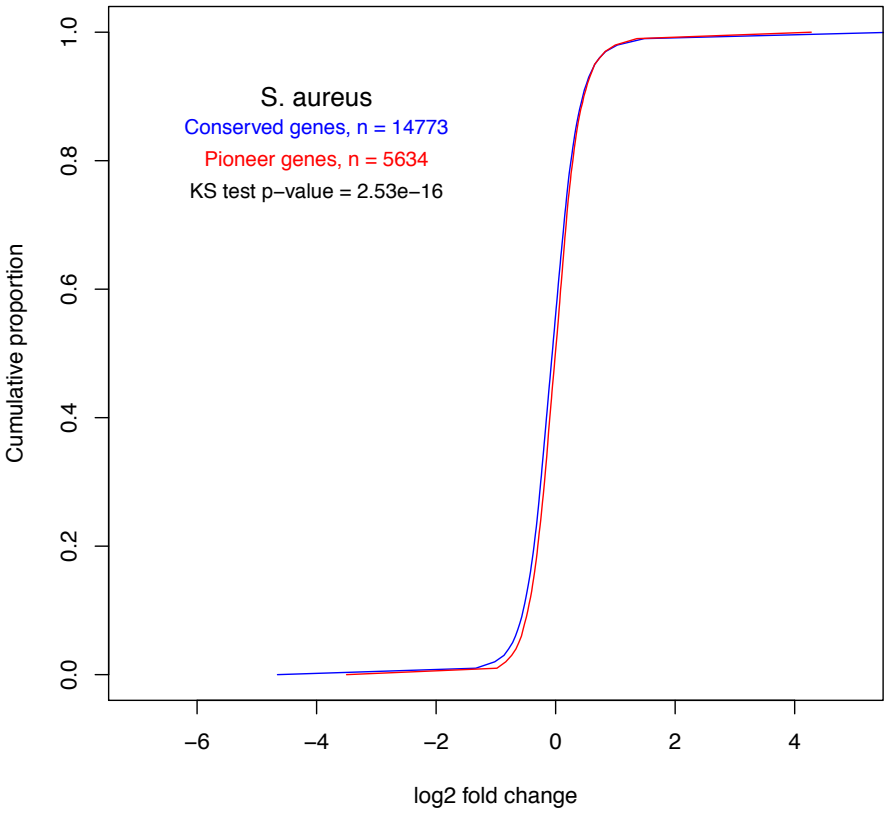
Supplementary Figure S5 .Pioneer genes are expressed at higher levels than non-pioneer genes in each of the pathogen-induced expression profiles on *P. pacificus*.

(A) *B. thuringiensis*



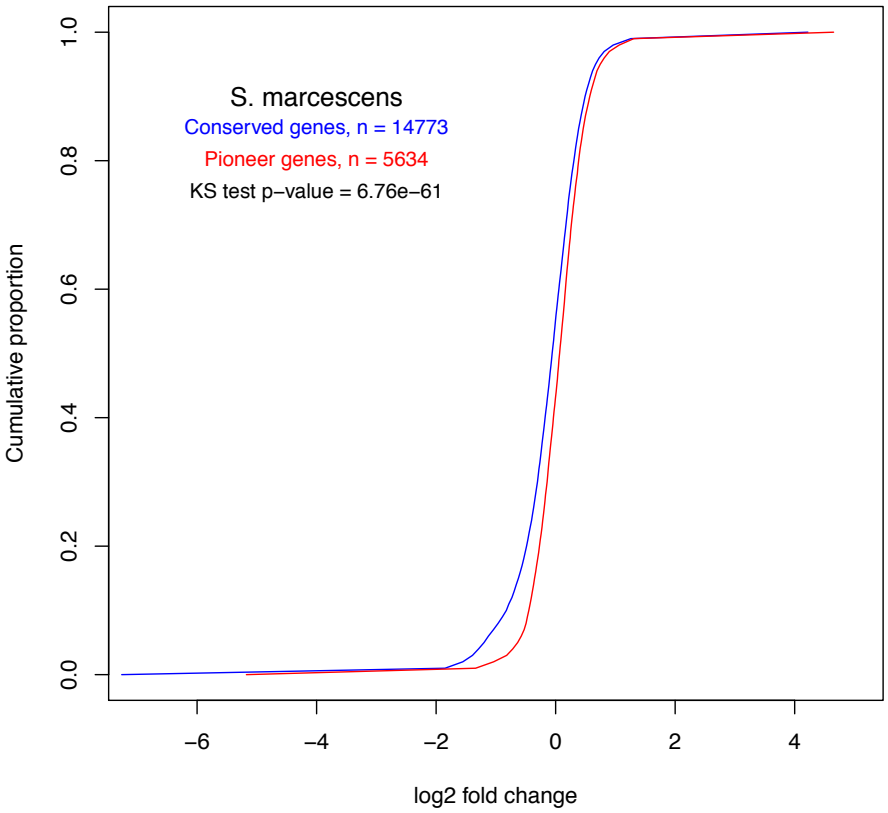
Supplementary Figure S5 .Pioneer genes are expressed at higher levels than non-pioneer genes in each of the pathogen-induced expression profiles on *P. pacificus*.

(B) : *S. aureus*



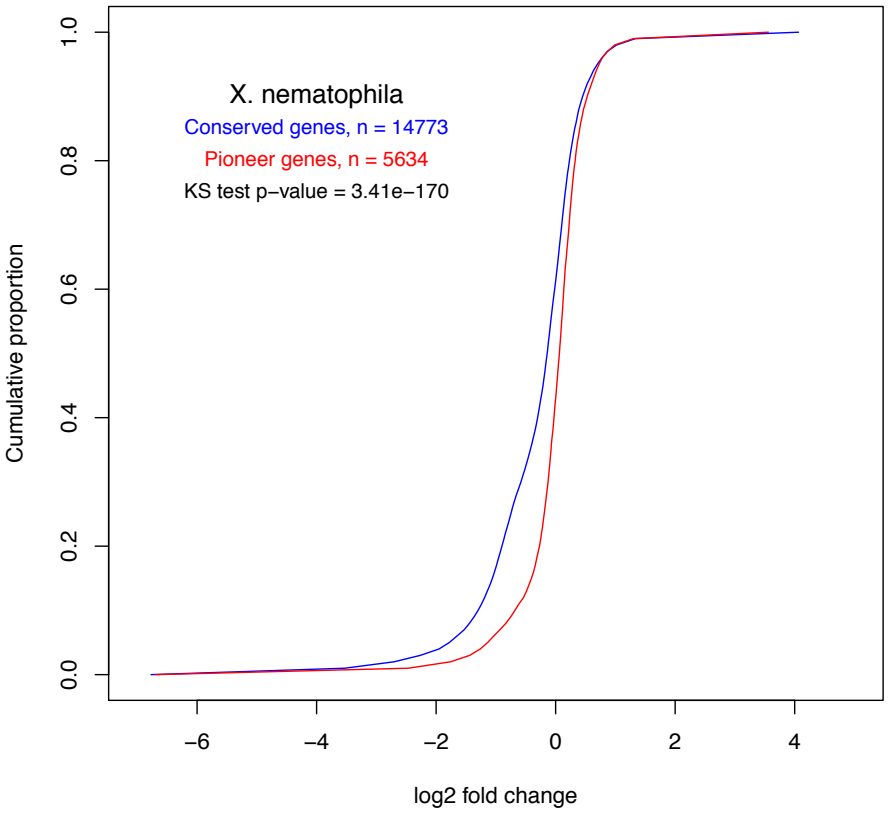
Supplementary Figure S5 .Pioneer genes are expressed at higher levels than non-pioneer genes in each of the pathogen-induced expression profiles on *P. pacificus*.

(C) : *S. marcescens*



Supplementary Figure S5 .Pioneer genes are expressed at higher levels than non-pioneer genes in each of the pathogen-induced expression profiles on *P. pacificus*.

(D) : *X. nematophila*



OPEN ACCESS Freely available online

PLoS PATHOGENS

Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode *Pristionchus pacificus*

Robbie Rae[☯], **Amit Sinha**[☯], **Ralf J. Sommer***

Department of Evolutionary Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany

Abstract

Removal of the reproductive system of many animals including fish, flies, nematodes, mice and humans can increase lifespan through mechanisms largely unknown. The abrogation of the germline in *Caenorhabditis elegans* increases longevity by 60% due to a signal emitted from the somatic gonad. Apart from increased longevity, germline-less *C. elegans* is also resistant to other environmental stressors such as feeding on bacterial pathogens. However, the evolutionary conservation of this pathogen resistance, its genetic basis and an understanding of genes involved in producing this extraordinary survival phenotype are currently unknown. To study these evolutionary aspects we used the necromenic nematode *Pristionchus pacificus*, which is a genetic model system used in comparison to *C. elegans*. By ablation of germline precursor cells and subsequent feeding on the pathogen *Serratia marcescens* we discovered that *P. pacificus* shows remarkable resistance to bacterial pathogens and that this response is evolutionarily conserved across the Genus *Pristionchus*. To gain a mechanistic understanding of the increased resistance to bacterial pathogens and longevity in germline-ablated *P. pacificus* we used whole genome microarrays to profile the transcriptional response comparing germline ablated versus un-ablated animals when fed *S. marcescens*. We show that lipid metabolism, maintenance of the proteasome, insulin signaling and nuclear pore complexes are essential for germline deficient phenotypes with more than 3,300 genes being differentially expressed. In contrast, gene expression of germline-less *P. pacificus* on *E. coli* (longevity) and *S. marcescens* (immunity) is very similar with only 244 genes differentially expressed indicating that longevity is due to abundant gene expression also involved in immunity. By testing existing mutants of *Ppa*-DAF-16/FOXO and the nuclear hormone receptor *Ppa*-DAF-12 we show a conserved function of both genes in resistance to bacterial pathogens and longevity. This is the first study to show that the influence of the reproductive system on extending lifespan and innate immunity is conserved in evolution.

Citation: Rae R, Sinha A, Sommer RJ (2012) Genome-Wide Analysis of Germline Signaling Genes Regulating Longevity and Innate Immunity in the Nematode *Pristionchus pacificus*. PLoS Pathog 8(8): e1002864. doi:10.1371/journal.ppat.1002864

Editor: James B. Lok, University of Pennsylvania, United States of America

Received April 16, 2012; **Accepted** July 2, 2012; **Published** August 9, 2012

Funding: This work was funded by the Max-Planck Society. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ralf.sommer@tuebingen.mpg.de

☯ These authors contributed equally to this work.

Introduction

Removal or alteration of sexual organs can cause a dramatic increase in lifespan in animals including fish, flies, nematodes, mice and humans, but underlying mechanisms remain unknown. For example, gonadectomy slows body wasting and intestinal atrophy in the lamprey *Lampetra fluviatilis* and increases the lifespan in Pacific salmon [1–3]. Castration halts major organ degeneration in male marsupial mice [4,5]. Also, transplantation of ovaries from young mice into older mice and genetic delay of the menopause can increase lifespan and dramatically reduce age related complications, respectively [6,7]. Even in humans castrated men have a 24% increase in median lifespan compared to un-castrated [8]. In the model organisms *Caenorhabditis elegans* and *Drosophila melanogaster* removal of the germline results in an increase in longevity of 40–60% [9,10]. This response depends on several genes including DAF-16/FOXO-like transcription factor, the ankyrin repeat KRI-1, the nuclear hormone receptor DAF-12, the cytochrome P450 DAF-9, the transcription elongation factor TCER-1 [11] and processes such as autophagy, and fat metabolism [12–14]. However, there has never been a systematic analysis of the whole genome transcriptional response to understand what genes are being expressed to decelerate aging and increase lifespan.

It has been shown in *C. elegans* that genes affecting lifespan also affect other phenotypes such as resistance against bacterial pathogens [15]. Specifically, long-lived *C. elegans* germline deficient animals can survive when fed various pathogens [16–19]. However, it is currently unknown how conserved this response of germline ablation-induced longevity and pathogen resistance is in other free-living nematodes and more distantly related animals. The diplogastrid nematode *Pristionchus pacificus* diverged from *C. elegans* 250–400 million years ago [20] and is used as a comparative model to *C. elegans*. This comparative research has revealed evolutionary changes in vulva development [21], gonad morphogenesis [22], and chemosensory behavior [23] compared to *C.*

Author Summary

Removal of the germline in the nematode *Caenorhabditis elegans* can increase lifespan and resistance to bacterial pathogens. Currently there is no information on what genes are regulated to produce this resistance phenotype in other nematodes and whether they are the same as genes involved in lifespan regulation. We used the necromenic nematode, *Pristionchus pacificus*, a species that diverged from *C. elegans* 250–400 MYA, ablated its germline and found increased resistance to the pathogens *Serratia marcescens* and *Xenorhabdus nematophila*. In a novel manner we performed cell ablation of the germline, exposure to bacterial pathogens and used whole genome microarrays of the same animals to find that this resistance is due to expression of genes involved in insulin signaling, nuclear pore complexes, ribosomal translation and lipid production. Furthermore, we see little difference between germline ablated lifespan and immunity leading us to believe that living longer is due to an abundance of genes also being involved with immunity. We could also show that, similar to *C. elegans*, the transcription factor DAF-16/FOXO and nuclear hormone receptor DAF-12, are integral for this response. Our study is the first to understand how the reproductive system regulates both lifespan and innate immunity transcriptionally and offers insights into the signaling cascades involved with resisting pathogen attack.

elegans. The toolkit for *P. pacificus* research includes a fully sequenced genome and a well characterized proteome [20,24], forward and reverse genetics, transgenic techniques [25], full genome microarray technology [26] and hundreds of naturally isolated *P. pacificus* strains isolated from around the world [27]. Interestingly, *C. elegans* and *P. pacificus* also differ in their ecological niches. *C. elegans* can be isolated from compost heaps, snails and rotten fruits [28], whereas *P. pacificus* is usually isolated from a range of scarab beetles [29]. *P. pacificus*, as well as other *Pristionchus* species lives in a necromenic lifestyle, that attach to passing beetles as dauers and feed on the plethora of microorganisms growing on the cadaver when the beetle dies [30].

The comparison of pathogen resistance of germline-ablated *P. pacificus* and *C. elegans* is of special interest given the different response of these two nematodes to bacteria under normal growth conditions. Specifically, *P. pacificus* can feed on *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Bacillus thuringiensis* Cry 5B toxin and *B. thuringiensis* DB27, whereas *C. elegans* dies on these bacterial strains [30–32]. There are six signaling pathways that have been identified as being critical for *C. elegans* survival when fed an array of bacterial and fungal pathogens e.g. ERK MAP kinase, p38 MAP kinase, TGF β, programmed cell death, DAF-2/DAF-16 insulin-like receptor signaling and JNK-like MAP kinase [33]. As some of these pathways also regulate aging in *C. elegans* (e.g. FOXO/DAF-16 insulin-like signaling [15]) it remains to be seen how these animals either use shared or distinct mechanisms to regulate innate immunity and aging.

Here, we investigated whether germline manipulation in *P. pacificus* would increase survival when fed the opportunistic human pathogen *Serratia marcescens*, whether this response is evolutionarily conserved across the Genus *Pristionchus*, and carried out a detailed analysis of the transcriptional mechanistic processes governing this response. Using whole genome microarrays comparing unablated and germline ablated *P. pacificus* we found that resistance to pathogenic bacteria is due to differential expression of genes involved in insulin signaling, pathogen response, lipid metabolism, and core cellular processes like ribosomal translation, proteasome function, nuclear pore complexes. Furthermore, we show that germline ablations of *P. pacificus daf-16* and *daf-12* mutants severely affect the survival when fed bacterial pathogens, thereby underpinning the importance of insulin signaling. Our study is the first to provide an understanding of how the reproductive system regulates both lifespan and innate immunity transcriptionally and offers insights into the signaling cascades involved with resisting pathogen attack.

Results/Discussion

Germline ablated *P. pacificus* are resistant to bacterial pathogens

As in *C. elegans* the gonad of *P. pacificus* consists of four cells (Z1, Z2, Z3 and Z4) with Z2 and Z3 giving rise to the germline and Z1 and Z4 making the somatic gonad, which can be removed using laser microsurgery (Figure 1 a) [34]. Previously it was shown that *P. pacificus* (Figure 1b) is long lived when the germline is ablated [9,35]. We were interested to see if survival on bacterial pathogens would also be enhanced. We ablated the germline (Figure 1c,d) and the somatic gonad separately and fed *S. marcescens*. *S. marcescens* can be isolated from soil, insects and *Pristionchus* nematodes emerging from beetles and is lethal to both *P. pacificus* and *C. elegans* [30].

Germline ablated *P. pacificus* survive significantly longer than un-ablated nematodes on *S. marcescens* (Figure 2a) ($P<0.001$, Supplementary Table S1). In contrast, ablation of Z1 and Z4 does not cause resistance to *S. marcescens* indicating that neither ablation nor sterility *per se* contributes to increased resistance (Figure 2a). Also, the effect is not limited to one pathogen as germline ablated *P. pacificus* also survive on the entomopathogenic nematode associated bacterium *Xenorhabdus nematophilum* that kills wild type *P. pacificus* in two days (Figure 2b). Thus, similarly to germline loss induced longevity, *P. pacificus* germline ablated nematodes are resistant to diverse bacterial pathogens.

It must be noted that in *C. elegans*, when Z1 and Z4 are ablated, the germline precursor cells (Z2 and Z3) will die [34]. Death of Z2 and Z3 is also observed upon Z1 and Z4 ablation in *P. pacificus*, except in a small fraction of animals that develop germline tumors at a rate of about 10% (e.g. 4 out of 35 animals, as reported in [22]).

Germline ablated resistance to *S. marcescens* is conserved across the Genus *Pristionchus*

Currently, over 400 strains of *P. pacificus* isolated worldwide [27] are available in a collection in the Sommer lab in Tuebingen, Germany. Therefore, to investigate if this bacterial resistance response was only present in the *P. pacificus* wild type strain PS312, we chose strains from Montenegro (M2), China (S264) and Japan (RS5160), ablated their germline precursor cells and fed them *S. marcescens* following germline ablation All three *P. pacificus* strains also show significant resistance to *S. marcescens* demonstrating the evolutionary conservation of somatic gonad signaling contributing to innate immunity across *P. pacificus* species (Figure 3 a–c). To investigate this further we expanded to two other *Pristionchus* species from a group of 25 *Pristionchus* species isolated from around the world, available in the Tuebingen collection. We chose *Pristionchus* sp. 3 and *Pristionchus* sp. 16 and repeated our previously described experiment. Again, we observe increased resistance to *S. marcescens* dependent on a signal from the somatic gonad when the germline is ablated. Both species survive significantly longer than un-ablated animals (Figure 3 d, e).

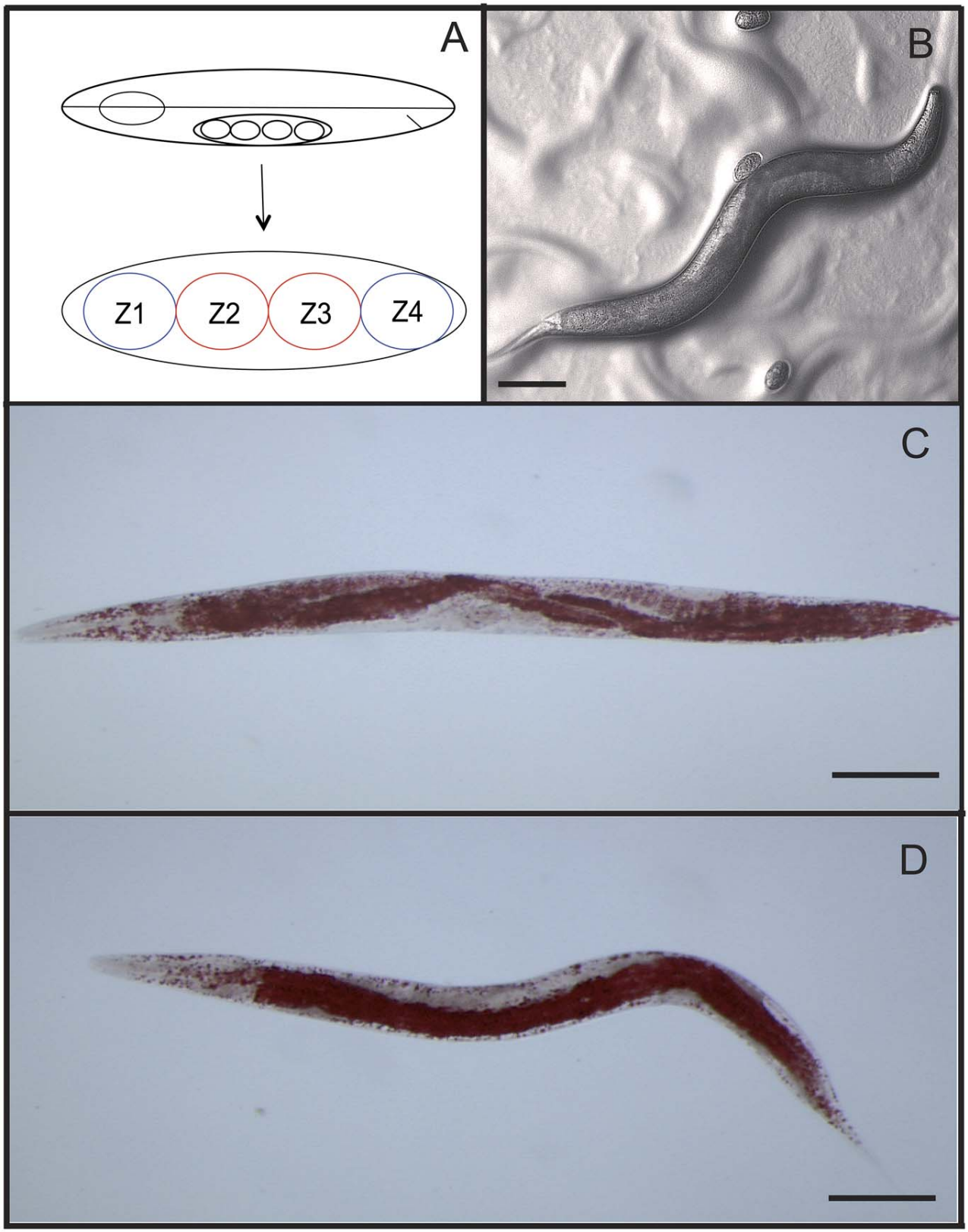


Figure 1. Effect of cell ablation on *P. pacificus*. Z2 and Z3 give rise germline and Z1 and Z4 give rise to somatic gonad (A). *P. pacificus* WT unablated (B). *P. pacificus* WT stained with Oil red O showing tryglyceride staining of intestine, hypodermis, gonad and eggs (C) and *P. pacificus* Z2 and Z3 and ablated similarly stained showing concentration of triglycerides in intestine, like germline ablated *C. elegans* (46) (D). Scale bar represents approx. 100 μm.
doi:10.1371/journal.ppat.1002864.g001

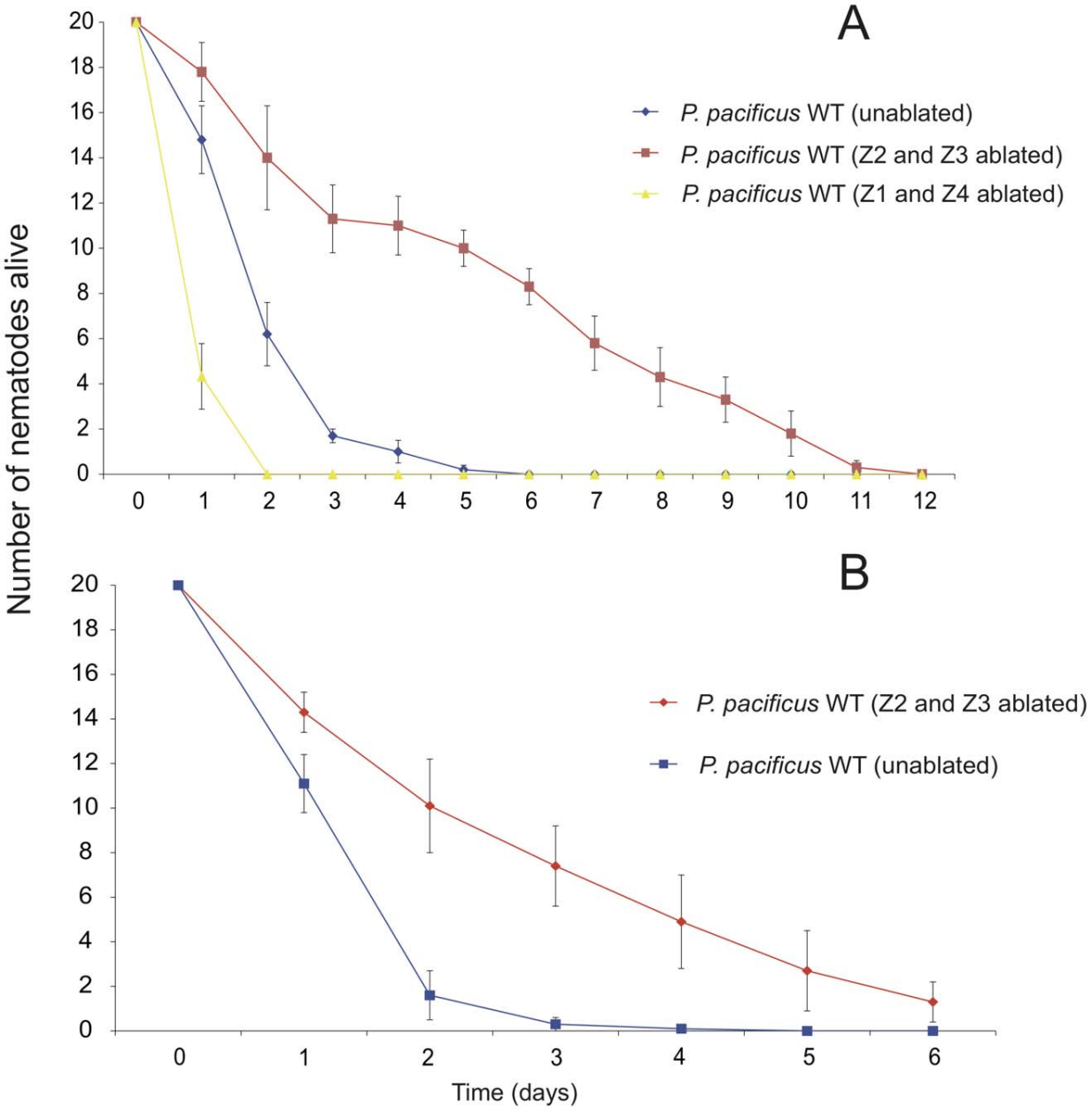


Figure 2. Effect of ablation of survival of *P. pacificus* fed bacterial pathogens. Number of alive *P. pacificus* WT (blue), Z2 and Z3 ablated (red) and Z1 and Z4 (yellow) exposed to *S. marcescens* (A) and *X. nematophila* (B). Batches consisting of twenty nematodes were added separately to three plates and survival was monitored daily beginning on day 0. Error bars represent ± S.E.M.
doi:10.1371/journal.ppat.1002864.g002

Transcriptional responses of germline ablated and un-ablated *P. pacificus* exposed to *E. coli* and *S. marcescens*

Results from germline ablation experiments indicate that the germline regulates some longevity and immunity related signals in a cross talk with the somatic cells of the animal. In order to gain a mechanistic understanding of what genes are regulated when the germline precursor cells (Z2 and Z3) are removed and nematodes are fed pathogens, we assessed the transcriptional response using whole genome microarrays. Many studies in *C. elegans* have looked for genes that mediate the enhanced longevity phenotypes of germline-less animals via RNAi screening [13,14,36]. Here, we have taken an unbiased approach to identify the set of all genes regulated in response to germline ablation, and tried to identify which of them might be functionally relevant. To our knowledge this is the first attempt to couple experimental cell ablation followed by pathogen exposure to microarray analysis.

In the first set of experiments, to identify longevity regulating genes, we ablated the *P. pacificus* germline precursor cells and compared them to un-ablated animals, feeding both of them on *E. coli* OP50. In the second set of experiments, we looked at the pathogen response of germline-ablated animals fed on *S. marcescens* in comparison to germline-ablated animals fed on the lab food source *E. coli* OP50, to check whether long-lived animals require additional transcriptional activity to defend against a pathogen (see

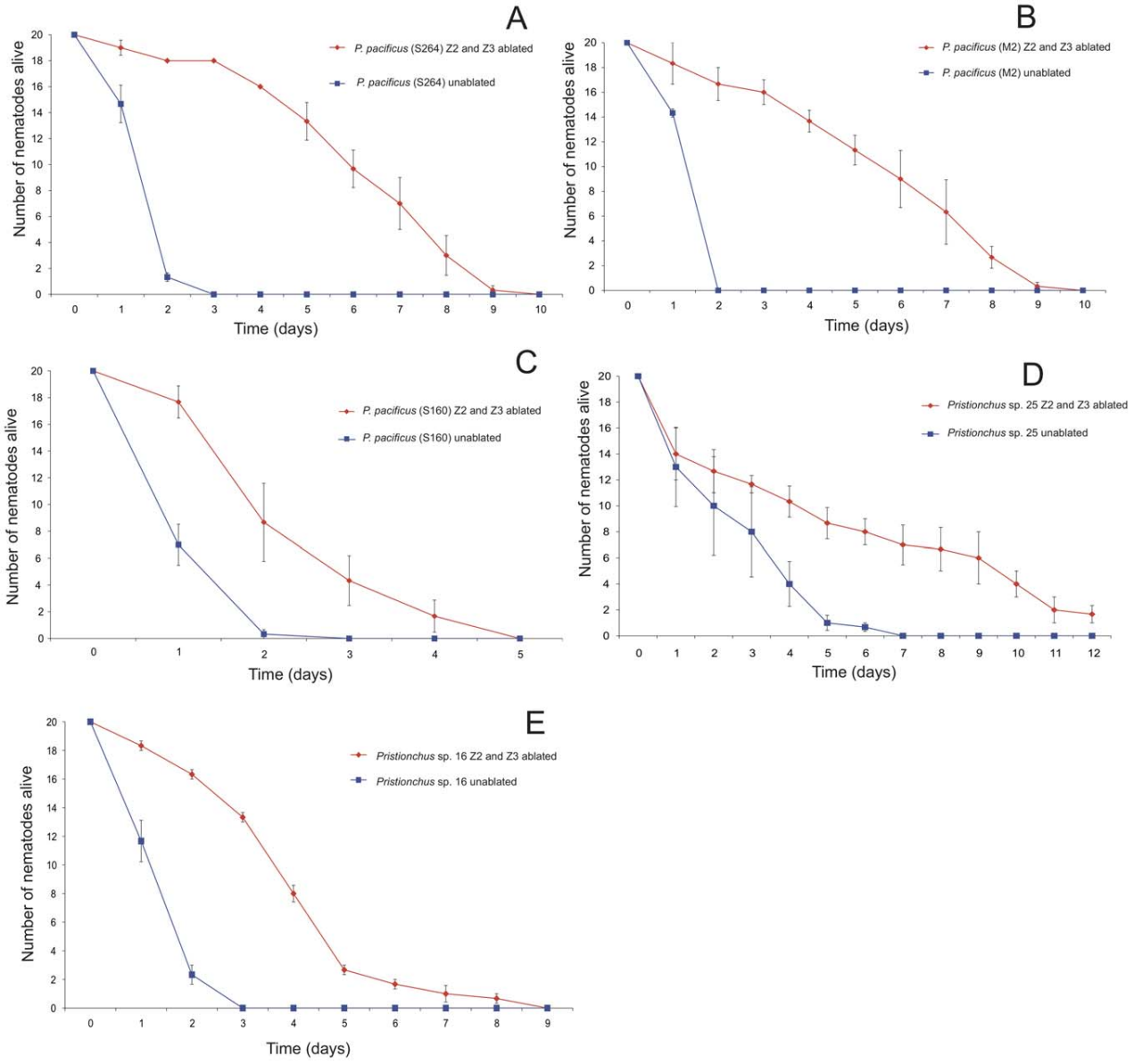


Figure 3. Effect of ablation of survival of *Pristionchus* species and strains exposed to *S. marcescens*. Survival of (A) *P. pacificus* S264 Z2 and Z3 ablated (red) and un-ablated (blue), (B) *P. pacificus* M2 Z2 and Z3 ablated (red) and un-ablated (blue), (C) *P. pacificus* RS5160 Z2 and Z3 ablated (red) and un-ablated (blue), (D) *Pristionchus* sp. 3 Z2 and Z3 ablated (red) and un-ablated (blue), (E) *Pristionchus* sp. 16 Z2 and Z3 ablated (red) and un-ablated (blue). Error bars represent ± S.E.M.
doi:10.1371/journal.ppat.1002864.g003

experimental design in Figure 4). For each condition, we used four independent biological replicates of a pool of about 100 animals each that were either ablated or unablated and exposed either to the pathogen *S. marcescens* or the control food source *E. coli* OP50 for four hours in our microarray experiments. We chose the standard lab food bacterium *E. coli* OP50 as the baseline food source to monitor lifespan so as to enable direct comparisons with various *C. elegans* studies that have also used *E. coli* OP50 as the standard food source [12–16]. Similarly, for survival assays on the pathogen *S. marcescens*, we again used *E. coli* OP50 as the control food source, as is typical in such studies in *C. elegans* [15,16]. In all experiments, young-adult *P. pacificus* animals were exposed to the respective bacterium for 4 hours, a time point when immediate-early pathogen response genes can be robustly detected and the expression profile is relatively unaffected by secondary effects of pathogenesis (AS, RR, II and RJS, unpublished data).

In the comparison of germline-ablated adults versus the un-ablated controls fed on *E. coli* OP50, we find 3,335 genes to be differentially expressed at a FDR corrected p-value threshold of 0.05 and a log2 gold change cut-off at 1.5 (corresponding to a relative fold change of about 2.8 on an absolute count scale, Expression profile E1 in Figure 4, Supplementary Table S2). Interestingly, in the second experiment, comparing germline ablated *P. pacificus* fed *E. coli* OP50 or *S. marcescens*, we found only 244 genes to be differentially expressed (Expression profile E2 in Figure 4, Supplementary Table S3). The relatively small number of genes regulated by pathogen exposure indicates that the contribution of germline-ablation to enhanced longevity is the

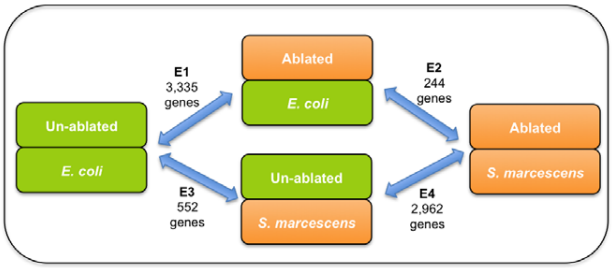


Figure 4. Enhanced longevity due to germline ablation is the major transcriptional component of increased pathogen resistance of these animals. In this schematic of microarrays comparisons, each experiment is represented by two parameters - the ablation status and the bacteria fed, while the double arrows show the samples co-hybridized on the same array. In the first experiment E1, 3,335 genes were differentially expressed when germline ablated animals fed on *E. coli* were compared against un-ablated animals also fed on *E. coli*. In second experiment, E2, only 244 genes were differentially expressed when ablated animals exposed to *S. marcescens* were compared against ablated animals exposed only to *E. coli*, thus indicating that effects of germline ablation are the major transcriptional component towards increased lifespan and pathogen resistance. We also compared these data to transcriptional response of wild-type *P. pacificus* worms exposed to *S. marcescens* (experiment E3, 552 genes, AS, RR, II and RJS, unpublished data). Using these data, we calculated the comparison E4 between ablated versus unablated animals, both fed on *S. marcescens*, and found 2,962 differentially expressed genes.
doi:10.1371/journal.ppat.1002864.g004

major component explaining their increased pathogen resistance. Thus, we characterized the expression profile from the first experiment in more detail, as described below.

a) Differential regulation of translation elongation and initiation and protein homeostasis in long-lived animals. Out of the ~3000 genes differentially expressed upon germline ablation, some would have a role in longevity extension, while others would be related only to the development of the germline. In agreement with this expectation, the most enriched biological processes in a Gene Ontology (GO) analysis (Supplementary Table S4a) are related to "determination of adult lifespan" or to germline and reproductive development (e.g. "hermaphrodite genitalia development", "germ cell development", "gonad development"). The differentially expressed gene set is also highly enriched for genes that encode for "structural constituent of ribosome" (Supplementary Table S4b) and that localize predominantly to the "ribosome", "small ribosomal subunit" and "ribonucleoprotein complex" (Supplementary Table S4c), while the molecular functions "translation initiation factor activity" and "translation elongation factor activity" are also significantly enriched (Supplementary Table S4b). These observations suggest that regulation of translation initiation and elongation at the ribosome is an important component of longevity enhancement program in *P. pacificus*. This interpretation also agrees with the findings in *C. elegans*, whereby the transcription elongation factor TCER-1 was found to be integral for longevity induced by ablation of the germline and to be essential for completion of RNA synthesis during gene expression [36]. Thus, the molecular components regulating germline appear to be conserved between *C. elegans* and *P. pacificus*.

b) Proteasome core subunits and cytoplasmic chaperonins are misregulated in germline-ablated animals. Regulation of protein homeostasis is also a key aspect of lifespan maintenance in *C. elegans* [37]. In our GO analysis of *P. pacificus* data, we find significant enrichment of terms related to proteasome function, such as "proteasome core complex", "protein folding", "protein refolding" (Supplementary Table S4a) and "unfolded protein binding" (Supplementary Table S4b). Further, the most enriched Pfam domains in the genes regulated by germline include the category "Proteasome" and other domains related to proteasome function ("ubiquitin", "PCI" and "Mov34", Supplementary Table S5). The eukaryotic 26S proteasome comprises a core 20S subunit and regulatory 19S regulatory subunits. We observe that upon germline-ablation, five of the seven subunits of the proteasome core alpha subunit (*pas* gene family), and six of the seven subunits of the beta subunit (pbs gene family) are significantly downregulated transcriptionally. Loss of function of any of the components of the proteasome core subunit is known to activate SKN-1 dependent oxidative stress and detoxifying response in a feedback loop [38]. The activation of SKN-1 in turn is responsible for enhanced longevity of insulin signaling mutants in parallel to the *Cel*-DAF-16 pathway [39]. Disruption of proteasomal function has also been shown to increase pathogen resistance [40,41]. Hence it is likely that the observed downregulation of proteasome core subunit components leads to SKN-1 activation resulting in regulation of expression, which contributes to the enhanced longevity of germline-ablated animals.

We further observed that all eight of the *cct* gene family members in *P. pacificus* are downregulated in germline-ablated animals. The loss of function of cytoplasmic chaperonin complex components in *C. elegans* also leads to SKN-1 activation [38] and loss of function of its components such as *Cel-cct-4* and *Cel-cct-6* has been linked to enhanced longevity phenotypes via SKN-1 activation [42]. Hence the observed downregulation of all *cct* family genes in germline abated *P. pacificus* suggests a causal link to enhanced longevity, presumably via SKN-1 activation. Also, it was recently shown that RNAi inactivation of *C. elegans* genes involved with essential processes such as translation, respiration and protein turnover can result in repulsion of nematodes from normally attractive bacteria [40] and that translational inhibition can activate the immune response [43,44].

c) Potential role for nucleolar and nuclear pore complex components in longevity enhancement. We observe downregulation of *Ppa-nol-5*, *Ppa-nol-6*, and *Ppa-nol-10*, which encode three out of six members of nucleolar RNA associated protein (NRAP) family. Inactivation of a *Cel-nol-6* not only affects ribosome biogenesis but also reduces intestinal pathogen accumulation resulting in enhanced pathogen resistance, by inhibiting the p53 homolog *Cel-cep-1* [45]. However, its role in lifespan regulation has not yet been characterized. Interestingly, RNAi inactivation of *Cel-nol-1*, another member of the same gene family, has independently been shown to enhance lifespan [46]. We thus expect the downregulation of *Ppa-nol-5*, *Ppa-nol-6* and *Ppa-nol-10* to contribute to the enhanced longevity of germline-ablated animals.

Intriguingly, we also find significant downregulation of 14 members of the nuclear pore complex protein family (*npp*) in the germline-ablated animals. There are 21 nuclear pore complex proteins in *C. elegans*, and 15 orthologs have been identified in the latest *P. pacificus* gene annotations so far (www.pristionchus.org). We were surprised to find that all the components of an essential nuclear pore complex are transcriptionally downregulated in long-lived animals. One possibility is that all of them are coordinately regulated by common factors, and the mis-regulation of these components (or their upstream regulator) activates certain stress resistance pathways that ultimately result in enhanced longevity.

d) Role of lipid metabolism pathways in enhanced lifespan of germline ablated animals. Given the role of fatty acid desaturation in *C. elegans* longevity [14], we looked into regulation of fatty acid desaturase enzymes in germline ablated *P.*

pacificus. While the *C. elegans* genome contains nine genes that encode a protein with a fatty acid desaturase enzyme domain (Pfam name FA_desaturase, Pfam ID = PF00487), we found 17 proteins with this domain in the predicted proteome of *P. pacificus*. One of these genes was robustly upregulated upon germline ablation in *P. pacificus* and shows the highest sequence similarity to *Cel-fat-7*. Fatty acid elongases such as *Cel-elo-2* also regulate lipid composition and lifespan in *C. elegans* [47]. Interestingly, we observe downregulation of the *P. pacificus* ortholog *Ppa-elo-1* in germline ablated animals, suggesting that the role of lipid metabolic pathways in lifespan regulation may also be conserved in *P. pacificus*. Indeed upon ablation of Z2 and Z3 in *P. pacificus* and subsequent staining with Oil Red O, we observe strong localization of triglycerides in the intestine (Figure 1c,d), indicating differences in fat content compared to unablated animals, similar to that seen in germline deficient *C. elegans* [14,48].

e) Enrichment of DAF-16/FOXO targets, dauer regulated genes, and genes regulated in response to various pathogens. Since the DAF-16/FOXO mediated pathway and the TGF-beta pathway play a role in lifespan regulation and innate immunity in *C. elegans* [9,35], we analyzed the extent and significance of overlap between our differentially expressed genes and the orthologous genes known to be regulated by each of these pathways in *C. elegans*. We indeed observe significant overlap between genes upregulated upon germline ablation and genes regulated by DAF-16 (''Class1'' genes from [49], see Table 1), and between genes downregulated upon germline regulation and genes downregulated by TGF-beta ligand DBL-1 (gene set derived from [50], see Table 1), indicating that regulation of these two pathways plays an important role in lifespan regulation in germline-ablated *P. pacificus*.

Since, dauers represent a stress-resistant long-lived stage in nematodes like *C. elegans* and *P. pacificus*, we checked for overlap of our data with the transcriptome data of *P. pacificus* dauers [26]. The dauer-regulated genes in *P. pacificus* are significantly over-represented in our set of germline-ablation regulated genes, suggesting that a common module of longevity regulating genes is activated in both dauers and germline-ablated animals.

We have recently characterized pathogen response genes in *P. pacificus* in response to fur different pathogens namely *Bacillus thuringiensis*, *Staphylococcus aureus*, *Serratia marcescens* and *Xenorhabdus nematophila* (AS, RR, II and RJS, unpublished data). Interestingly, we found extensive and highly significant overlap between the genes regulated upon germline ablation and the pathogen response genes known in *P. pacificus* (our unpublished data), particularly those regulated by exposure to gram-negative pathogens *S. marcescens* and *X. nematophila* (Table 1). Further, homologues of genes regulated by the p38 MAPK *sek-1* and the JNK MAPK *kgb-1* in *C. elegans* [51] are also over-represented in the germline ablation dataset (Table 1). These observations suggest that the germline ablation leads to activation of a significantly large number of pathogen response genes, which might be the reason for their enhanced pathogen resistance as well.

Other relevant gene expression clusters that were significantly enriched in the set of differentially expressed genes included those involved in germline development [52,53] in *C. elegans* (Table 1). This suggests that our expression cluster enrichment analysis is indeed able to capture biologically relevant expression patterns.

f) Regulation of antimicrobial response genes. We also find significant upregulation of the antimicrobial lysozyme *Ppa-lys-7*, whose corresponding ortholog, *Cel-lys-7* is a known DAF-16 target [49] and is induced upon *S. marcescens* infection [54]. Similarly, genes encoding several other potential antimicrobial proteins such as the lectins *Ppa-clec-1*, *Ppa-clec-149*, *Ppa-clec-160* and *Ppa-clec-41* are also significantly upregulated upon germline ablation. Thus it appears that ablation of the germline also results in constitutive activation of various components of pathogen response machinery, which potentially contributes to longevity as well as their enhanced pathogen resistance.

g) *Ppa-age-1* is downregulated in germline-ablated animals in *P. pacificus*. Apart from this system level analysis of differentially expressed genes, we also looked for regulation of genes that have a known role in longevity in *C. elegans*. Interestingly, we observed significant downregulation of the *P. pacificus* homolog of the *C. elegans* PI3 Kinase subunit *age-1*. *age-1* is a downstream component of insulin signaling whose loss of function leads to increased lifespan in a DAF-16/FOXO dependent manner, as well as increased survival in the presence of pathogenic bacteria [15]. We thus expect the downregulation of *Ppa-age-1* in germline-ablated *P. pacificus* animals to be a major contributor to their extended lifespan and enhanced pathogen resistance.

Immune response of germline ablated animals to the pathogen *S. marcescens*

Since germline ablated animals also have an enhanced resistance to pathogens in addition to enhanced longevity, we wanted to ascertain whether this enhanced resistance is a separable component from enhanced longevity. We thus, exposed the germline-ablated adults of *P. pacificus* either to the pathogen *S. marcescens* or to the control lab food, *E. coli* OP50, for four hours and compared the transcriptional differences. Although germline ablation results in mis-regulation of a large number of genes (~3,330, see previous section), the pathogen response of ablated worms comprises only 244 differentially regulated genes (using the same p-value and fold-change cut-offs as in germline ablation effect experiments).

The differential expression of a relatively smaller number of genes in ablated animals exposed to pathogen suggests that the genes differentially expressed due to ablation alone might be sufficient not only for enhanced longevity but also for enhanced pathogen resistance. It is also possible that although the number of regulated genes is small in absolute number, they may still have large phenotypic effects on pathogen resistance, a possibility that awaits functional validation.

Interestingly, of the 244 genes regulated in ablated animals in response to pathogen, only 54 have a corresponding ortholog in *C. elegans*. We find the *P. pacificus* ortholog of the stress-responsive transcription factor *Cel-pqm-1* to be significantly upregulated, implicating activation of the stress response pathway in response to the pathogen. We also observe upregulation of the lectins *Ppa-clec-6*, *Ppa-clec-41*, *Ppa-clec-160*, but surprisingly we do not see induction of any lysozymes. Since *Ppa-lys-7* was already highly induced in germline-ablated animals (see previous section), it is plausible that no further induction of such genes is needed to counter the pathogens. We indeed observe that although there is no differential expression, the absolute expression levels of *Ppa-lys-7* are relatively quite high in these animals (log2(Average_expression) >13.5, Supplementary Table S3). Many genes involved in lipid metabolism such as *Ppa-fat-7*, *Ppa-elo-1, Ppa-idh-1, Ppa-alh-4, Ppa-acs-14* and *Ppa-ech-6* are also downregulated when ablated worms are exposed to *S. marcescens*.

In a previous set of experiments, we have characterized the expression profile of wild-type, un-ablated *P. pacificus* worms in response to the pathogen *S. marcescens* (AS, RR, II and RJS, unpublished data), where we find 552 genes to be differentially expressed, using the same statistical cut-offs of FDR corrected p-value <0.05 and absolute log2(FoldChange) >1.5). Thus the

Table 1. Microarray expression clusters showing significant overlap with genes up or down regulated upon germline ablation in *P. pacificus*.

DAF-16, TGF-beta and Dauer related clusters		
Expression cluster	**sigScore.UP**	**sigScore.DOWN**
Murphy_etal_cgc5976_Class1	1.35	0
Murphy_etal_cgc5976_Class2	0	1.55
Sinha_etal_ppa_dauers_UP	4.97	0
Sinha_etal_ppa_dauers_DOWN	29.91	302.83
Roberts_etal_2010_DBL-1-DOWN	0	49.64
Roberts_etal_2010_DBL-1-UP	0	12.24
Immunity and Pathogen response related clusters		
Expression cluster	**sigScore.UP**	**sigScore.DOWN**
Kao_etal2011_kgb1_regulated	0	27.03
Kao_etal2011_sek1_regulated	0	26.42
Sinha_et_al_serratia_ppa_up*	43.16	0
Sinha_et_al_serratia_ppa_down*	0	275.3
Sinha_et_al_xeno_ppa_up*	57.95	0
Sinha_et_al_xeno_ppa_down*	0	302.83
Sinha_et_al_staph_ppa_up*	37.5	0
Sinha_et_al_staph_ppa_down*	1.45	0
Sinha_et_al_bac27_ppa_up*	9.16	0
Sinha_et_al_bac27_ppa_down*	4.8	0
Germline and reproductive development related		
Expression cluster	**sigScore.UP**	**sigScore.DOWN**
cgc6390-oogenesis-enriched	0	190.3
cgc6390-spermatogenesis-enriched	47.63	0
WBPaper00037611:GLD-2-associated	0	41.04
WBPaper00037611:RNP-8-associated	0	59.01

Significance scores are −log10 of the p-values obtained in a 2×2 Fisher's exact test, and have been set to zero in case of non-significant enrichment where p-value >0.05.

*Expression clusters derived from our as yet unpublished data on pathogen response of wild-type *P. pacificus* to *Serratia marcescens*, *Xenorhabdus nematophila*, *Staphylococcus aureus*, and *Bacillus thuringiensis* DB27 (AS, RR, II and RJS, manuscript submitted).

doi:10.1371/journal.ppat.1002864.t001

pathogen response of germline-ablated worms (244 genes) is also relatively smaller than that of the wild-type worms. Nonetheless, we find 101 genes to be commonly regulated between the two data sets, the overlap being highly significant (Fisher's exact test p-value = 4.34E-94). The set of overlapping genes include the upregulated genes such as lectin *Ppa-clec-41* and the stress responsive transcription factor *Ppa-pqm-1*, and downregulation of lipid metabolic genes *Ppa-fat-7*, *Ppa-acs-14*, *Ppa-alh-4*, *Ppa-idh-1* and *Ppa-elo-1* (Supplementary Table S6). Thus, the regulation of these genes appears to be crucial for enhanced longevity and pathogen resistance of ablated as well as un-ablated animals. Taken together, our analysis, combining cell ablation, pathogen exposure and microarray analysis in a single experiment, suggests a substantial overlap between lifespan extension and pathogen response in *P. pacificus*.

Effect of different bacteria on transcriptional differences between ablated and unablated worms

While measuring the transcriptional changes in ablated worms exposed to *S. marcescens* versus ablated worms exposed to *E. coli*, it is possible that some of the differences might not be due to pathogenicity factors but due to inherent differences in the two species of the bacteria used (e.g. nutritional differences between *E. coli* and *S. marcescens*). To characterize such differences, we derived the expression profile of ablated worms versus unablated worms when both are exposed to *S. marcescens* for 4 hours (Experiment E4 in Figure 4, see Methods). We found this expression profile (Supplementary Table S7) to be qualitatively quite similar to the longevity expression profile from the comparison of ablated worms versus unablated worms when both are exposed to *E. coli* for 4 hours (Experiment E1 in Figure 4). Specifically, the fold changes of each gene across the two profiles show an almost perfect correlation (Pearson correlation 0.9, Spearman rank correlation = 0.89, Supplementary Figure S2). Given the excellent correlation between the fold changes across the two profiles E1 and E4, we calculated the overlap between all significantly differentially expressed genes across the two conditions, irrespective of the fold-changes, and found only 292 genes expressed exclusively upon exposure to *S. marcescens* but not on *E. coli* (Supplementary Table S8). Sixty-seven of these genes have a *C. elegans* homolog and belong to diverse functional classes, but do not have any obvious or known role in response to pathogens.

This extensive overlap with the longevity profile E1 includes downregulation of *Ppa-age-1*, components of cytoplasmic chaperonin complex (*cct-* family), genes regulating translation elongation and those involved in proteasomal function. Similarly, this gene set is also enriched for orthologs of DAF-16 targets known from *C. elegans*, as well as other genes involved in the pathogen response in *P. pacificus* (Supplementary Table S9).

The FOXO transcription factor DAF-16 and nuclear hormone receptor DAF-12 are responsible for increased bacterial resistant in germline ablated *P. pacificus*

In *C. elegans* it has previously been demonstrated that DAF-16 and DAF-12 are responsible for the germline ablated induced longevity in *C. elegans* [9] and in the increased survival of *glp* (germline proliferation) mutants when fed various pathogens [16–19], although this depends on pathogen growth conditions [16].

Our microarray data analysis suggests a role for *Ppa*-DAF-16/FOX transcription factor and *Ppa*-DAF-12 nuclear hormone receptor in the enhanced longevity and pathogen resistance observed upon germline ablation. First, we found *Ppa-age-1* to be downregulated in germline ablated animals, which is expected to activate DAF-16 dependent transcriptional activity. Second, we found an enrichment of DAF-16 target genes in germline ablated *P. pacificus* (Table 1). Third, we also observe a significant overlap between germline-regulated genes and genes regulated in *P. pacificus* dauers (Table 1), and *Ppa-daf-16* and *Ppa-daf-12* have already been shown to be key regulators of dauer formation [55,56]. Hence, all this evidence combined together implicates *Ppa-daf-16* and *Ppa-daf-12* in germline-loss dependent increase in longevity and pathogen resistance.

As a functional test for the roles of *Ppa-daf-16* and *Ppa-daf-12* in these processes, we ablated the germline precursor cells of two alleles of *Ppa-daf-16* (*tu302* and *tu901*) and *Ppa-daf-12* (*tu390* and *tu389*), as well as a double mutant of both genes and assayed their survival when exposed to the pathogen *S. marcescens*. Both the alleles of germline ablated *Ppa-daf-12* (*tu390* and *tu389*) showed significantly less resistance to *S. marcescens* than germline ablated *P. pacificus* wild type animals ($P<0.001$) (Figure 5). This was even more apparent when the germlines of *Ppa-daf-16* (*tu302* and *tu901*) were ablated, as they show very weak resistance compared to the wild type *P. pacificus* ($P<0.001$). Thus, like longevity in *C. elegans* [9,35], the germline induces a signal through the somatic gonad that depends on the transcription factor *Ppa-daf-16* and the nuclear hormone receptor *Ppa-daf-12* when fed lab food *E. coli* or pathogenic bacteria. It must be noted however, that survival of the ablated *Ppa-daf-16* mutants is significantly greater than unablated *Ppa-daf-16* mutants (Supplementary Figure S1) ($P<0.001$), meaning that the ablation of these alleles still induces an increase in immune responsiveness. Hence, there must be another pathway acting in parallel. When the double mutant (*Ppa-daf-16*; *Ppa-daf-12*) is ablated there is no significant difference between *Ppa-daf-16* or *Ppa-daf-12*, demonstrating that in *P. pacificus* these genes are in the same pathway (data not shown), similar to *C. elegans* [57]. We conclude that somatic gonad signaling, causing increased longevity and pathogen resistance, is largely dependent on DAF-16/FOXO signaling in *P. pacificus*.

Germline signaling genes regulate longevity and innate immunity

Aging affects susceptibility to diseases, vaccine failure, potentially autoimmunity and cancer as well as a decreasing the function of epithelial skin barriers, lung or gastrointestinal tract allowing pathogens to enter mucosal tissues, causing increased risk for aged innate immune systems [58]. We show that the increase in longevity and bacterial resistance upon germline ablation is an evolutionarily conserved response over the *Pristionchus* genus and, similar to the increased longevity and resistance of *C. elegans* germline deficient mutants to bacterial pathogens [9,35], the response is strongly reliant on DAF-16/FOXO and DAF-12/NHR in *P. pacificus* [16–19]. Although the DAF-16 effect depends on the pathogen used in the analysis [16], DAF-16 regulates many stress-response genes including antimicrobial defenses [49,59,60] and is responsible for regulating the formation of stress resistant dauer juveniles [57]. We found an enrichment of genes regulated by DAF-16 associated with our enhanced pathogen phenotype e.g. *Ppa-lys-7*, *Ppa-clec-1*, *Ppa-clec-149*, *Ppa-clec-160* and *Ppa-clec-41*. Our study thus strongly indicates a conserved function of DAF-16/FOXO among nematodes, although it is clear that another pathway works in parallel with DAF-16 as our Z2 and Z3 ablated *Ppa-daf-16* (*tu302* and *tu901*) still showed an increase in resistance compared to un-ablated mutants. In *C. elegans* for example, it has been shown that in *glp* mutants the p38 MAPK pathway acts in parallel to DAF-16 [16]. Unfortunately however, we have no *P. pacificus* MAPK mutant so that it remains unclear whether MAPK signaling also acts in parallel to DAF-16/FOXO in *P. pacificus*.

The cell ablation data presented in this study indicate that the *P. pacificus* germline produces a signal that accelerates ageing [9] and depresses immunity to pathogens. We therefore argue that in wild type animals, the germline signals inhibit DAF-16 activity, but when Z2 and Z3 are ablated then the somatic gonad releases a signal and DAF-16 levels rise and regulate an abundance of genes that affect lifespan and bacterial resistance. The somatic gonad signal is poorly understood and how it is created, released or what the potential targets are remains elusive. It may indeed even be an emergency signal released upon removal of germline cells and secreted from injured cells to neighbors.

This is the first study to try to understand at the level of transcriptional response, how the reproductive system controls both immunity and longevity. Previous studies in *C. elegans* using RNAi screens have identified numerous genes and processes that affect the longevity of *glp-1* mutants including steroid hormone signaling [11], translation elongation [36], autophagy [12], oleic acid synthesis [13] and triglyceride metabolism [14]. By using a novel approach that combines experimental cell ablation, pathogen exposure and microarray analysis, we have expanded on this knowledge and shown an array of processes to be affected in germline-ablated animals. Many of these components have also been shown to be involved in longevity and immunity phenotypes in *C. elegans*, e.g proteasomal and cytoplasmic chaperonin function [16,37,38,40–42], lipid metabolism [13,14,61], and translation initiation and elongation [36,40,43,44]. In addition, we propose a role for disruption of nucleolar proteins (*nol* gene family) and nuclear pore complex proteins (*npp* gene family) in activating promoting longevity and pathogen resistance. This interpretation is also consistent with the recent reports suggesting that disruption of core cellular processes leads to activation of protective responses [40]. However, the mechanisms through which germline ablation affects these essential cellular processes remain to be discovered.

C. elegans mutants that exhibit remarkable lifespan extension e.g. *daf-2*, *age-1* and *glp-1* are also resistant to pathogens, oxidative and thermal stress [11,16–19,62]. Has *P. pacificus* evolved two separate pathways to enhance lifespan or increase immunity? The regulation of innate immunity and longevity by the *Cel*-DAF-2 insulin signaling pathway has caused many to think they are the same [63,64] and studies have shown that longer lived *Caenorhabditis* species are more resistant to abiotic (heavy metals and heat shock) and biotic stresses (*P. aeruginosa* and *S. aureus*) [65].

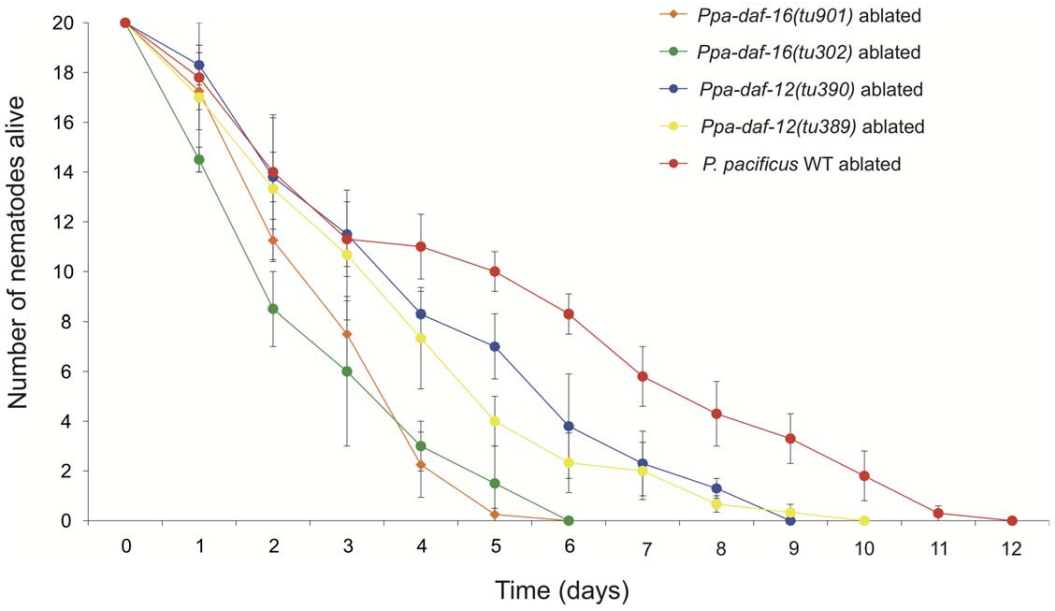


Figure 5. Effect of Z2 and Z3 germline ablation on survival of *P. pacificus* insulin signaling mutants exposed to *S. marcescens*. Survival of Z2 and Z3 ablated *P. pacificus* WT (red), *Ppa-daf-16* (*tu901*) (orange), *Ppa-daf-16* (*tu302*) (green), *Ppa-daf-12* (*tu390*) (blue) and *Ppa-daf-12* (*tu389*) (yellow). Error bars represent ± S.E.M.
doi:10.1371/journal.ppat.1002864.g005

However, there are a few examples in *C. elegans*, where contribution of genes towards either longevity or innate immunity could be separated. For example *sgk-1(ok538)* and *pdk-1(sa680)* mutants are long lived, but are not resistant to *P. aeruginosa* [17]. Similarly, loss of the GATA transcription factor ELT-2 enhances susceptibility to pathogens without shortening lifespan [66] and *sek-1(km4)* and *pmk-1(km25)* mutants are hypersusceptible to pathogens, but have a relatively normal lifespan [67,68]. In our experiments comparing germline-ablated animals fed either *S. marcescens* or *E. coli* OP50, we see only 244 differentially expressed genes comprising of lectins and lipid metabolism genes. Similarly, the expression profile E4 of ablated versus unablated animals on *S. marcescens* was very similar to profile E1 (ablated versus unablated animals on *E. coli*), with only 292 genes specific to E4. With such a small number of genes it seems likely that longevity and pathogen resistance are regulated in a similar manner, although the small differences might be functionally important and could be contributed by a parallel pathway.

Taken together, manipulations of the gonad in an array of diverse organisms such as nematodes [9], flies [10], mice [6,7] and humans [8] have demonstrated increases in lifespan but the mechanisms have remained elusive. By taking a combined approach of using laser microsurgery, pathogen exposure and whole genome microarrays we have demonstrated that upon germline ablation *P. pacificus* can live longer and resist pathogens by regulating numerous downstream effectors that affect an array of processes including translation initiation factors in the ribosome, proteasome maintenance, insulin signaling, nuclear pore complexes and lipid metabolism, which is dependent on the transcription factor DAF-16. It has been well documented that insulin signaling and DNA modifications in FOXO affect longevity in humans [69], but little is known about the role and contribution of genes involved in immunity. We show that processes integral for increasing lifespan and enhancing innate immunity are largely similar. Therefore, upregulation of pathogen defense systems might be an essential factor for living longer.

Materials and Methods

Nematode and bacteria strains

P. pacificus WT RS312, RS5160, S264, M2, *Pristionchus* sp. 3, *Pristionchus* sp. 16, *Ppa-daf-16* (*tu302* and *tu901*) and *Ppa-daf-12* (*tu390* and *tu389*) were maintained on 5 cm NGM agar plates laced with *E. coli* (strain OP50) at 20°C. *S. marcescens* strain C2 was isolated from an *Oryctes* beetle from La Reunion and *X. nematophila* strain XN2 was a gift from Becker Underwood, U.K. and were maintained on LB plates.

Survival assays and analysis

Bacteria (*S. marcescens* and *X. nematophila* XN2) were grown in LB at 30°C overnight in a shaking incubator. The following day 100 µl were spread evenly onto predried 5 cm NGM plates and incubated overnight at 30°C. Three independent biological replicates of 20 worms per plate were exposed to either pathogen or *E. coli* OP50 and were monitored for survival. Worms which failed to respond to a touch of the worm-pick were considered dead. Survival of *P. pacificus* fed *E. coli* OP50 or pathogens was compared using the log rank test.

Cell ablation and RNA collection

P. pacificus J2 stage were picked into 2.8 µl PBS on a agar pad containing 1 mM NaN_3. Ablations would take place within 1 hour of hatching at 20°C. After ablation nematodes were stored at 20°C and successful ablation was verified 48 hours later. Nematodes unablated were grown in parallel and acted as controls. For microarrays 20 *P. pacificus* (either Z2 or Z3 ablated or unablated) were picked onto 5 separate NGM plates either spread with *E. coli* OP50 or *S. marcescens* and incubated at 25°C for 4 hours. Nematodes were then picked into 500 µl of Trizol and stored at −80°C until further analysis. The treatments therefore included (i) Z2 and Z3 ablated *P. pacificus* fed *E. coli* OP50 (ii) Z2 and Z3 ablated *P. pacificus* fed *S. marcescens* (iii) unablated *P. pacificus* fed *E. coli* OP50 (iv) unablated *P. pacificus* fed *S. marcescens*. Development

of ablated and unablated *P. pacificus* fed both *E. coli* OP50 and *S. marcescens* was the same. Each treatment consists of a pool of approximately 100 animals and there were 4 biological replicates of each such pool.

Microarray experiments

A total of 8 microarray hybridizations were carried out for two of the comparisons (E1, E2) depicted in Figure 4. The third comparison (E3, transcriptional response of wild type *P. pacificus* to *S. marcescens*) has been previously characterized in our lab (AS, RR, II and RJS, unpublished data). The fourth comparison E4 (Ablated versus unablated animals on *S. marcescens*) could be computationally derived from this experimental design as the contrast E2+E1−E3. We used our custom designed oligonucleotide microarrays manufactured by Agilent Technologies, which contain ~93,000 unique probes for the ~23,000 *P. pacificus* predicted genes (NCBI GEO accession GPL14372, see [26] for design details of custom microarrays). Equal amounts of total RNA (500 ng to 800 ng) from four biological replicates of each experimental and control samples was used to produce Cy5 or Cy3 dye labeled cRNA using Quick Amp Labelling Kit (Agilent Technologies Inc., USA) as per manufacturer's instructions. Depending upon the amount of total RNA used, appropriate amounts of positive control RNA (Spike Mix-A and Spike Mix-B, from Agilent Technologies) were added to the mix before reverse transcribing the total RNA, as per manufacturer's instructions. The experiments were carried out in a two-color format where Cy5 and Cy3 dye labeled cRNA from experimental and control sample is co-hybridized on the same microarray. The four biological replicates per experiment included two dye-swap experiments to account for differences in dye labeling. Hybridization and washing of the arrays was carried according to manufacturer-supplied protocol. The arrays were scanned on a GenePix 4000B Microarray Scanner, and raw data extracted using GenePix Pro software (version 6).

Microarray data analysis

We used the Bioconductor package limma [70] for analysis of our microarray data. Array quality was checked for parameters such as uniform background and foreground intensities over the entire array. The raw signal was background corrected using the normexp method [71] and the arrays were then lowess normalized individually (''normalizeWithinArrays'' option), with differential weights assigned to probes and to positive control spike-ins, which are expected to show no fold change [72]. This differential weighing of probes is particularly necessary to account for differences in relative proportion of mRNA versus total RNA, and/or differences in the amount of RNA produced per worm under different experimental conditions. Without this differential weighing scheme, the fold change calculations can be erroneous [26,73]. The weight parameters were optimized based on MA-plots such that spike-in controls show their expected fold change values. lmFit function was used to fit a linear model to probe intensities across arrays, differential expression was calculated by empirical Bayes method using the eBayes function [74], and control of FDR was employed as the multiple testing correction. MA-plots were also used as diagnostic to identify and remove outlier arrays before fold-change calculations, such that at least three biological replicates were used for each experiment. Further data analysis was carried out using custom scripts in Perl and the statistical package R. Raw and processed data from all the experiments from this publication have been deposited in a MIAME compliant format at NCBI's Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/), with accession numbers GSE37331 and GSE3733. Expression data from our as yet unpublished experiments (AS, RR, II and RJS, unpublished data, in submission) comparing transcriptomes of wild type *P. pacificus* exposed to *S. marcescens* versus *E. coli* OP50 are available under accession number GSE36521.

Functional analysis of microarray data

We have previously used a pairwise best BLASTP strategy to identify 7,176 pairs of orthologs in *C. elegans* and *P. pacificus* [26]. Probes for 6,126 of these gene pairs exist on microarrays of *P. pacificus*. Gene Ontology annotations were transferred to *P. pacificus* genes using these orthology relations and topGO tool was used for enrichment analysis [75]. Pfam domain annotations are the same as described before [26]. For enrichment analysis, only the domains for which minimum 5 protein coding genes were represented on each microarray were used. Statistical significance of enrichment of Pfam domains in each expression profile determined using a 2×2 Fishers exact test, at a p-value cut-off of 0.05. Expression cluster data from relevant experiments [26,49–53] was compiled from WormBase or from Supplementary Materials of the respective publications. When needed, we inferred expression clusters for *P. pacificus* from *C. elegans* datsests based on the set of orthologs. For the enrichment analysis, only genes with at least one expression-cluster annotation were used as the background set. P-values for expression cluster enrichment in each expression profile were computed with a 2×2 Fisher exact test with a p-value cut-off of 0.05 as the significance threshold.

Supporting Information

Figure S1 Effect of Z2 and Z3 germline ablation on survival of *P. pacificus* insulin signaling mutants exposed to *S. marcescens*. Survival of *Ppa-daf-16* (*tu901*) Z2 and Z3 ablated (blue) and un-ablated (red), and *Ppa-daf-16* (*tu302*) Z2 and Z3 ablated (yellow) and un-ablated (green) exposed to *S. marcescens*. Error bars represent ± S.E.M.
(TIFF)

Figure S2 Comparison of fold-changes in expression profiles E4 (ablated versus unablated animals exposed to *S. marcescens*) and E1 (ablated versus unablated animals exposed to *E. coli*). The two profiles are quite similar, with almost similar fold-changes for the majority of the genes. (Pearson correlation = 0.90, Spearman's rank correlation = 0.89).
(TIFF)

Table S1 Summary statistics of *P. pacificus* ablation experiments monitoring survival when fed *S. marcescens* and *X. nematophila*. Mean survival and standard errors for all conditions tested, and p-values from log Rank test assessing significance of difference between various comparisons. The rows 8 and 9 (marked with an ''*'') correspond to the pathogen *X. nematophila*.
(TIFF)

Table S2 Genes significantly differentially expressed in the comparison of germline-ablated animals fed on *E. coli* OP50 versus un-ablated animals fed *E. coli* OP50 (Expression profile E1). Processed microarray data for experiment E1 for each *P. pacificus* gene with its log2 fold change, FDR corrected p-value and average expression value (log2 scale).
(XLS)

Table S3 Genes significantly differentially expressed in the comparison of germline-ablated animals fed on *S. marcescens* versus germline-ablated animals fed *E. coli*

OP50 (Expression profile E2). Processed microarray data for experiment E2 for each *P. pacificus* gene with its log2 fold change, FDR corrected p-value and average expression value (log2 scale).
(XLS)

Table S4 Enrichment for GO terms from categories (a) Biological Process (b) Molecular Function (c) Cellular Component, in genes differentially regulated upon germline ablation. The total number of genes in *P. pacificus* genome with a given GO term are in the "Annotated" column, the number of genes observed to be significantly differentially expressed are in the column "Significant" and the number of genes expected by random chance are given in the column "Expected". The p-value for enrichment was calculated using the method "elimFisher" in the "topGO" tool in Bioconductor.
(XLS)

Table S5 Pfam domains enriched in the genes regulated upon germline ablation. The total number of genes in *P. pacificus* genome whose products contain a given Pfam domain are in the "Total" column, the number of genes observed to be significantly differentially expressed are in the column "Observed" and the number of genes expected by random chance are given in the column "Expected". "Enrichment" is the ratio of Observed to Expected. P-values for enrichment are from a 2×2 Fisher's Exact test and corrected for False Discover Rate. Proteasome/Ubiquitin system related domains are highlighted in orange. Domains involved in RNA metabolism are highlighted in blue.
(XLS)

Table S6 Genes common between pathogen response of germline-ablated animals (experiment E2) and pathogen response of un-ablated animals (experiment E3). About 100 significantly differentially expressed genes are common between the expression profiles E2 and E3. All genes except one show a similar direction of fold change. The corresponding ortholog in *C. elegans* exists for only 30 of these genes.
(XLS)

Table S7 Genes significantly differentially expressed in the comparison of germline-ablated animals fed on *S. marcescens* versus un-ablated animals fed *S. marcescens* (Expression profile E4). Processed microarray data for experiment E4 for each *P. pacificus* gene with its log2 fold change, FDR corrected p-value and average expression value (log2 scale).
(XLS)

Table S8 List of the 292 genes exclusive to expression profile E4 (ablated versus unablated animals exposed to *S. marcescens*) in a comparison against profile E1 (ablated versus unablated animals exposed to *E. coli*). The log2 fold change, FDR corrected p-value and average expression value (log2 scale) of the 292 genes exclusive to profile E4.
(XLS)

Table S9 Microarray expression clusters showing significant overlap with genes up or down regulated in expression profile E4 (ablated versus unablated *P. pacificus* exposed to *S. marcescens*). The profile E4 is also enriched for genes that are known targets of DAF-16, and TGF-beta pathway in *C. elegans*, and genes regulated in response to various pathogens in *P. pacificus*. Significance scores are −log10 of the p-values obtained in a 2×2 Fisher's exact test, with a zero value indicating non-significant enrichment. The results are qualitatively very similar to that seen for profile E1 (compare with Table 1).
(XLS)

Acknowledgments

We thank Hahn Witte for construction of *Ppa-daf-16;daf-12* double mutant.

Author Contributions

Conceived and designed the experiments: RR AS RJS. Performed the experiments: RR AS. Analyzed the data: RR AS. Contributed reagents/materials/analysis tools: RR AS. Wrote the paper: RR AS RJS.

References

1. Larsen LO (1974) Effects of testosterone and oestradiol on gonadectomized and intact male and female river lampreys (Lampetra fluviatilis (L.) Gray). Gen Comp Endocrinol 24: 305–313.
2. Kime DE, Larsen LO (1987) Effect of gonadectomy and hypophysectomy on plasma steroid levels in male and female lampreys (Lampetra fluviatilis, L.). Gen Comp Endocrinol 68: 189–196.
3. Robertson OH (1961) Relation of gonadal maturation to length of life in Pacific salmon. Fed Proc 20(Suppl 8): 29–30.
4. Lee AK, Cockburn A (1985) Evolutionary Ecology of Marsupials. Cambridge University Press. 283 p.
5. Fahy GM (2010) Precedents for the Biological Control of Aging: Experimental Postponement, Prevention, and Reversal of Aging Processes. In: Fahy GM, West MD, Coles LS, Harris SB, editors. The Future of Aging. Netherlands: Springer. pp. 127–223.
6. Cargill SL, Carey JR, Müller H-G, Anderson G (2003) Age of ovary determines remaining life expectancy in old ovariectomized mice. Aging Cell 2: 185–190.
7. Perez GI, Jurisicova A, Wise L, Lipina T, Kanisek M, et al. (2007) Absence of the proapoptotic Bax protein extends fertility and alleviates age-related health complications in female mice. Proc Natl Acad Sci U S A 104: 5229–5234. doi:10.1073/pnas.0608557104.
8. Hamilton JB, Mestler GE (1969) Mortality and survival: comparison of eunuchs with intact men and women in a mentally retarded population. J Gerontol 24: 395–411.
9. Hsin H, Kenyon C (1999) Signals from the reproductive system regulate the lifespan of C. elegans. Nature 399: 362–366. doi:10.1038/20694.
10. Flatt T, Min K-J, D'Alterio C, Villa-Cuesta E, Cumbers J, et al. (2008) Drosophila germ-line modulation of insulin signaling and lifespan. Proc Natl Acad Sci U S A 105: 6368–6373. doi:10.1073/pnas.0709128105.
11. Kenyon C (2010) A pathway that links reproductive status to lifespan in Caenorhabditis elegans. Ann N Y Acad Sci 1204: 156–162. doi:10.1111/j.1749-6632.2010.05640.x.
12. Lapierre LR, Gelino S, Meléndez A, Hansen M (2011) Autophagy and Lipid Metabolism Coordinately Modulate Life Span in Germline-less C. elegans. Curr Biol 21: 1507–1514. doi:10.1016/j.cub.2011.07.042.
13. Goudeau J, Bellemin S, Toselli-Mollereau E, Shamalnasab M, Chen Y, et al. (2011) Fatty Acid Desaturation Links Germ Cell Loss to Longevity Through NHR-80/HNF4 in C. elegans. PLoS Biol 9: e1000599. doi:10.1371/journal.pbio.1000599.
14. Wang MC, O'Rourke EJ, Ruvkun G (2008) Fat metabolism links germline stem cells and longevity in C. elegans. Science 322: 957–960. doi:10.1126/science.1162011.
15. Garsin DA, Villanueva JM, Begun J, Kim DH, Sifri CD, et al. (2003) Long-lived C. elegans daf-2 mutants are resistant to bacterial pathogens. Science 300: 1921. doi:10.1126/science.1080147.
16. Alper S, McElwee MK, Apfeld J, Lackford B, Freedman JH, et al. (2010) The Caenorhabditis elegans Germ Line Regulates Distinct Signaling Pathways to Control Lifespan and Innate Immunity. J Biol Chem 285: 1822–1828. doi:10.1074/jbc.M109.057323.
17. Evans EA, Chen WC, Tan M-W (2008) The DAF-2 insulin-like signaling pathway independently regulates aging and immunity in C. elegans. Aging Cell 7: 879–893. doi:10.1111/j.1474-9726.2008.00435.x.
18. Miyata S, Begun J, Troemel ER, Ausubel FM (2008) DAF-16-dependent suppression of immunity during reproduction in Caenorhabditis elegans. Genetics 178: 903–918. doi:10.1534/genetics.107.083923.
19. TeKippe M, Aballay A (2010) C. elegans Germline-Deficient Mutants Respond to Pathogen Infection Using Shared and Distinct Mechanisms. PLoS ONE 5: e11777. doi:10.1371/journal.pone.0011777.
20. Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, et al. (2008) The Pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism. Nat Genet 40: 1193–1198. doi:10.1038/ng.227.
21. Wang X, Sommer RJ (2011) Antagonism of LIN-17/Frizzled and LIN-18/Ryk in Nematode Vulva Induction Reveals Evolutionary Alterations in Core

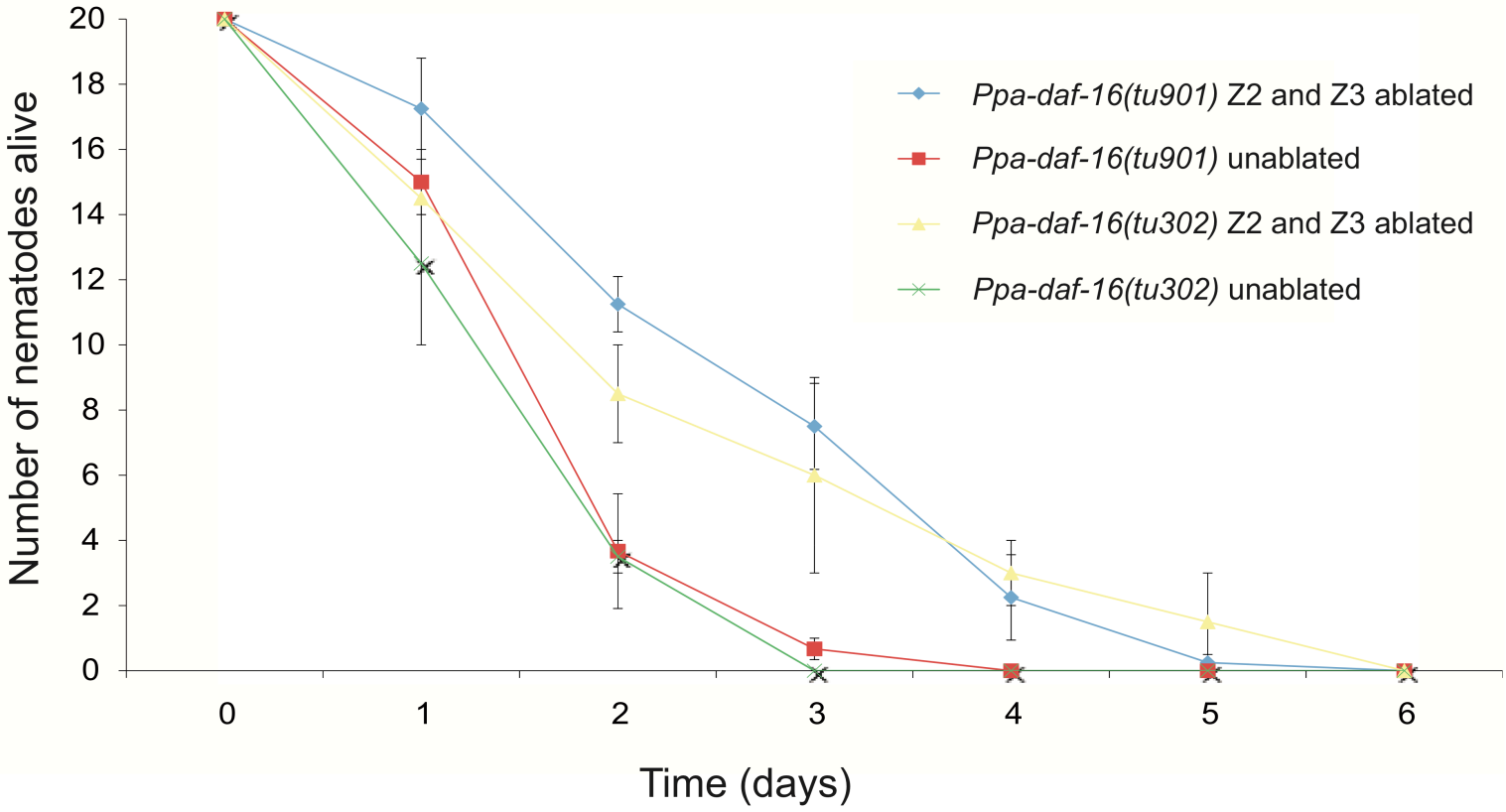
Developmental Pathways. PLoS Biol 9: e1001110. doi:10.1371/journal.pbio.1001110.
22. Rudel D, Riebesell M, Sommer RJ (2005) Gonadogenesis in Pristionchus pacificus and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad. Dev Biol 277: 200–221. doi:S0012-1606(04)00658-X.
23. Hong RL, Witte H, Sommer RJ (2008) Natural variation in Pristionchus pacificus insect pheromone attraction involves the protein kinase EGL-4. Proc Natl Acad Sci U S A 105: 7779–7784. doi:0708406105.
24. Borchert N, Dieterich C, Krug K, Schütz W, Jung S, et al. (2010) Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models. Genome Res 20: 837–846. doi:10.1101/gr.103119.109.
25. Schlager B, Wang X, Braach G, Sommer RJ (2009) Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode Pristionchus pacificus. Genesis 47: 300–304. doi:10.1002/dvg.20499.
26. Sinha A, Sommer RJ, Dieterich C (2012) Divergent gene expression in the conserved dauer stage of the nematodes Pristionchus pacificus and Caenorhabditis elegans. BMC Genomics 13: 254. doi:10.1186/1471-2164-13-254.
27. Morgan K, McGaughran A, Villate L, Herrmann M, Witte H, et al. (2012) Multi locus analysis of Pristionchus pacificus on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. Mol Ecol 21: 250–266. doi:10.1111/j.1365-294X.2011.05382.x.
28. Félix M-A, Braendle C (2010) The natural history of Caenorhabditis elegans. Curr Biol 20: R965–969. doi:10.1016/j.cub.2010.09.050.
29. Herrmann M, Mayer WE, Sommer RJ (2006) Nematodes of the genus Pristionchus are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. Zoology (Jena) 109: 96–108. doi:S0944-2006(06)00006-7.
30. Rae R, Riebesell M, Dinkelacker I, Wang Q, Herrmann M, et al. (2008) Isolation of naturally associated bacteria of necromenic Pristionchus nematodes and fitness consequences. J Exp Biol 211: 1927–1936. doi:211/12/1927.
31. Rae R, Iatsenko I, Witte H, Sommer RJ (2010) A subset of naturally isolated Bacillus strains show extreme virulence to the free-living nematodes Caenorhabditis elegans and Pristionchus pacificus. Environ Microbiol 12: 3007–21. doi:10.1111/j.1462-2920.2010.02278.x.
32. Wei J-Z, Hale K, Carta L, Platzer E, Wong C, et al. (2003) Bacillus thuringiensis crystal proteins that target nematodes. Proc Natl Acad Sci U S A 100: 2760–2765. doi:10.1073/pnas.0538072100.
33. Partridge FA, Gravato-Nobre MJ, Hodgkin J (2010) Signal transduction pathways that function in both development and innate immunity. Dev Dyn 239: 1330–1336. doi:10.1002/dvdy.22232.
34. Kimble JE, White JG (1981) On the control of germ cell development in Caenorhabditis elegans. Dev Biol 81: 208–219. doi:10.1016/0012-1606(81)90284-0.
35. Patel MN, Knight CG, Karageorgi C, Leroi AM (2002) Evolution of germ-line signals that regulate growth and aging in nematodes. Proc Natl Acad Sci U S A 99: 769–774. doi:10.1073/pnas.012511099.
36. Ghazi A, Henis-Korenblit S, Kenyon C (2009) A transcription elongation factor that links signals from the reproductive system to lifespan extension in Caenorhabditis elegans. PLoS Genet 5: e1000639. doi:10.1371/journal.pgen.1000639.
37. Ghazi A, Henis-Korenblit S, Kenyon C (2007) Regulation of Caenorhabditis elegans lifespan by a proteasomal E3 ligase complex. Proc Natl Acad Sci U S A 104: 5947–5952. doi:10.1073/pnas.0700638104.
38. Kahn NW, Rea SL, Moyle S, Kell A, Johnson TE (2008) Proteasomal dysfunction activates the transcription factor SKN-1 and produces a selective oxidative-stress response in Caenorhabditis elegans. Biochem J 409: 205–213. doi:10.1042/BJ20070521.
39. Tullet JMA, Hertweck M, An JH, Baker J, Hwang JY, et al. (2008) Direct inhibition of the longevity-promoting factor SKN-1 by insulin-like signaling in C. elegans. Cell 132: 1025–1038. doi:10.1016/j.cell.2008.01.030.
40. Melo JA, Ruvkun G (2012) Inactivation of Conserved C. elegans Genes Engages Pathogen- and Xenobiotic-Associated Defenses. Cell 149: 452–466. doi:10.1016/j.cell.2012.02.050.
41. Alper S, Laws R, Lackford B, Boyd WA, Dunlap P, et al. (2008) Identification of innate immunity genes and pathways using a comparative genomics approach. Proc Natl Acad Sci U S A 105: 7016–7021. doi:10.1073/pnas.0802405105.
42. Curran SP, Wu X, Riedel CG, Ruvkun G (2009) A soma-to-germline transformation in long-lived Caenorhabditis elegans mutants. Nature 459: 1079–1084. doi:10.1038/nature08106.
43. McEwan DL, Kirienko NV, Ausubel FM (2012) Host Translational Inhibition by Pseudomonas aeruginosa Exotoxin A Triggers an Immune Response in Caenorhabditis elegans. Cell Host Microbe 11: 364–374. doi:10.1016/j.chom.2012.02.007.
44. Dunbar TL, Yan Z, Balla KM, Smelkinson MG, Troemel ER (2012) C. elegans Detects Pathogen-Induced Translational Inhibition to Activate Immune Signaling. Cell Host Microbe 11: 375–386. doi:10.1016/j.chom.2012.02.008.
45. Fuhrman LE, Goel AK, Smith J, Shianna KV, Aballay A (2009) Nucleolar Proteins Suppress Caenorhabditis elegans Innate Immunity by Inhibiting p53/CEP-1. PLoS Genet 5: e1000657. doi:10.1371/journal.pgen.1000657.
46. Curran SP, Ruvkun G (2007) Lifespan regulation by evolutionarily conserved genes essential for viability. PLoS Genet 3: e56. doi:10.1371/journal.pgen.0030056.
47. Horikawa M, Nomura T, Hashimoto T, Sakamoto K (2008) Elongation and desaturation of fatty acids are critical in growth, lipid metabolism and ontogeny of Caenorhabditis elegans. J Biochem 144: 149–158. doi:10.1093/jb/mvn055.
48. O'Rourke EJ, Soukas AA, Carr CE, Ruvkun G (2009) C. elegans major fats are stored in vesicles distinct from lysosome-related organelles. Cell Metab 10: 430–435. doi:10.1016/j.cmet.2009.10.002.
49. Murphy CT, McCarroll SA, Bargmann CI, Fraser A, Kamath RS, et al. (2003) Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature 424: 277–283. doi:10.1038/nature01789.
50. Roberts AF, Gumienny TL, Gleason RJ, Wang H, Padgett RW (2010) Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in Caenorhabditis elegans. BMC Dev Biol 10: 61. doi:10.1186/1471-213X-10-61.
51. Kao C-Y, Los FCO, Huffman DL, Wachi S, Kloft N, et al. (2011) Global functional analyses of cellular responses to pore-forming toxins. PLoS Pathog 7: e1001314. doi:10.1371/journal.ppat.1001314.
52. Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development 131: 311–323. doi:10.1242/dev.00914.
53. Kim KW, Wilson TL, Kimble J (2010) GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. Proc Natl Acad Sci U S A 107: 17445–17450. doi:10.1073/pnas.1012611107.
54. Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, et al. (2002) Inducible Antibacterial Defense System in C. elegans. Current Biol 12: 1209–1214. doi:10.1016/S0960-9822(02)00928-4.
55. Ogawa A, Streit A, Antebi A, Sommer RJ (2009) A Conserved Endocrine Mechanism Controls the Formation of Dauer and Infective Larvae in Nematodes. Current Biol 19: 67–71. doi:10.1016/j.cub.2008.11.063.
56. Ogawa A, Bento G, Bartelmes G, Dieterich C, Sommer RJ (2011) Pristionchus pacificus daf-16 is essential for dauer formation but dispensable for mouth form dimorphism. Development 138: 1281–1284. doi:10.1242/dev.058909.
57. Riddle DL, Swanson MM, Albert PS (1981) Interacting genes in nematode dauer larva formation. Nature 290: 668–671. doi:10.1038/290668a0.
58. Weiskopf D, Weinberger B, Grubeck-Loebenstein B (2009) The aging of the immune system. Transpl Int 22: 1041–1050. doi:10.1111/j.1432-2277.2009.00927.x.
59. McElwee J, Bubb K, Thomas JH (2003) Transcriptional outputs of the Caenorhabditis elegans forkhead protein DAF-16. Aging Cell 2: 111–121. doi:10.1046/j.1474-9728.2003.00043.x.
60. Yu H, Larsen PL (2001) DAF-16-dependent and independent expression targets of DAF-2 insulin receptor-like pathway in Caenorhabditis elegans include FKBPs. J Mol Biol 314: 1017–1028. doi:10.1006/jmbi.2000.5210.
61. McCormick M, Chen K, Ramaswamy P, Kenyon C (2012) New genes that extend Caenorhabditis elegans' lifespan in response to reproductive signals. Aging Cell 11: 192–202. doi:10.1111/j.1474-9726.2011.00768.x.
62. Barsyte D, Lovejoy DA, Lithgow GJ (2001) Longevity and heavy metal resistance in daf-2 and age-1 long-lived mutants of Caenorhabditis elegans. FASEB J 15: 627–634. doi:10.1096/fj.99-0966com.
63. Bolm M, Chhatwal GS, Jansen WTM (2004) Bacterial Resistance of daf-2 Mutants. Science 303: 1976–1976. doi:10.1126/science.303.5666.1976a.
64. Lithgow GJ (2003) Does Anti-Aging Equal Anti-Microbial? Sci Aging Knowl Environ 2003: pe16. doi:10.1126/sageke.2003.25.pe16.
65. Amrit FRG, Boehnisch CML, May RC (2010) Phenotypic covariance of longevity, immunity and stress resistance in the caenorhabditis nematodes. PLoS ONE 5: e9978. doi:10.1371/journal.pone.0009978.
66. Evans EA, Kawli T, Tan M-W (2008) Pseudomonas aeruginosa Suppresses Host Immunity by Activating the DAF-2 Insulin-Like Signaling Pathway in Caenorhabditis elegans. PLoS Pathog 4: e1000175. doi:10.1371/journal.ppat.1000175.
67. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, et al. (2002) A conserved p38 MAP kinase pathway in Caenorhabditis elegans innate immunity. Science 297: 623–626. doi:10.1126/science.1073759.
68. Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, et al. (2006) p38 MAPK Regulates Expression of Immune Response Genes and Contributes to Longevity in C. elegans. PLoS Genet 2: e183. doi:10.1371/journal.pgen.0020183.
69. Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, et al. (2009) Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. Aging Cell 8: 460–472. doi:10.1111/j.1474-9726.2009.00493.x.
70. Smyth GK (2005) Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer. pp. 397–420.
71. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, et al. (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23: 2700–2707. doi:10.1093/bioinformatics/btm412.
72. Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31: 265–273. doi:10.1016/S1046-2023(03)00155-5.
73. van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, et al. (2003) Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep 4: 387–393. doi:10.1038/sj.embor.embor798.

74. Smyth GK (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Stat Appl Genet Mol Biol 3: Article 3. doi:10.2202/1544-6115.1027.

75. Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22: 1600–1607. doi:10.1093/bioinformatics/btl140.

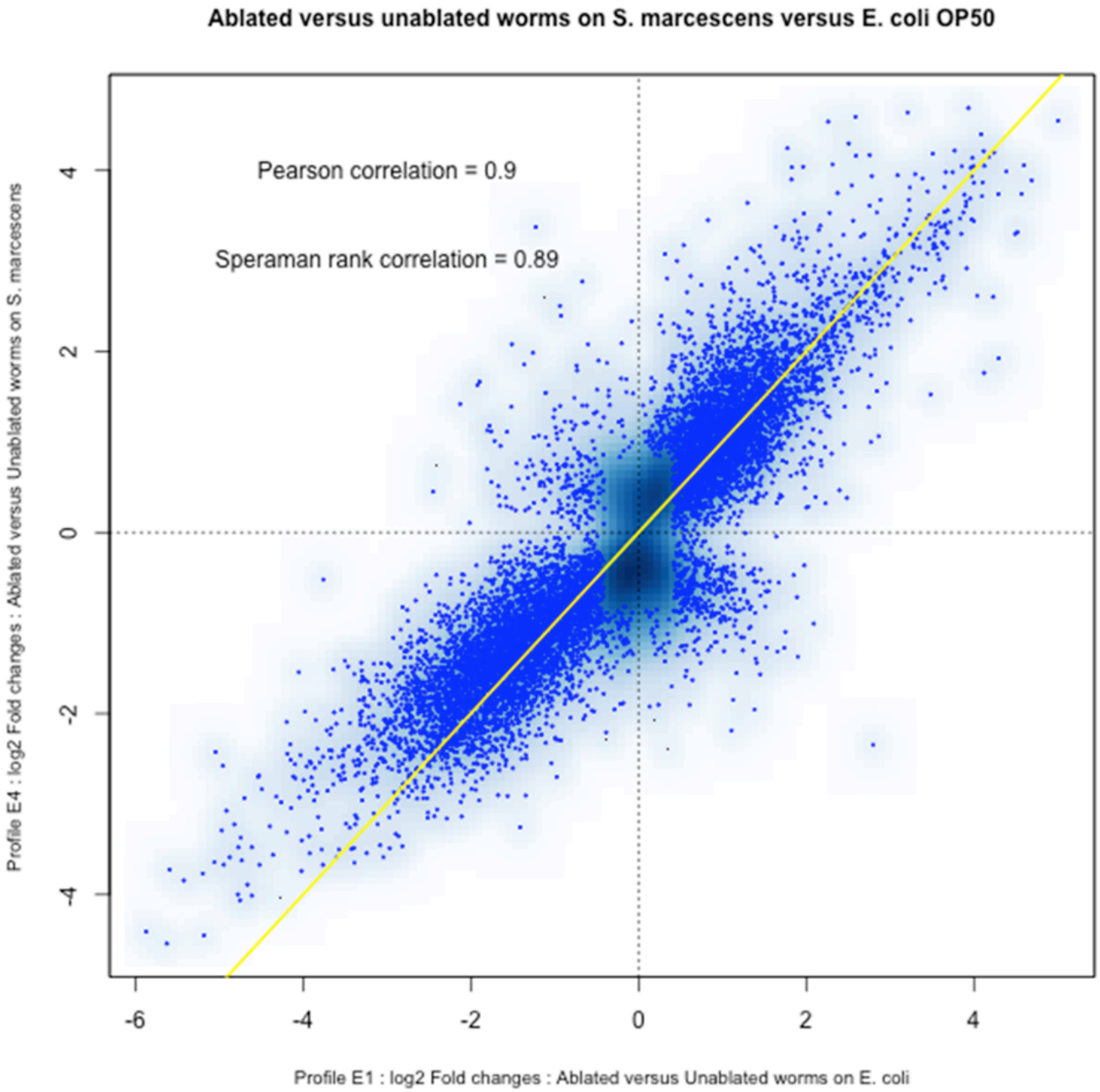
Supporting Information:
Rae et al PLoS Pathogens 8(8) 2012

Figure S1.
Effect of Z2 and Z3 germline ablation on survival of *P. pacificus* insulin signaling mutants exposed to *S. marcescens*. Survival of *Ppa-daf-16 (tu901)* Z2 and Z3 ablated (blue) and un-ablated (red), and *Ppa-daf-16 (tu302)* Z2 and Z3 ablated (yellow) and un-ablated (green) exposed to S. marcescens. Error bars represent ± S.E.M.



Supporting Information:
Rae et al PLoS Pathogens 8(8) 2012

Figure S2.
Comparison of fold-changes in expression profiles E4 (ablated versus unablated animals exposed to *S. marcescens*) and E1 (ablated versus unablated animals exposed to *E. coli*). The two profiles are quite similar, with almost similar fold-changes for the majority of the genes. (Pearson correlation = 0.90, Spearman's rank correlation = 0.89).



Supporting Information:
Rae et al PLoS Pathogens 8(8) 2012

Table S1.

Summary statistics of *P. pacificus* ablation experiments monitoring survival when fed *S. marcescens* and *X. nematophila*. Mean survival and standard errors for all conditions tested, and p-values from log Rank test assessing significance of difference between various comparisons. The rows 8 and 9 (marked with an "*") correspond to the pathogen *X. nematophila.*

	Pristionchus treatment	Number of nematodes tested	Mean survival (days)	Standard error	Comparision	Log rank P value
1	*P. pacificus* WT unablated	120	2.1	0.09		
2	*P. pacificus* WT Z2+Z3 ablated	80	5.3	0.37	1 vs. 2	0.000
3	*P. pacificus* WT Z1+Z4 ablated	60	1.2	0.05	1 vs. 3	0.000
4	*Ppa-daf-16(tu302)* Z2+Z3 ablated	40	2.6	0.24	1 vs. 4	0.000
5	*Ppa-daf-16(tu901)* Z2+Z3 ablated	80	2.9	0.14	1 vs. 5	0.000
6	*Ppa-daf-12(tu389)* Z2+Z3 ablated	60	3.9	0.28	1 vs. 6	0.000
7	*Ppa-daf-12(tu390)* Z2+Z3 ablated	80	4.3	0.26	1 vs. 7	0.002
8	*P. pacificus* WT unablated*	60	1.9	0.12		
9	*P. pacificus* WT Z2+Z3 ablated*	60	2.8	0.25	8 vs. 9	0.002
10	*Pristionchus* sp. 3 unablated	60	2.8	0.22		
11	*Pristionchus* sp. 3 Z2+Z3 ablated	54	4.8	0.51	10 vs. 11	0.000
12	*Pristionchus* sp. 16 unablated	60	1.7	0.08		
13	*Pristionchus* sp. 16 Z2+Z3 ablated	60	5.0	0.24	12 vs. 13	0.000
14	*P. pacificus* PS160 unablated	60	1.3	0.06		
15	*P. pacificus* PS160 Z2+Z3 ablated	60	2.6	0.14	14 vs. 15	0.000
16	*P. pacificus* M2 unablated	60	1.7	0.05		
17	*P. pacificus* M2 Z2+Z3 ablated	60	5.7	0.32	16 vs. 17	0.000
18	*P. pacificus* S264 unablated	60	1.8	0.07		
19	*P. pacificus* S264 Z2+Z3 ablated	60	6.2	0.28	18 vs. 19	0.000
20	*Ppa-daf-16(tu302)* unablated	40	1.8	0.11		
21	*Ppa-daf-16(tu302)* Z2+Z3 ablated	40	2.6	0.24	20 vs. 21	0.001
22	*Ppa-daf-16(tu901)* unablated	60	1.9	0.09		
23	*Ppa-daf-16(tu901)* Z2+Z3 ablated	80	2.9	0.14	22 vs. 23	0.000
24	*Ppa-daf-12(tu389)* unablated	60	2.2	0.09		
25	*Ppa-daf-12(tu389)* Z2+Z3 ablated	60	3.9	0.28	24 vs. 25	0.000
26	*Ppa-daf-12(tu390)* unablated	80	2.1	0.07		
27	*Ppa-daf-12(tu390)* Z2+Z3 ablated	80	4.3	0.26	26 vs. 27	0.000

This paper is available on line at http://www.mcponline.org

Phosphoproteome of *Pristionchus pacificus* Provides Insights into Architecture of Signaling Networks in Nematode Models

Nadine Borchert‡§, Karsten Krug§¶, Florian Gnad‖, Amit Sinha‡, Ralf J. Sommer‡‡‡, and Boris Macek¶‡‡**

Pristionchus pacificus* is a nematode that is increasingly used as a model organism in evolutionary biology. The genome of *P. pacificus* differs markedly from that of *C. elegans*, with a high number of orphan genes that are restricted to *P. pacificus* and have no homologs in other species. To gain insight into the architecture of signal transduction networks in model nematodes, we performed a large-scale qualitative phosphoproteome analysis of *P. pacificus*. Using two-stage enrichment of phosphopeptides from a digest of *P. pacificus* proteins and their subsequent analysis via high accuracy MS, we detected and localized 6,809 phosphorylation events on 2,508 proteins. We compared the detected *P. pacificus* phosphoproteome to the recently published phosphoproteome of *C. elegans.* The overall numbers and functional classes of phosphoproteins were similar between the two organisms. Interestingly, the products of orphan genes were significantly underrepresented among the detected *P. pacificus* phosphoproteins. We defined the theoretical kinome of *P. pacificus* and compared it to that of *C. elegans*. While tyrosine kinases were slightly underrepresented in the kinome of *P. pacificus*, all major classes of kinases were present in both organisms. Application of our kinome annotation to a recent transcriptomic study of dauer and mixed stage populations showed that Ser/Thr and Tyr kinases show similar expression levels in *P. pacificus* but not in *C. elegans*. This study presents the first systematic comparison of phosphoproteomes and kinomes of two model nematodes and, as such, will be a useful resource for comparative studies of their signal transduction networks. *Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.022103, 1631–1639, 2012.

Pristionchus *pacificus* is a nematode that is established as a model in evolutionary developmental biology (2). Like *Caenorhabditis elegans*, which was the first multicellular organism to have its genome completely sequenced (3), it has several advantageous features: it is easy to cultivate in the laboratory, it feeds on *E. coli*, it has a short generation time of 4 days (at 20 °C), and, because it is a self-fertilizing hermaphrodite, it is amenable to forward and reverse genetic techniques. Its genome has recently been sequenced, revealing a high number of predicted genes that share no sequence similarity to genes from any other organisms ("orphan" or "pioneer" genes) (4). Like many other nematodes, *P. pacificus* exhibits phenotypic plasticity of its life cycle and is able to quickly adapt to different environmental conditions. Under favorable conditions, *P. pacificus* undergoes direct development, but it can arrest development to form a stress resistant dauer stage when the environmental conditions turn unfavorable. These examples of phenotypic plasticity have allowed nematodes to invade many different habitats (5). *P. pacificus* occupies a completely different ecological niche than *C. elegans*. It has a necromenic lifestyle in which the developmentally arrested dauer larva infests a scarab beetle and resumes development upon the beetle's death, feeding on the microorganisms that decompose the beetle's carcass (6). The estimated evolutionary distance between *C. elegans* and *P. pacificus* is 250 to 420 million years, which makes them very attractive models in evolutionary developmental biology (4).

We recently performed a comprehensive analysis of the proteome and transcriptome of *P. pacificus*, with the aim of refining its genome annotation. Retraining the gene prediction algorithm with gene expression data estimated the number of predicted open reading frames to 24,000. Comparison of our data to the predicted proteome of *C. elegans* revealed differences in the proteome structures of the two nematodes. Whereas the predicted proteome of *P. pacificus* showed a unimodal distribution of protein sizes, the proteome of *C. elegans* followed a clearly bimodal distribution. Interestingly, this bimodal distribution seemed to be connected to functions related to protein phosphorylation, suggesting a potential difference in protein phosphorylation between the two organisms (7).

To gain further insights into the proteome of *P. pacificus*, we performed a large-scale analysis of *P. pacificus* phosphopro-

From the ‡Department for Evolutionary Biology, Max Planck Institute for Developmental Biology, Tuebingen, Germany; ¶Proteome Center Tuebingen, University of Tuebingen, Germany; ‖Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany

Received July 6, 2012, and in revised form, August 21, 2012

Published, MCP Papers in Press, August 25, 2012, DOI 10.1074/mcp.M112.022103

teome using phosphopeptide enrichment and high accuracy mass spectrometry. Here we report the first comprehensive phosphoproteome map of *P. pacificus,* measured to a depth of almost 7,000 localized phosphorylation sites, and compare it to the recently reported phosphoproteome of *C. elegans* (8). We show that the two phosphoproteomes are of similar sizes but differ significantly in the frequencies of phosphorylated serine, threonine, and tyrosine residues. We define direct orthologs between the two organisms and show that this discrepancy is also pronounced at the ortholog level. We show that the products of orphan genes are significantly underrepresented among the detected *P. pacificus* phosphoproteins. Finally, we define the predicted kinome of *P. pacificus* and show that it is slightly smaller than that of *C. elegans* but contains all major classes of kinases.

MATERIALS AND METHODS

Culturing of Worms and Preparation of Protein Extracts—*P. pacificus* strain PS312 was grown on 10 cm NGM agar plates spotted with 2 ml *E. coli* OP50 solution. Plates were inoculated with between 50 and 100 worms and incubated at 25 °C. The mixed stage population was harvested shortly after the bacterial lawn was consumed, avoiding the starvation of the animals. After thorough washing with distilled water and 0.9% sodium chloride, worms were pelleted and prepared for proteomics measurements.[1]

For protein isolation, 100 μl of animals were solubilized in 300 μl denaturation buffer (6 M urea, 2 M thiourea, 10 mM Tris pH 8.0). After three cycles of freezing (liquid nitrogen) and thawing (37 °C), 100 μl of glass beads were added, and the solution was vortexed for 20 min. After centrifugation (20 min, 20.800 $\times$ *g*, 4 °C), the protein concentration of the supernatant was determined using the Bradford assay and further processed using the filter-aided sample preparation (FASP)[2] method (9) (see below).

Protein Digestion—The soluble protein fraction was digested as described previously (7). Briefly, 5 mg of protein was reduced with a final concentration of 1 mM DTT and alkylated with a final concentration of 5.5 mM iodoacetamide. After the pH was adjusted to 8.0, 1 μg of trypsin was added per 100 μg of protein, and the mixture was incubated overnight at 37 °C.

The insoluble protein fraction was processed with a modified FASP protocol (9). The protein pellet was solubilized in 4% SDS, 100 mM DTT, and 100 mM Tris pH 7.6. An aliquot of the pellet was precipitated with chloroform/methanol and solubilized in denaturation buffer for Bradford analysis. Based on the Bradford measurement, a protein-SDS solution containing 5 mg of protein was diluted with urea buffer A (8 M urea in 100 mM Tris, pH 8.5) to a final volume of 6 ml and pipetted into the 15-ml Centriprep column YM-30 (Millipore, Billerica, MA). After the sample had been spun for 15 min at 6,000 g, 600 μl of iodoacetamide solution (550 mM) was added, and the sample was incubated for 1 h in the dark and then centrifuged for 15 min at 3,000 g. The protein was washed with UA three times, and the last centrifugation step was increased to 20 min. Six ml of ammonium bicarbonate was added, and the sample was centrifuged for another 15 min at 3,000 g. After that, Trypsin was added at a final concentration of 1 μg per 100 μg total protein and incubated overnight at 37 °C. After the next centrifugation step (15 min, 3,000 g), the peptides were collected in the flowthrough. Centrifugation was repeated with 3 ml water, and the flowthrough was collected for strong cation exchange (SCX).

Phosphopeptide Enrichment—After 5 mg of digested total protein lysate had been acidified to pH 2.7 with trifluoroacetic acid, the sample was loaded onto an ÄKTApurifier (GE Healthcare, Little Chalfont, UK) HPLC for SCX. The 16 resulting fractions were pooled according to the elution profile to 10 fractions for titanium dioxide enrichment. Five mg of TiO_2 beads were resuspended in 50 μl of a 30 mg/ml 2.5 dihydrobenzoic acid, 80% acetonitrile in water solution. After 10 min of incubation at room temperature, the TiO_2 loading solution was added to the sample and mixed for 30 min at room temperature using an orbital shaker. The beads were precipitated with centrifugation at 13,000 rpm for 2 min and washed with 1 ml Wash Solution I (30% acetonitrile (ACN), 3% TFA) for 10 min in a shaker and Wash Solution II (10) (80% ACN, 0.1% TFA) for 10 min in a shaker. The beads were then resuspended in 50 μl Wash Solution II and transferred to a 200 μl pipette tip plugged with one layer of Empore C8 tip. After the beads had been washed three times with 100 μl 40% ammonia solution (25% in water) in ACN pH 10.5, the eluate was reduced to 5 μl in a SpeedVac.

NanoLC-MS/MS Analysis—Enriched phosphopeptide mixtures were separated via Easy-LC nano-HPLC (Proxeon Biosystems, Odense, DK) coupled to an LTQ-Orbitrap-XL (Thermo Fisher Scientific) through a nano-LC-MS interface (Proxeon Biosystems). Chromatographic separation of the peptides was performed on a 15 cm fused silica emitter with a 75 μm inner diameter (Proxeon Biosystems), in-house packed with reversed-phase ReproSil-Pur C18-AQ 3 μm resin (Dr. Maisch GmbH, Ammerbuch-Entringen, DE). The peptide mixtures were injected onto the column in HPLC solvent A (0.5% acetic acid) at a flow rate of 500 nl/min and subsequently eluted with a 107 min segmented gradient of 2% to 80% of HPLC solvent B (80% acetonitrile in 0.5% acetic acid) at a flow rate of 200 nl/min.

The MS was operated in the data-dependent mode so as to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra were acquired in the mass range from *m/z* 300 to 2,000 in the orbitrap mass analyzer at a resolution of 60,000. An accumulation target value of 10^6 charges was set, and the lock mass option was used for internal calibration (11). The five most intense ions were sequentially isolated and fragmented in the linear ion trap using collision-induced dissociation (CID) at an ion accumulation target value of 5,000 and default CID settings. Multistage activation (at −98, −49, and −32.66 Th relative to the precursor ion) was used to optimize fragmentation of Ser/Thr phosphopeptides. The ions already selected for MS/MS were dynamically excluded for 90 s. The resulting peptide fragment ions were recorded in the linear ion trap. In total, 41 LC-MS measurements were performed, corresponding to 4 days of measurement time.

Data Processing and Analysis—MS data were processed with MaxQuant (12), version 1.0.14.3. Peak lists were generated and subsequently submitted to the Mascot search engine (Matrix Science, London, UK) to query a database consisting of the latest annotation of *P. pacificus* (dataset "HYBRID1 proteomics gene models"; 24,231

[1] The data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hash: pxGey/Jh9q186pz5hyUKK13Idzf8sjFVLW+ZZbNgv0IkOAH71q31oIfK2vNyvp8wb7ItfBczkQ8O5W/llVxLtpPhpEoAAAAAAAA1Ow==.

The hash may be used to prove exactly what files were published as part of this manuscript's data set, and the hash may also be used to check that the data have not changed since publication. The data can also be viewed through the PHOSIDA database www.phosida.com (1).

[2] The abbreviations used are: CID, collision-induced dissociation; EC, enzyme commission number; EGF, epidermal-growth factor; FASP, filter-aided sample preparation; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; Pfam, protein families and domains; SCX, strong cation exchange; VPC, vulva precursor cell.

protein entries) (7), 4,256 *E. coli* proteins, 262 commonly observed protein contaminants, and 28,749 reversed sequences. The initial precursor mass tolerance was set to 7 ppm for Orbitrap data (full scans); fragment ion mass tolerance was set to 0.5 Da for ion trap data (MS/MS scans). Full trypsin specificity was required, and up to two missed cleavages were allowed. Carbamidomethylation on cysteine was defined as fixed modification; methionine oxidation, protein N-terminal acetylation, and phosphorylation on serine, threonine, and tyrosine were defined as variable modifications. The database search results were parsed by MaxQuant to assemble protein groups, peptides, and phosphorylation sites at a false discovery rate of 1%. All phosphorylation events having a reported localization probability of at least 0.75 were considered as localized (assigned to a specific amino acid). Subsequent downstream analysis of the result tables was done in R v2.11.1 (13).

Determination of Orthologous and Homologous Relationships—Pairwise orthologs and homologs between *P. pacificus* and *C. elegans* were inferred using bidirectional and unidirectional BLASTP, respectively (14, 15). We used Wormbase WS200 for *C. elegans* and the latest genome annotation for *P. pacificus* (7) as input. Global alignments between orthologous proteins were derived using Needle (16, 17).

Determination of Orphan/Pioneer Proteins—Orphan proteins were defined by two BLAST analyses. First we regarded every *P. pacificus* protein having no homologue in the NCBInr database (BLASTP E-value $< 1 \times 10^{-3}$) as a potential orphan. Second, we used the information derived from the pairwise BLAST analysis of the theoretical proteomes of *P. pacificus* and *C. elegans* as described above. Orphan proteins were required to have no homologues in the NCBInr database or in the Wormbase WS200.

Functional Annotation of the P. pacificus Proteome—Blast2GO software was used to derive Gene Ontology (GO) (18) terms via a BLAST search of the theoretical proteome of *P. pacificus* against the nonredundant NCBI protein database (downloaded on April 29, 2010) using default parameters. Information on specific pathways on the basis of Kyoto Encyclopedia of Genes and Genomes (KEGG) terms (19) was obtained from the KEGG Automatic Annotation Server (20) using default parameters. The classification of proteins into protein families was performed using Pfam (21). The significance E-value threshold was gathered by the software automatically. All types of annotation were merged and exported to an Excel sheet using R.

Functional Enrichment Analysis of the Detected Phosphoproteome—The frequencies of functional annotation terms assigned to the detected phosphoproteome were tested against the corresponding frequencies in the entire proteome using Fisher's exact test (one-sided). A minimum of five occurrences of each term was required in order for the term to be taken into account for analysis. Derived *p* values were further adjusted for multiple hypothesis testing using the method proposed by Benjaminii and Hochberg (22).

Draft Kinome Annotation—We considered all proteins having predicted Pfam domains "Pkinase," "Pkinase_C," or "Pkinase_Tyr" as potential kinases. In order to classify these kinases into kinase groups, families, and subfamilies, we performed a BLAST search of predicted kinase domains against all nematode-specific kinase domains contained in Kinbase. BLAST hits were considered significant if the reported E-value was below 1×10^{-20}, resulting in a minimal bit score of 90.9. For further validation, we did a second BLAST search by querying the kinase domains contained in Kinbase against the predicted Pfam domains in the *P. pacificus* proteome and checking whether the results were consistent. All predicted Pfam domains that met these criteria were classified according to Kinbase annotation.

Phylogenetic distances between the domains were estimated by ClustalW and exported to Nexus format. Distances were logarithmized and imported into the Interactive Tree of Life online tool (23) to produce the phylogenetic trees. The trees were annotated with kinase groups using the classification obtained by the BLAST analysis described above.

Secondary Protein Structure Prediction—The secondary structures of all phosphorylated proteins detected in *P. pacificus* and *C. elegans* were calculated using PsiPred v3.3 (24) and PSIBLAST v2.2.23. Initial PSIBLAST searches were done against the Unriref90 database. Prior to the search, low complexity regions were removed from that database as described in the README file of the Psipred software.

Comparison with Transcriptome Data—The gene expression data from a dauer versus mix-stage comparison in both *C. elegans* and *P. pacificus* were obtained from Sinha *et al.* (25). For *P. pacificus*, the gene predictions and, hence, the gene identifiers used in Ref. 25 are different from those used in Ref. 7, although the underlying genome assembly is the same. Thus, the mapping from a microarray probe to a gene prediction corresponding to Ref. 7 was calculated using stringent BLAST criteria (E-value $< 1 \times 10^{-10}$, 100% identity between the 60 bp microarray probe and the target gene). Probes that matched multiple genes were removed from the analysis, and fold-changes were calculated using the same parameters and methods as in Ref. 25. Pfam domain annotations of *C. elegans* were based on wormpep-210. We used kinase domain annotation for *P. pacificus* from supplemental Table 2. The average expression values of all the kinase genes (expression ratio "Dauer/Mix-stage") for all genes annotated with a particular kinase domain ("Pkinase" or "Pkinase_Tyr") were compared within species, and the significance of the difference was assessed based on two-sample Wilcoxon tests. The number *n* in Fig. 5 is the total number of genes belonging to a particular gene family. The "average expression" is defined as log2(RedSignal * GreenSignal) on an arbitrary scale. Hence the values can be compared only within a nematode species and should not be compared across nematodes. The *P. pacificus* and *C. elegans* fold-change and average expression data on kinases are included in supplemental Table 5.

RESULTS

In this study we aimed to provide the reference phosphoproteome of the nematode model *P. pacificus* and compare it to the recently published phosphoproteome of *C. elegans* (8). To minimize experimental bias and enable direct comparison between the datasets, we employed similar sample preparation, measurement, and data processing workflows as in the phosphoproteomic study of *C. elegans*. Briefly, we lysed a well-fed mixed stage *P. pacificus* culture by rupturing the cuticle with freeze-thaw cycles and glass bead treatment. We extracted the proteins from the insoluble fraction in 4% SDS and processed them via the FASP protocol as described by Wisniewski *et al.* (9). We digested the soluble protein fraction in solution with trypsin and separately subjected both fractions to two stages of phosphopeptide enrichment, consisting of strong cation exchange and TiO_2 chromatographies (26, 27). We performed LC-MS analysis on an Easy-LC (Proxeon Biosystems) coupled to an LTQ-Orbitrap XL MS (Thermo Fisher Scientific) and processed the data using the MaxQuant software suite (12). The workflow employed in this study is depicted in Fig. 1.

Detected Phosphoproteome of P. pacificus—The analysis of the *P. pacificus* phosphoproteome resulted in 60,358 identified MS/MS spectra that detected 9,872 nonredundant pep-

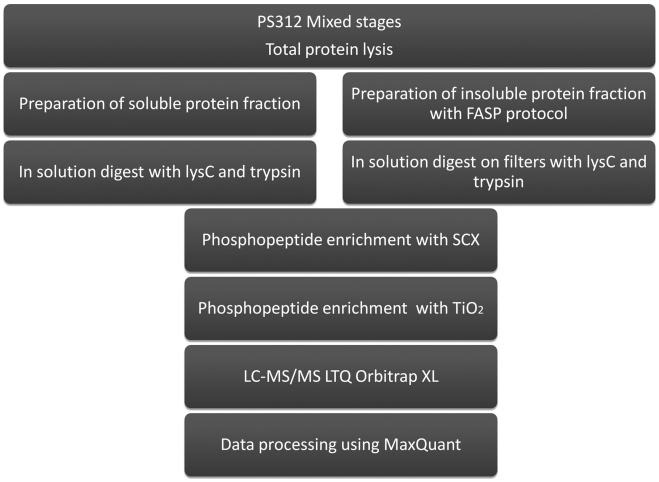


FIG. 1. **Biochemical workflow used in this study.** A mixed population of *P. pacificus* worms was harvested, and the protein extract was split into soluble and insoluble fractions, which were processed further using the FASP protocol. After LysC/trypsin digestion, phosphopeptides were enriched by SCX and TiO2 chromatographies and measured on an LTQ Orbitrap XL mass spectrometer.

TABLE I

Number of (phospho)proteins detected in this study (at 1% false discovery rate (FDR))

MS data were searched against a decoy database containing *P. pacificus* and *E. coli* protein entries.

	Phosphoproteins	All proteins
P. pacificus	2,508	3,158
E. coli	11	23

tide sequences with a median absolute mass deviation of 255 ppb (supplemental Fig. 1). We detected 3,158 *P. pacificus* protein groups at a false discovery rate of 1%; of these, 2,508 were phosphorylated (Table I) and 1,518 were not detected in our previous large-scale proteomics study (7). This resulted in extension of the catalogue of *P. pacificus* proteins detected by MS to 5,547 (supplemental Fig. 2). In total, we localized 6,809 phosphorylation events to a specific amino acid residue with a median confidence level of 99.8%. The frequencies of phosphorylated serines, threonines, and tyrosines were found to be 87.8% (5,981 events), 11.1% (756 events), and 1.06% (72 events), respectively. All detected phosphorylation sites are presented in supplemental Table 1.

Functional Classes and Kinase Motifs of Detected P. pacificus Phosphoproteins—To gain insight into the functional distribution of proteins phosphorylated in *P. pacificus*, we first retrieved the latest functional annotation according to GO terms, KEGG pathways, enzyme commission numbers (ECs), and protein families and domains (Pfam) (supplemental Table 2). We then performed enrichment analyses of the GO, KEGG, Pfam, and EC terms of proteins detected as phosphorylated (supplemental Table 3). The GO term analysis showed an enrichment of functions related to protein and nucleoside binding, transcription repressor activities, and kinase regulator activities, terms commonly enriched in large phosphoproteome datasets. The Enzyme Class analysis showed significant enrichment of only two classes, protein tyrosine kinases (EC 2.7.10.0; 23 detected phosphoproteins) and protein serine kinases (EC 2.7.11.0; 35 detected phosphoproteins). This was expected because kinases and phosphatases themselves are commonly regulated by phosphorylation, and many kinases show autophosphorylation activity. In agreement with this, the Pfam analysis showed an enrichment of protein kinase domains, as well as phosphatase domains. Proteins with domains involved in protein–protein interactions and signaling were also overrepresented in comparison with the total gene predictions. Among the detected domains, WD40, VWD, Ankyrin, and PDZ domains were highly represented. Moreover, RNA binding domains such as rrm-1, helicase, and DEAD were also overrepresented. The results of the functional enrichment analysis are summarized in Fig. 2.

We next tested the representation of *P. pacificus* orphan gene products in the phosphoproteome. In total, we detected phosphorylation on 234 products of orphan genes (9.3% of the detected phosphoproteome). Compared with all orphan genes in the *P. pacificus* genome (9,957; 41.09% of the genome), this presented a significant underrepresentation ($p < 3.64 \times 10^{-303}$). However, it has to be noted that this class of gene products showed a similar underrepresentation at the proteome level (7), pointing to the fact that their underrepresentation in the phosphoproteome results from the lack of expression, not phosphorylation.

Next, we tested the enrichment of specific kinase target motifs on *P. pacificus* phosphoproteins detected in our dataset, as described by Zielinska *et al.* (8). On proteins phosphorylated on serine, three motifs were overrepresented—CAMK2 (RXX[pS]), CK2 ([pS]XXE), and PKA (RX[pS])—whereas on proteins phosphorylated on threonine, only the CAMK2 motif was overrepresented. For both phosphorylated residues there was also significant overrepresentation of proline adjacent to the phosphorylation site ([pS]P and [pT]P) (supplemental Fig. 3). No significant motifs were detected on proteins phosphorylated on tyrosine residues, most likely because of the small size of the dataset.

Comparison of P. pacificus and C. elegans Phosphoproteomes—We next compared the phosphoproteome of *P. pacificus* to the recently published phosphoproteome of *C. elegans* (8), in which 6,699 phosphorylation sites were localized on 2,365 proteins (Table II). The sizes of the two phosphoproteomes were very similar, and the enriched functional classes of detected phosphoproteins were almost identical, demonstrating that both nematodes likely use protein phosphorylation in similar biological processes. Interestingly, the two phosphoproteomes differed in frequencies of S/T/Y phosphorylation events. In *C. elegans*, the reported pSer, pThr, and pTyr frequencies were 80.2%, 18%, and 1.8%, whereas in *P. pacificus* they were 87.8%, 11.1%, and 1.1%, respectively.

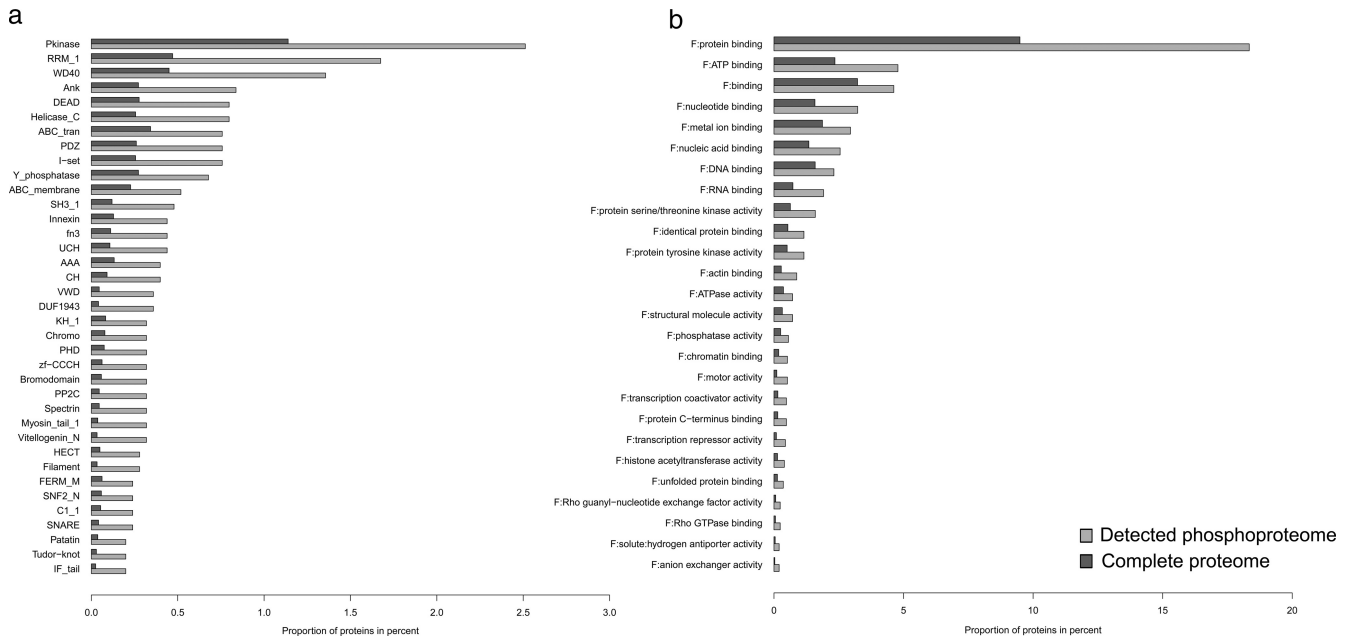


FIG. 2. **Functional enrichment analysis of the detected *P. pacificus* phosphoproteome.** *a*, enrichment of Pfam terms; *b*, enrichment of GO terms (molecular function).

TABLE II

Numbers and frequencies of phosphorylation sites localized on serine, threonine, and tyrosine in P. pacificus and C. elegans

To test whether the frequencies of pS, pT, and pY were significantly different between the two nematodes, we calculated *p* values using a two-sided binomial test.

	Total	pS	pT	pY
Wormbase200	23,973 proteins	7.81%	5.85%	2.75%
Ppa database	24,231 proteins	8.12%	5.89%	3.13%
C. elegans (Zielinska et al. (8))	6,699 (2,365 proteins)	5,372 (80.19%)	1,207 (18.02%)	120 (1.79%)
P. pacificus (this study)	6,809 (2,401 proteins)	5,981 (87.84%)	756 (11.1%)	72 (1.06%)
Binomial *p* value		$p < 2.2 \times 10^{-16}$	$p < 2.2 \times 10^{-16}$	$p < 1.18 \times 10^{-6}$

The frequencies of all phosphorylated amino acids were significantly different despite very similar overall frequencies of these amino acids in the proteomes of *P. pacificus* and *C. elegans* (Table II).

To gain insight into the potential origin of this discrepancy, we investigated the frequencies of pSer, pThr, and pTyr in orthologs shared between *P. pacificus* and *C. elegans* and therefore likely present in their common ancestor. Based on the bidirectional BLASTP approach, 619 phosphoproteins from our dataset were defined as orthologs between *P. pacificus* and *C. elegans* and phosphorylated in both species. On these orthologs, 340 phosphorylation sites were determined as conserved (Fig. 3; supplemental Table 4). Interestingly, the frequencies of pSer, pThr, and pTyr at the ortholog level (90%, 9.1%, and 0.9%, respectively) resembled more closely the frequencies measured in the phosphoproteome of *P. pacificus* than those in the phosphoproteome of *C. elegans*. This means that the basal phosphoproteome of *P. pacificus* might resemble the phosphoproteome of the common ancestor of *P. pacificus* and *C. elegans*.

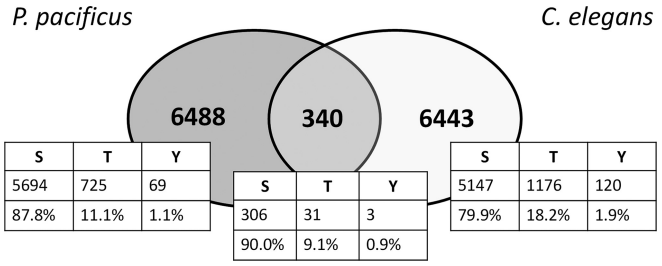


FIG. 3. **Evolutionary conserved phosphorylated residues between *P. pacificus* and *C. elegans*.** Venn diagram depicting the overlap of conserved phosphorylation sites on direct orthologs found to be phosphorylated in the phosphoproteome datasets.

Whereas different frequencies of tyrosine phosphorylation may be explained by different usages of this modification in signal transduction (see Discussion), different frequencies of detected serine and threonine phosphorylation are more difficult to explain, mostly because of the dual specificity of Ser/Thr kinases. A potential reason could be different representation of these amino acids in the unstructured protein regions that are more accessible to protein kinases. To test

TABLE III
Number of predicted protein kinases in different nematodes according to Pfam annotations of protein kinase domains

	Domains	Proteins
C. elegans	441 (119 pTyr)	413 (117 pTyr)
P. Pacificus	408 (102 pTyr)	368 (94 pTyr)
B. malayi	406 (89 pTyr)	378 (83 pTyr)
M. incognita	392 (57 pTyr)	361 (53 pTyr)

this hypothesis, we calculated the frequencies of all serine, threonine, and tyrosine residues in coiled coils and in helical and strand regions of proteins from detected phosphoproteomes of *P. pacificus* and *C. elegans* and compared them to detected phosphorylation sites (supplemental Fig. 4). As expected, this analysis did not reveal any significant differences in the frequencies of serine, threonine, and tyrosine in the two organisms, demonstrating that different accessibility is not the reason for the observed differences in frequencies of phosphorylated amino acids.

Predicted Kinome of P. pacificus and Its Comparison with C. elegans—We next compared the predicted kinomes of several sequenced model nematodes. To define the predicted kinomes of *P. pacificus*, we used Pfam annotation and considered all proteins containing a "P-kinase" domain as potential kinases (see Methods). After collapsing all *C. elegans* kinase isoforms, we compared the predicted kinome to that of *P. pacificus*. The kinome of *P. pacificus* contained 368 kinases (supplemental Table 2) and was 11% smaller than that of *C. elegans*, which contained 413 kinases (Table III); interestingly, the number of predicted tyrosine kinases was 20% lower in *P. pacificus* (94 kinases) and therefore was underrepresented relative to *C. elegans* (117 kinases). Of the 368 predicted kinases in *P. pacificus*, 77 were detected as phosphorylated in our study. Of those, 61 had direct orthologs and 30 were detected as phosphorylated in *C. elegans* (8). Interestingly, two of the three (66.6%) conserved pTyr residues were located on kinases (cdk-1, mbk-1), one of the 31 (3.2%) conserved pThr residues was located on a kinase (sek-1), and 12 of 306 (3.9%) conserved pSer residues were located on kinases (unc-82, unc-22, pkc-1, grk-1, ZK524.4, gcy-28, ZC581.9, B0495.2).

To classify *P. pacificus* kinases into groups, families, and subfamilies, we performed a bidirectional BLAST analysis of predicted kinase domains against *C. elegans* kinase domains contained in Kinbase. The BLAST analysis resulted in 282 highly confident hits, indicating that the catalytic domains of predicted kinases appeared to be conserved between the two nematodes. All eight major protein kinase groups present in *C. elegans* were also present in *P. pacificus* (Fig. 4; supplemental Fig. 5).

Expression of Different Kinase Classes in C. elegans and P. pacificus—To assess the expression of different kinase classes in *C. elegans* and *P. pacificus*, we analyzed a recently published transcriptome dataset that addresses global changes in gene expression in the dauer and mixed populations of these two nematodes (25). Applying our Pfam-based kinome annotation, we extracted expression data for 404 kinases in *C. elegans* (288 Ser/Thr and 116 Tyr kinases) and 316 kinases in *P. pacificus* (239 Ser/Thr and 77 Tyr kinases). As expected, the transcriptome analysis showed good coverage of the kinome in both organisms, albeit slightly higher in *C. elegans* (404/413, 98%) than in *P. pacificus* (316/368, 86%). Interestingly, in *C. elegans*, the average expression of "Pkinase_Tyr" genes was significantly higher than the average expression of "Pkinase" genes in the dauer population. However, in *P. pacificus*, all kinase genes were expressed at a significantly higher level in the dauer population, and there was no difference in average expression between the two kinase categories, pointing to the fact that tyrosine kinases are expressed at levels similar to those of Ser/Thr kinases (Fig. 5). These data reveal that both nematodes express all classes of kinases and point to their potentially different usage in the dauer stage of the life cycle.

DISCUSSION

In this study, we have reported the first global phosphoproteomic dataset of the mixed stage population of the *P. pacificus* nematode. In order to increase the number of identified phosphorylation sites, we performed three biological replicates, two with the soluble and one with the insoluble protein fraction. In this way, we made all cellular compartments accessible to protein analysis. By using mixed stages, we aimed to get an in-depth catalog of phosphorylation sites of *P. pacificus* and compare it to the previously reported phosphoproteome of *C. elegans*, analyzed under similar conditions.

Although the two phosphoproteomes were very similar in terms of size, classes of phosphorylated proteins, and overrepresented kinase motifs, they were different in the extent of serine, threonine, and tyrosine phosphorylation. Interestingly, this difference might reflect the observed alterations in signal transduction during postembryonic development of these two species. Work over the past decade has compared signaling networks during vulva development and dauer formation between *P. pacificus* and *C. elegans* and identified substantial differences (28). In *C. elegans*, three vulva precursor cells (VPCs) are induced to form vulval tissue by a signal from the gonadal anchor cell. This signal is a secreted epidermal-growth factor (EGF)-type factor that is transmitted within the VPCs by EGFR-RAS-MAP kinase signaling and finally results in the initiation of cell division. A series of phosphorylation events by LIN-45/RAF, MEK-2/MAP kinase, and MPK-1/MAP kinase is at the center of *C. elegans* vulva induction (29). In *P. pacificus*, in contrast, vulva formation is regulated by a completely different regulatory mechanism (for a review, see (30)). While the same VPCs form vulval tissue, their induction requires regulatory input from Wnt signaling rather than EGF-MAP kinase signaling (31). This involves an unusual regulatory

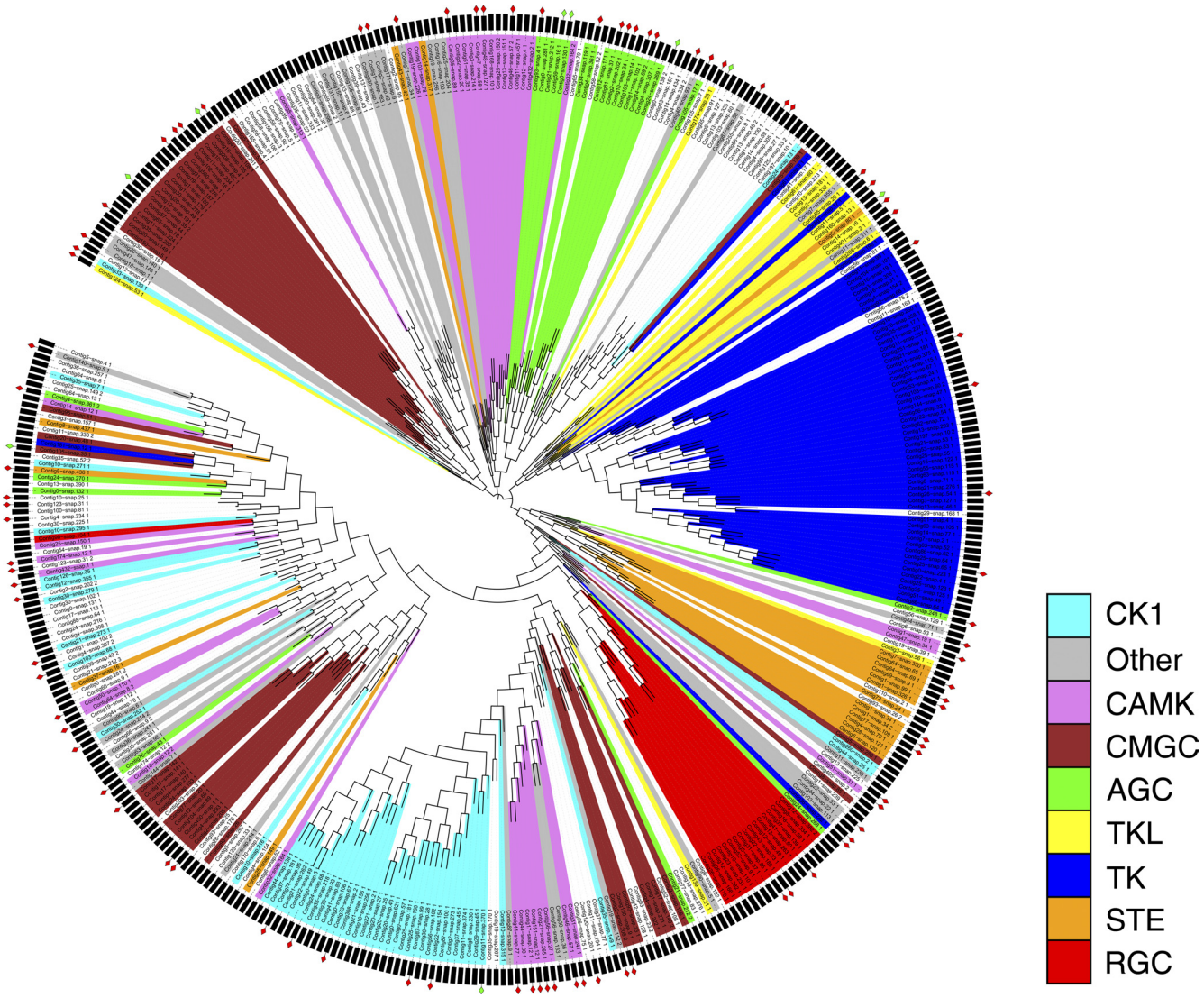


FIG. 4. **Phylogentic tree of the predicted *P. pacificus* kinome.** The tree shows the phylogentic relationships of predicted kinase domains and their classification into kinase groups according to the *C. elegans* kinome. Phylogenetic distances are based on multiple sequence alignments of predicted kinase domains. The classification of domains into kinase groups is shown by the different colors of the branches. Red rectangles at the outer edge of the circle indicate kinases that are detected as phosphorylated; green rectangles indicate kinases that are detected as nonphosphorylated in our study.

linkage of Wnt-type ligands and Frizzled-type receptors, as well as novel protein-interaction domains in LIN-18/Ryk/Derailed-type co-receptor (28). Thus, vulva induction in *C. elegans* is regulated by a kinase pathway involving a high extent of tyrosine phosphorylation, whereas the same process in *P. pacificus* depends much less on tyrosine phosphorylation. It has to be noted, however, that *P. pacificus* contains 1:1 orthologs for all of the EGF/Ras pathway genes/proteins known from *C. elegans*. Interestingly, *Ppa*-MPK-1 was the only kinase of the EGF/RAS pathway shown to be phosphorylated in our dataset. The functional significance of this finding, if any, has yet to be identified.

Similarly, work on dauer formation revealed potential differences in signaling activity during development. In *C. elegans*, the formation of dauer larvae, an arrested alternative life stage that facilitates the survival of harsh environmental conditions, involves insulin and TGF-β signaling activity that is coupled to transcriptional activity of the nuclear hormone receptor DAF-12 and the FOXO-transcription factor DAF-16 (5). In *P. pacificus*, both transcription factors have similar roles during dauer regulation, as indicated by the phenotype of mutations in the corresponding genes, whereas there is no report that would suggest similar roles of insulin and TGF-β signaling (6). However, as indicated above for vulva development, these differences in signaling activity in these two nematodes are not reflected in the copy number of genes encoding signaling components in the respective genomes. Thus, differences in phosphorylation patterns as revealed in our study can occur

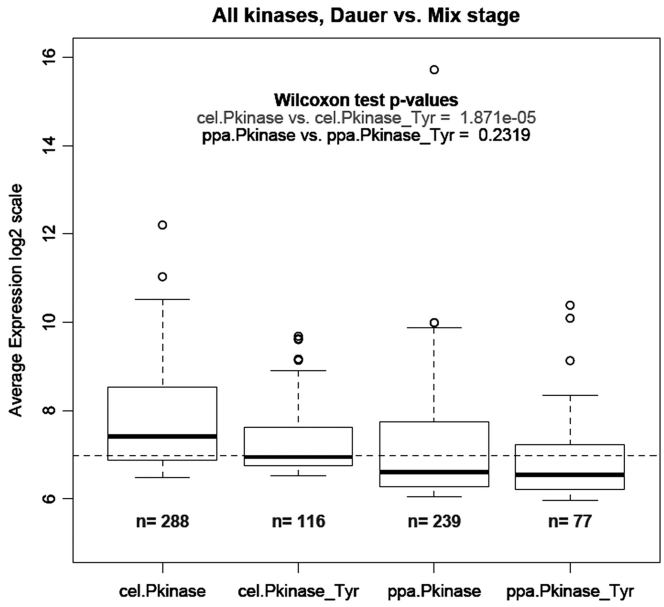


FIG. 5. **Average expression of different protein kinase classes in dauer versus mixed stages of *C. elegans* (cel) and *P. pacificus* (ppa).** Kinome annotation from supplemental Table 2 was applied to quantitative transcriptomics data derived from Sinha *et al.* (25).

in the absence of major changes in the signaling pathways that act during development.

The comparative analysis of the predicted kinomes of *P. pacificus* and *C. elegans* indicates that all major protein kinase groups are conserved between these two nematodes (supplemental Table 2), and recent transcriptome analysis suggests that all kinase classes are expressed (and presumably active) in both nematodes during the dauer stage of the life cycle (25). When compared with other protein classes, the kinome shows a relatively high level of conservation and low copy number variations. For example, many of the detoxification enzymes, such as cytochrome P450 proteins, show a more than 3-fold difference between the *P. pacificus* and *C. elegans* proteomes with 197 and 67 protein predictions, respectively (4). We speculate that the difference in cytochrome P450 enzymes reflects the adaptation to the different environments in which these nematodes are found. In contrast, the overall similarity of the two kinomes represents the conserved molecular and cellular processes, which evolved largely independent of ecological alterations. This evolutionary pattern becomes even stronger when data available for additional nematodes are considered: the numbers of predicted protein kinases of *P. pacificus*, *C. elegans,* the human parasite *B. malayi*, and the plant parasite *M. incognita* are surprisingly similar (Table III). Thus, the kinome represents a stable part of the nematode proteome, and most likely the analysis of a small number of selected model organisms will provide comprehensive insight into processes of phosphorylation.

Acknowledgments—We thank Matthias Mann for access to the Phosida database. We also thank Jonathan Goldberg for fruitful discussions on the kinome prediction.

* This work was supported by funding from Baden-Württemberg Stiftung (Juniorprofessoren-Programm), the German Research Foundation, and PRIME-XS (to B.M.).

S This article contains supplemental Figs. 1 to 5 and Tables 1 to 5.

§ These authors contributed equally to this work.

** Current address: Department of Bioinformatics and Computational Biology, Genentech Inc., San Francisco, CA.

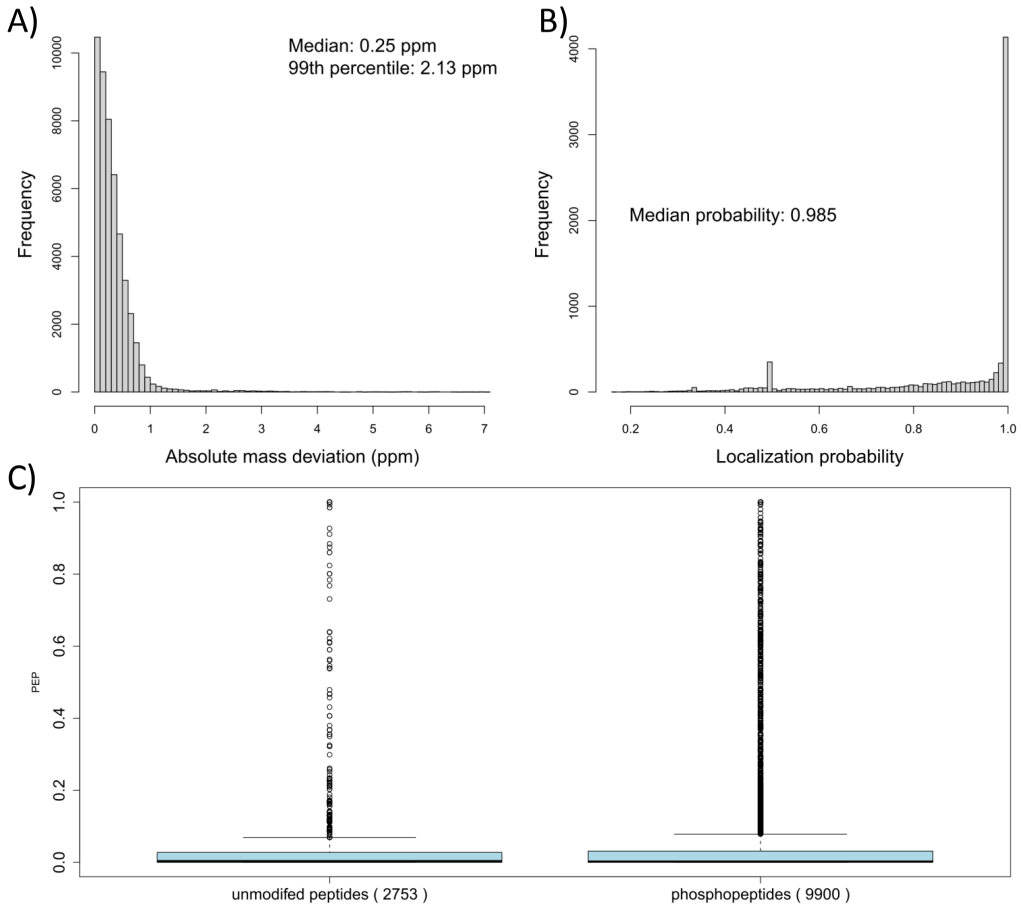
‡‡ To whom correspondence may be addressed: Prof. Dr. Boris Macek, Proteome Center Tuebingen, Interfaculty Institute for Cell Biology, University of Tuebingen, Auf der Morgenstelle 15, 72076 Tuebingen, Germany, Tel.: +49/(0)7071/29-70558, Fax: +49/(0)7071/29-5779, E-mail: boris.macek@uni-tuebingen.de

§§ To whom correspondence may be addressed: Prof. Dr. Ralf J. Sommer, Dept. for Evolutionary Biology, Max Planck Institute for Developmental Biology, Spemannstrasse 37, Tuebingen D-72076, Germany, Tel.: +49-7071-601371, Fax: +49-7071-601498, E-mail: ralf.sommer@tuebingen.mpg.de.

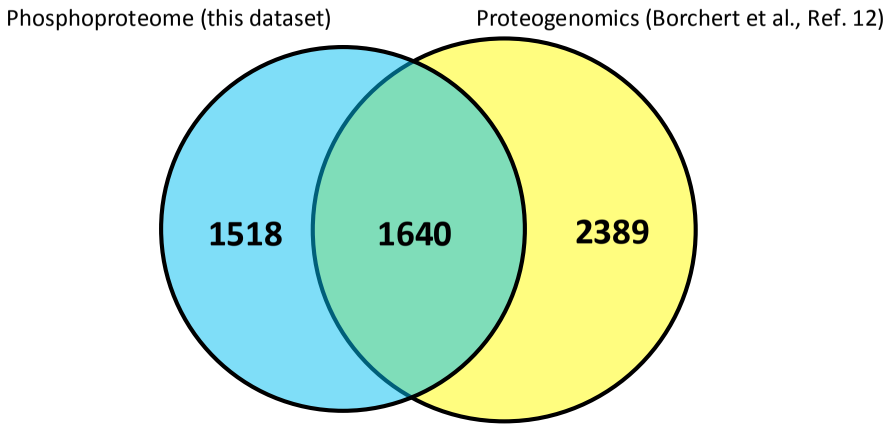
REFERENCES

1. Gnad, F., Ren, S. B., Cox, J., Olsen, J. V., Macek, B., Oroshi, M., and Mann, M. (2007) PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* **8,** R250
2. Hong, R. L., and Sommer, R. J. (2006) Pristionchus pacificus: a well-rounded nematode. *Bioessays* **28,** 651–659
3. The C. elegans Sequencing Consortium (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282,** 2012–2018
4. Dieterich, C., Clifton, S. W., Schuster, L. N., Chinwalla, A., Delehaunty, K., Dinkelacker, I., Fulton, L., Fulton, R., Godfrey, J., Minx, P., Mitreva, M., Roeseler, W., Tian, H., Witte, H., Yang, S. P., Wilson, R. K., and Sommer, R. J. (2008) The Pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism. *Nat. Genet.* **40,** 1193–1198
5. Coghlan, A. (2005) Nematode genome evolution. *WormBook* **2005,** 1–15
6. Sommer, R. J., and Ogawa, A. (2011) Hormone signaling and phenotypic plasticity in nematode development and evolution. *Curr. Biol.* **21,** R758–R766
7. Borchert, N., Dieterich, C., Krug, K., Schutz, W., Jung, S., Nordheim, A., Sommer, R. J., and Macek, B. (2010) Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models. *Genome Res.* **20,** 837–846
8. Zielinska, D. F., Gnad, F., Jedrusik-Bode, M., Wisniewski, J. R., and Mann, M. (2009) Caenorhabditis elegans has a phosphoproteome atypical for metazoans that is enriched in developmental and sex determination proteins. *J. Proteome Res.* **8,** 4039–4049
9. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6,** 359–362
10. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **78,** 2113–2120
11. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4,** 2010–2021
12. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–1372
13. R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria
14. Remm, M., Storm, C. E., and Sonnhammer, E. L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314,** 1041–1052
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410
16. Needleman, S. B., and Wunsch, C. D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins.

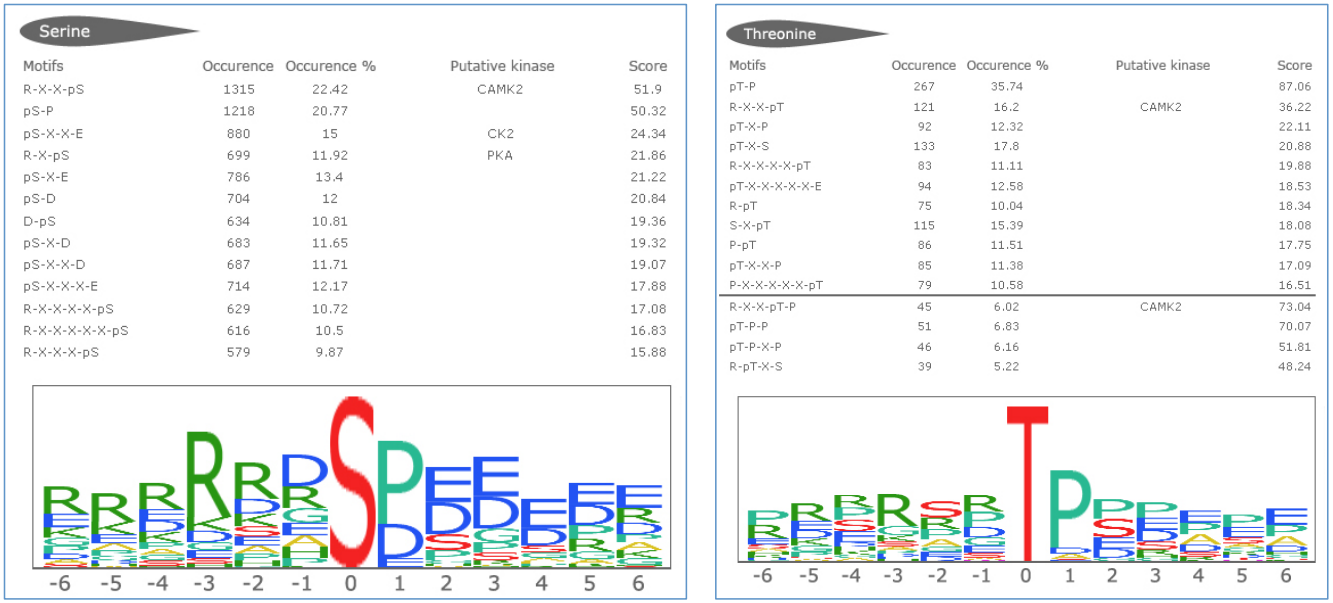
J. Mol. Biol. **48,** 443–453
17. Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16,** 276–277
18. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25,** 25–29
19. Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., and Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* **34,** D354–D357
20. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35,** W182–W185
21. Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res.* **38,** D211–D222
22. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B Met.* **57,** 289–300
23. Letunic, I., and Bork, P. (2007) Interactive Tree of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23,** 127–128
24. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292,** 195–202
25. Sinha, A., Sommer, R. J., and Dieterich, C. (2012) Divergent gene expression in the conserved dauer stage of the nematodes Pristionchus pacificus and Caenorhabditis elegans. *BMC Genomics* **13,** 254
26. Macek, B., Mann, M., and Olsen, J. V. (2009) Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu. Rev. Pharmacol. Toxicol.* **49,** 199–221
27. Olsen, J. V., and Macek, B. (2009) High accuracy mass spectrometry in large-scale analysis of protein phosphorylation. *Methods Mol. Biol.* **492,** 131–142
28. Wang, X., and Sommer, R. J. (2011) Antagonism of LIN-17/Frizzled and LIN-18/Ryk in nematode vulva induction reveals evolutionary alterations in core developmental pathways. *PLoS Biol.* **9,** e1001110
29. Sternberg, P. W. (2005) Vulval development. *WormBook* **2005,** 1–28
30. Sommer, R. J. (2008) Homology and the hierarchy of biological systems. *Bioessays* **30,** 653–658
31. Tian, H., Schlager, B., Xiao, H., and Sommer, R. J. (2008) Wnt signaling induces vulva development in the nematode Pristionchus pacificus. *Curr. Biol.* **18,** 142–146



Supplementary Figure 1. Technical details of the detected phosphorylation events. A) distribution of measured mass deviations of all identified phosphopeptides; B) distribution of localization probabilities of all phosphorylation events; C) distributions of peptide PEP values of all identified phosphopeptides



Supplementary Figure 2. Additional *P. pacificus* protein identifications derived from the phosphoproteome dataset (compared to Borchert et al.).



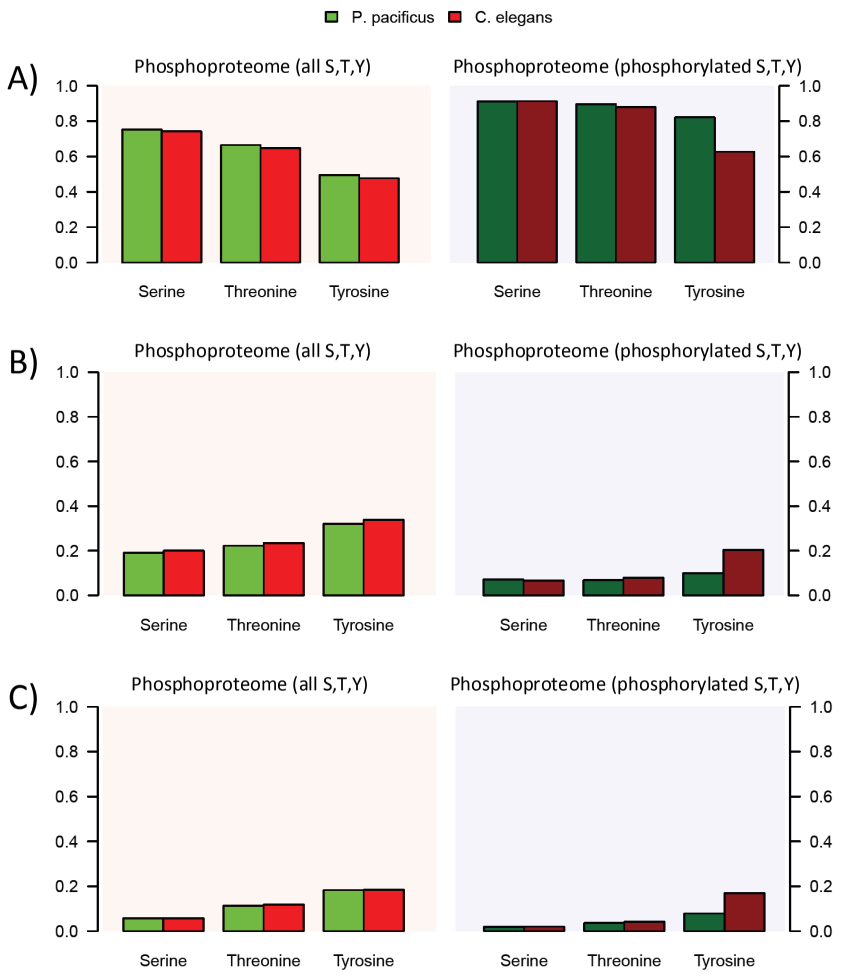
Serine

Motifs	Occurence	Occurence %	Putative kinase	Score
R-X-X-pS	1315	22.42	CAMK2	51.9
pS-P	1218	20.77		50.32
pS-X-X-E	880	15	CK2	24.34
R-X-pS	699	11.92	PKA	21.86
pS-X-E	786	13.4		21.22
pS-D	704	12		20.84
D-pS	634	10.81		19.36
pS-X-D	683	11.65		19.32
pS-X-X-D	687	11.71		19.07
pS-X-X-X-E	714	12.17		17.88
R-X-X-X-X-pS	629	10.72		17.08
R-X-X-X-X-X-pS	616	10.5		16.83
R-X-X-X-pS	579	9.87		15.88

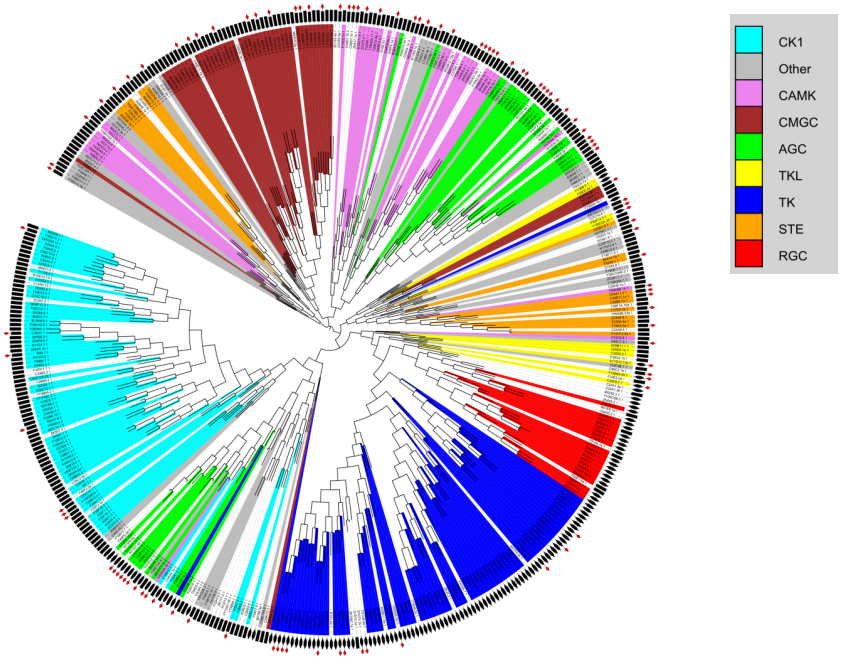
Threonine

Motifs	Occurence	Occurence %	Putative kinase	Score
pT-P	267	35.74		87.06
R-X-X-pT	121	16.2	CAMK2	36.22
pT-X-P	92	12.32		22.11
pT-X-S	133	17.8		20.88
R-X-X-X-X-pT	83	11.11		19.88
pT-X-X-X-X-X-E	94	12.58		18.53
R-pT	75	10.04		18.34
S-X-pT	115	15.39		18.08
P-pT	86	11.51		17.75
pT-X-X-P	85	11.38		17.09
P-X-X-X-X-X-pT	79	10.58		16.51
R-X-X-pT-P	45	6.02	CAMK2	73.04
pT-P-P	51	6.83		70.07
pT-P-X-P	46	6.16		51.81
R-pT-X-S	39	5.22		48.24

Supplementary Figure 3. Overrepresented kinase motifs in *P. pacificus* phosphoproteome.

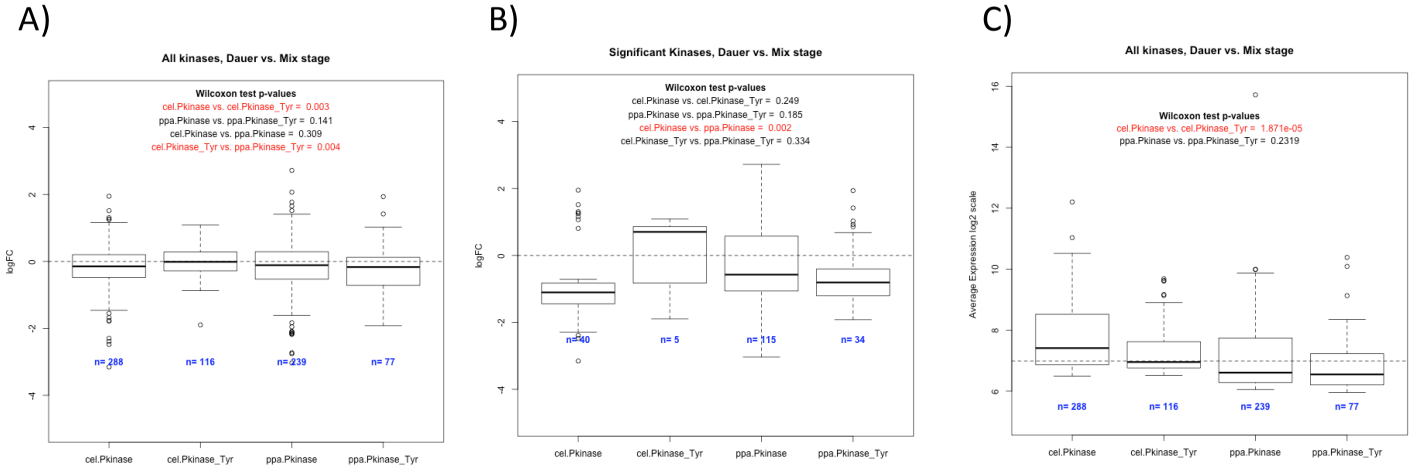


Supplementary Figure 4.



C. elegans

Supplementary Figure 5. Annotated kinome of *C. elegans*. The kinome tree is based on kinase domains predicted using Pfam. The annotation of kinases is based on kinbase (http://kinase.com/).



Supplementary Figure 6. Expression of different protein kinase classes in dauer vs. mixed stage of *C. elegans* (cel) and *P. pacificus* (ppa). Data are derived from Sinha et al. (27). A) changes in expression of all annotated kinases from the dataset; B) Changes in expression of significantly regulated kinases from the dataset; C) Changes in average expression of kinases