The function of mechanosensory systems in the startle behavior of planktonic larvae

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Luis Alberto Bezares Calderón, MSc

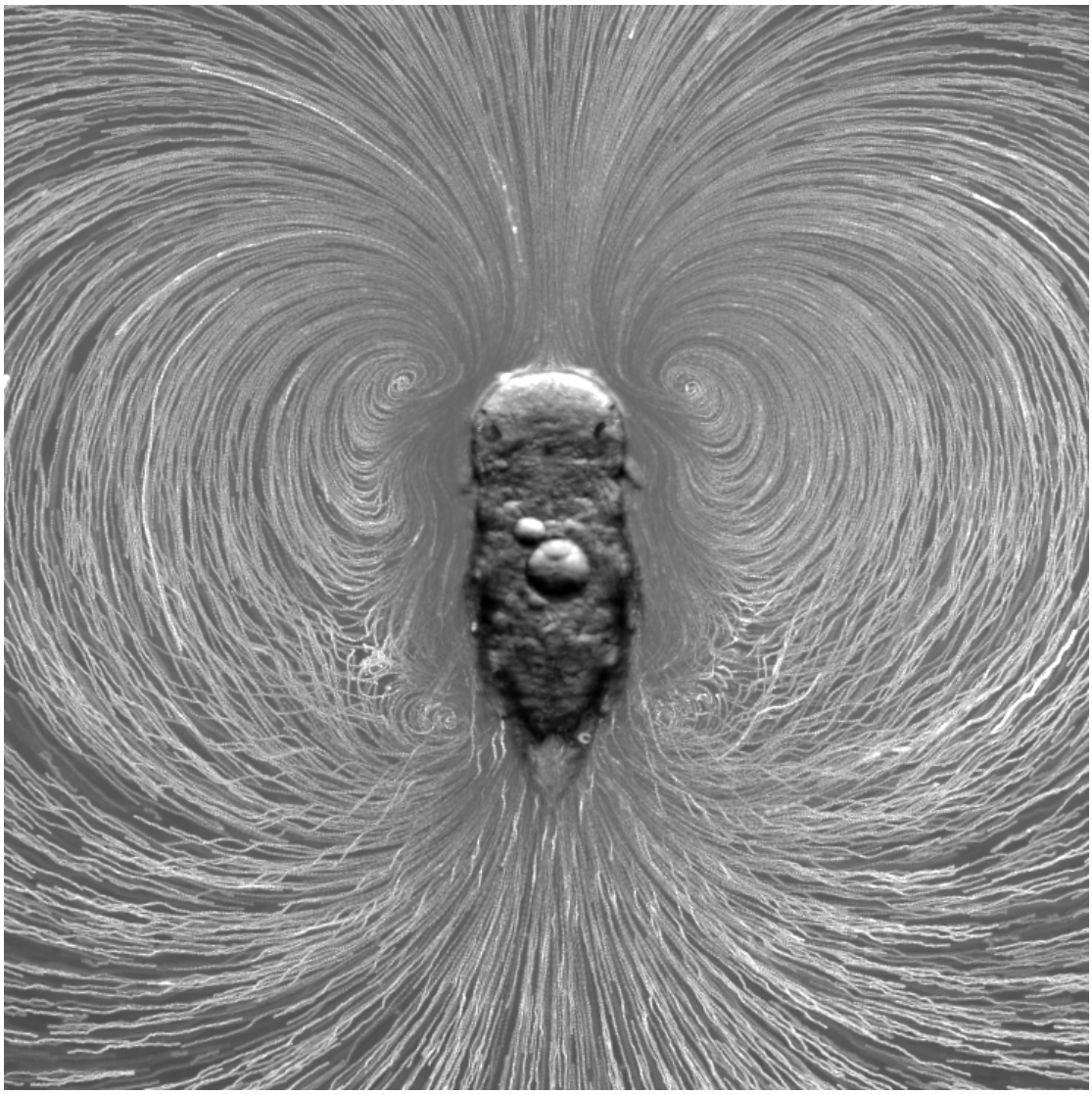
aus Mexiko-Stadt, Mexiko

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	10.05.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr Nico Michiels
2. Berichterstatter:	Prof. Gáspár Jékely



"The scientist's art is first of all to find himself a good master, and then to find himself a good problem"

-M Perutz

"No estudio para saber más, sino para ignorar menos."

-S. Ines de la Cruz

Table of Contents

Abstract

The sensation of mechanical stimuli is a central function of all animal nervous systems. Mechanosensory systems are in charge of this function, using for that a set of specialized molecules and cells that transmit the signal to downstream circuits that initiate and guide a wide variety of behaviors, ranging from navigation, to social interactions. Despite the intensive study of mechanosensory systems in the main genetics models, a clear unified picture of these sensory systems is still lacking. Exploring mechanosensory systems in animals spread across the phylogeny may help reveal such common principles. To contribute towards this aim, the mechanosensory systems of the planktonic larva of the marine annelid *Platynereis dumerilii* are analysed in this work using genetics, circuit and behavioral approaches.

At the behavioral level, a startle response elicited by mechanical stimuli is described in *Platynereis* larvae using high-speed recordings. This startle response is a fast and well-coordinated behavior involving the control of both the muscular and ciliary locomotor systems of the larva. The startle response is shown to be modulated according to the intensity and site of stimulation.

Such responses have been observed in other planktonic organisms, but the mechanosensory cells responsible for initiating the response are not known. A group of penetrating uniciliated neurons in the *Platynereis* larva are shown by calcium imaging to respond to the mechanical stimuli eliciting the startle response. Their morphology is quite similar to putative mechanosensory cells found in other animals, thus suggesting a deep evolutionary conservation.

It is not entirely understood what molecular and cellular mechanisms are required for transforming mechanical cues to cellular signals. Here it is shown that *Platynereis* has homologs to the main molecules that have been implicated in mechanotransduction. The ciliated hydrodynamic receptors identified in this study express *PKD1-1* and *PKD2-1*, two members of the polycystin family that have been implicated in mechanotransduction in other animals. The CRISPR system is used to generate frame-shift mutations in these genes. The mutants no longer display the startle response upon mechanical stimulation, thus suggesting that *PKD2-1* and *PKD1-1* are essential for the transmission of the mechanical information to downstream circuitries.

Startle behaviors generally have a role in avoiding, escaping or deterring predators. It is however not clear what specific adaptations are most useful to increase survival. Here, I used the mutants defective in the startle response to assess the survival value of this behavior. Competition experiments using a rheotactic planktonic predator showed that the mutants are predated more than their wildtype counterparts. These results show that seemingly simple behavioral adaptions can have a high adaptive value.

Due to their relatively simplicity, startle responses such as the one described for *Platynereis* have been dissected at the circuit level. Here, the startle circuit of *Platynereis* larvae is reconstructed at the synapse level using serial transmission electron microscopy. The resulting circuit shows direct and indirect pathways that explain how ciliary bands and muscles are controlled in a coordinated and synchronous manner. A novel group of interneurons and motoneurons is described that provides candidates for further functional exploration of this circuit.

Zusammenfassung

Die Wahrnehmung mechanischer Reize ist eine zentrale Funktion der Nervensysteme aller Tiere. Mechanosensorische Systeme nutzen ein Set von speziellen Molekülen und Zellen, die das Signal des mechanischen Reizes an das postsynaptische neuronale Netzwerk weiterleiten und eine Vielzahl von Verhalten, von Navigation bis zu sozialer Interaktion, induzieren und beeinflussen. Trotz intensiver Forschung der mechanosensorischen Systeme in den wichtigsten genetischen Modelorganismen fehlt ein klares, einheitliches Bild. Dies kann durch die Untersuchung der mechanosensorischen Systeme in Tieren aus weiteren phylogenetischen Gruppen erweitert werden und hilft somit allgemeine Prinzipien hervorzuheben. Um sich diesem Ziel zu nähern, werden in dieser Arbeit die mechanosensorischen Systeme der planktischen Larve des marinen Ringelwurms *Platynereis dumerilii*, auf der Ebene von Genetik, neuronalen Schaltreisen und Verhalten, analysiert.

Auf Verhaltensebene wird eine Schreckreaktion durch mechanische Reize in der *Platynereis* Larve ausgelöst und mit Hilfe von Hochgeschwindigkeitsaufnahmen beschrieben. Diese Schreckreaktion ist ein schnelles und gut koordiniertes Verhalten, welches die Kontrolle von Muskeln und Zilien in der Larve involviert und durch die Reizintensität und dem Ort der Reizapplikation verändert werden kann.

Andere planktische Organismen zeigen ähnliche Reaktion, aber die mechanosensorischen Zellen, die dieses Verhalten verursachen, sind nicht bekannt. In *Platynereis* Larven konnte eine Gruppe von Nervenzellen mit je einer, die Kutikula durchdringenden, Zilie identifiziert werden. Durch Kalziumindikatoren konnte gezeigt werden, dass diese Neurone auf mechanische Reize reagieren, welche die Schreckreaktion auslösen. Ihre Morphologie ist den, in anderen Tieren gefundenen, mutmaßlichen mechanosensorischen Zellen sehr ähnlich, welches eine evolutionäre Konservierung annehmen lässt.

Es ist nicht vollständig verstanden, welche molekularen und zellulären Mechanismen für die Umwandlung mechanischer Reize in zellulare Signale verantwortlich sind. Hier wird gezeigt, dass *Platynereis* Moleküle besitzt, welche homolog zu den Molekülen sind, denen eine Beteiligung an der Weiterleitung des mechanosensorischen Reizes nachgesagt wird. Die bewimperten hydrodynamischen Rezeptoren, die in dieser Studie identifiziert worden, exprimieren *PKD1-1* und *PKD2-1*, zwei Mitglieder der Polycystin Familie, welche in anderen Tieren mutmaßlich an der mechanosensorischen Weiterleitung beteiligt sind. Larven mit einer Rastermutation auf diesen Genen, welche mittels dem CRISPR System erzeugt wurde, zeigen keine Schreckreaktion nach mechanischer Stimulation. Das lässt vermuten, dass *PKD2-1* und *PKD1-1* für die Weiterleitung der mechanischen Reizinformation an das postsynaptische neuronale Netzwerk essentiell sind.

Schreckreaktionsverhalten spielt im allgemeinen eine Rolle in der Vermeidung, Flucht oder Abschreckung von Räubern. Jedoch ist nicht eindeutig geklärt welche spezifischen Adaptionen besonders nützlich für das Überleben sind. In der vorliegenden Studie werden die mutierten Larven ohne Schreckreaktion genutzt, um die Bedeutung des Verhaltens für das Überleben zu bewerten. Die Experimente mit einem rheotaktischen Räuber zeigen, dass mehr mutierte Larven als Wildtyp-Larven gefressen werden. Diese Ergebnisse zeigen, dass scheinbar einfache Verhaltensanpassungen einen großen Effekt haben können.

Hier wird die Schreckreaktion in *Platynereis* Larven auf dem Level von neuronalen Schaltkreisen mit synaptischer Auflösung durch Serien-Transmission-Elektronen-Mikroskopie untersucht. Der daraus resultierende neuronale Schaltkreis zeigt direkte und indirekte Wege, die erklären wie Zilienbänder und Muskeln koordiniert und synchron kontrolliert werden. Außerdem wird eine neue Gruppe von Interneuronen und Motorneuronen beschrieben und liefert Kandidaten für weitere funktionale Erkundigungen des neuronalen Schaltkreises.

Acknowledgements

This work is the result of a challenging journey that lasted a bit longer than 7 years. During this time the intellectual and personal challenges seemed at points too daunting, that I am surprised that the project finally reached a safe haven. But once reflecting on the reasons of this achievement makes it clear that if I managed to complete this project was because of the great intellectual and moral support from my colleagues, friends, and family.

First of all, I totally agree with Max Perutz when he says that a good master is essential to do good science. My supervisor Gaspar Jekely first ignited my interest in the study of mechanosensory systems in the weird animal model that *Platynereis* was at that time. At that point, I could not recognize the potential this project had. Over the years, and with Gaspar constantly renewing my interest and suggesting interesting experiments to go forward, I started to see the relevance this project had. I suspect that Gaspar had the project completely clear right from the beginning, but he was patiently waiting for me to catch up with his insight. I appreciate the freedom he gave me to find my way and his disposition to discuss ideas. I am also grateful for the patience, support and flexibility he showed over the years whenever I wanted to attend conferences, workshops or do internships.

I also would like to acknowledge the specific contributions to this project of all my colleagues. Firstly, the patience required to manually trace all the neurons in the circuit is admirable. The reconstructions made in great part by Csaba Verasztó, with significant input from Sanja Jasek were essential to the completeness of this project. Martin Gühmann, Gáspár Jékely and Reza Shahidi also greatly contributed to this effort. I also thank Reza and Jürgen Berger for his disposition to sit at the electron microscope and take those nice images.

I am also thankful for all the support I received from staff at the Max Planck Institute. In particular, I thank Dorothee Koch and Sinja Mattes, who took excellent care of the worm culture, which was essential to have a reliable source of larvae. I thank Aurora Panzera for all her help during her stay at the Jekely lab, from microinjection to microscopy. I thank her and Christian Liebig for all their advice on microscopy issues, and for providing clever solutions for my crazy imaging experiments. I thank the staff of the Genome Center at the MPI, and in particular to Andrea Belkacemi for the efficiency and speed in which she processed all my samples. I also thank the staff at the mechanical and electrical workshops of the institute including Mark Münster, Klaus Schneider, Maila Götzendörfer, and Luis Antoniotti for their technical support and professionalism.

I also thank colleagues at other institutions, including Detlev Arendt for sharing genome data, Javier Bernardo for sharing tips and tricks on immunochemistry, Nadine Randel for translating the

abstract of this thesis to German (as well as to Katharina Eichler for proof-reading it), and Rodrigo Almeda for his advice on the predator experiments.

I am also indebted to my trainees Sara Mendes and Ada Kozlowska, who contributed to important aspects of this project. They were very patient with my lack of experience to supervise students and showed enthusiasm and support for the project. I also thank them for being excellent students, this made my task as supervisor much more manageable.

During this long PhD I had always the unconditional support, patience and love of Alejandra Manjarrez. I cannot thank her enough for this, and for reminding me to always follow my ideals. I also apologize to her for all the times I was not by her side due to my workaholism. And I thank her for letting me discover Europe by her side. Without those moments of joy away from the lab, I could not have had the patience and courage needed to finish this project.

And although I cannot express my gratitude anymore to her, I am deeply grateful to my mother, María de la Luz Calderón, for giving me the freedom to do what I liked the most. I thank her for showing me to be perseverant, and creative when facing problems. I regret that the wings she gave me to fly took me too far from her when she needed me the most.

And not in the least, I also thank my grandmother for always receiving me back home with all her love, humour and good food. I also thank my sister, my brother and my father, for their love and continued support during the most difficult times.

General Introduction

The evolution of mechanosensation

Mechanosensation: what it is and why it is important

Independently of where a given living organism thrives, be it on earth, air or in water, it will be exposed to mechanical forces. These forces exert tension, contraction or shear stress on a body and need a medium to be propagated.

The nature, and the parameters of the mechanical forces such as frequency, intensity or direction, will be more or less relevant to a given organism depending on its size, and on the complexity and biophysical properties of its receiving sensor. For small cells like the bacterium *Escherichia coli*, one of the most relevant mechanical forces is osmotic pressure (Wood 2015). For bigger unicellular organisms such as the protist *Paramecium*, local deformations in the membrane such as those caused by bumping into an obstacle become relevant (Martinac et al. 1987; Naitoh and Eckert 1969; Jennings 1900). For large protists and multicellular organisms like plants and animals, other mechanical forces such as pressure, gravity, texture or sound are available and provide more detailed information about the surroundings. Internal mechanical forces are also relevant for large organisms, as they give information about the physiological and spatial state of the internal organs.

Mechanosensation is a broad concept simply defined as the ability of an organism to sense and respond to a mechanical force. This sense is thought to have been present since the dawn of life, and in terms of the molecular mechanisms of detection, it represents a sensing mode fundamentally different to that of photosensation and chemosensation (Kung 2005). While the last two senses rely on a lock-and-key type mechanism of activation, mechanosensation requires to detect mechanical forces by their effects on the cell membrane.

The ubiquity of mechanical forces has driven the evolution of multiple molecular, cellular and organismal mechanisms to counteract their damaging effects and to leverage the information they provide to increase fitness. For instance, osmotic stress in bacteria is controlled at the molecular level by the channels MscL and MscS (Martinac et al. 1987; Levina et al. 1999). These channels are opened when the membrane bilayer in which they are embedded is deformed—such as during an osmotic shock—releasing osmolites in a non-specific manner and thus restoring the turgidity of the cell (Booth et al. 2007). In yeast, a member of the TRP family expressed in the vacuole has an analogous function (Palmer et al. 2001; Denis and Cyert 2002). Cells have evolved in addition to stretch-activated channels additional means to sense mechanical forces. Focal adhesions in eukaryotes are points of attachment to the substrate that concentrate mechanical stress (Sarkar 1999). Through interactions with the extracellular matrix, focal adhesions can detect the mechanical properties of a substrate, and by their association with the intracellular scaffolding machinery, drive local and global changes in the cell (Chen 2008). Cilia in eukaryotic cells also serve as

mechanosensory organelles sensing shear and compression forces in the surroundings that drive changes in cell physiology (Spasic and Jacobs 2017). At the organismal level, plants evolved a way to sense gravity by using sedimentable amyloplasts—a sort of statolith—in specialized cells in the root and shoot (Anon 1900b; Anon 1900a). Amyloplast sedimentation in turn activates signaling cascades—possibly through the opening of as-yet-unidentified mechanosensitive channels—that reach the corresponding elongation zones to direct growth in a particular direction (Baldwin et al. 2013; Su et al. 2017).

Mechanosensation in the nervous system

One of the major transitions in evolution was the origin of animals from unicellular protists (Smith and Szathmáry 1998). As such, this transition brought with it new sensory abilities and new sources of mechanical information that were exploited. Novel and inherited functions from their unicellular ancestors were segregated into cell types, of which the detection of force was no exception. In fact, this modality is hypothesized to have been present in one of the first cell types to have appeared in animals (Arendt et al. 2015).In this scenario, these cells were specialized in sensing and at the same time responding to mechanical cues, a sort of mechanosensory-motor cell type that directly preceded the first neurons (Mackie 1970).

The evolution of the nervous system gave animals more elaborate ways to react and control their behavior upon detecting stimuli. Neuron-based mechanosensory systems diversified into the different types we observe nowadays, including the complex hearing organs in mammals, and in insects, the touch detectors in the skin, or in the *C.elegans* nose, or the proprioceptive systems that control locomotion. But because of the different manifestations of mechanical forces, and the diversity in the mechanosensory structures, it has been challenging to understand and to unify this sense under an evolutionary framework. Even a single animal has a wide variety of mechanoreceptors with partially overlapping molecular machineries. For instance, of the 302 neurons in *C.elegans,* at least 30 are mechanosensory cells, specializing to detect gentle touch, harsh touch, body stretch and substrate texture (Goodman 2006). Moreover, as it will be introduced in Chapter 3 there is not a single transducing molecule identified that is tightly associated to the mechanosensory system. This sharply contrasts with the case observed in the photosensory system, in which opsins are responsible for phototransduction in almost all animal photoreceptor cells (Fain et al. 2010). Nor there is a conserved set of regulatory genes—such as Pax/Six/Eya/Otx/Mitf for the photoreceptor cells (Vopalensky and Kozmik 2009)—that is essential for the development of mechanosensory cells (Fritzsch et al. 2007).

Such conclusion has been drawn from work in the main genetic systems—the fly, the nematode worm and the mouse—which only represent a small fraction of all the animal diversity. While new molecular profiling technologies applied to many animal phyla may help untangle the evolutionary

relationships between neuronal cell types, including mechanosensory cells (Arendt et al. 2016), the need for functional characterization of mechanosensory systems in those animal groups is imperative to properly assign function to cell types, and to further explore the diversity in mechanosensory systems.

Neuroethology of planktonic organisms

Most of the animal phyla as well as their unicellular relatives are of aquatic origin[1,2]. Many, if not all of them, form part at some point in their life cycles of the morphologically and taxonomically diverse group called zooplankton (Kiørboe 2011) (**Figure 1-1**). The neuronal basis of behavior in planktonic organisms is relatively unexplored, even when it has been long recognized that in such a diverse community probably lies a rich and interesting treasure of neurobiological and behavioral mechanisms (Bullock 1997).

The sensory ecology of zooplankton

To understand how animal sensory systems as a whole work, it is essential to place them in their natural and physiological context: the kind of information that is available to the organism based on the organism's physical features (size, shape, etc), the sophistication of its sensory systems, and the nature of the physical forces around it (Uexküll and Kriszat 1934).

The planktonic environment presents unique sensory challenges, which are quite different to the environment in which animal sensory systems have been traditionally studied. First of all, in contrast to land, this environment is almost featureless in the horizontal direction (Ringelberg 2010), at least at first glance, so orientation is non-trivial. In addition, this environment is sparsely populated. For feeding zooplankton, the area that needs to be scanned in search for particulate food is in the orders of magnitude bigger than its size (Kiørboe 2011). Additional consideration has to be given to the fact that many planktonic organisms are small, thus limiting the complexity and type of sensory structures they can have, as well as the stimuli they can perceive (Martens et al. 2015). At small scales viscous forces dominate over inertial forces (Lauga and Powers 2009), which has implications for their means of locomotion (Chia et al. 1984; Purcell 1976).

Despite their small size and the locomotion challenges brought with it, planktonic organisms are able to undergo daily migrations up and down the water column, in a behavior called diel vertical migration (DVM) (Brierley 2014). DVM occurs in both oceans and lakes and is thought to be an adaptation to minimize encounters with predators (Bollens and Frost 1989; Zaret and Suffern 1976). These rhythmic migrations are driven by both external factors—of which light is considered a main driver (Ringelberg 2010)—and by internal clocks (Häfker et al. 2017; L. Zhang et al. 2013; Tosches et al. 2014), and may be modulated by neuroendocrine signaling (Conzelmann et al. 2011).

---

[1] https://simple.wikipedia.org/wiki/List_of_animal_phyla

[2] Census of Marine Zooplankton. http://www.cmarz.org/index.html

Detailed observations and experiments in more controlled environments have revealed other relevant planktonic behaviors such as feeding, predator avoidance, mating, dispersal or settlement, as well as the underlying sensory systems involved. For instance, experiments with a range of chemical cues excreted by potential predators suggest planktonic organisms may use chemoreception to trigger vertical migrations (Dodson 1988). Chemoreception also plays a role in filter-feeding copepods that select the food particles based on chemical signals (Poulet and Marsot 1978).Their swimming activity is also increased upon the presence of food odors (Buskey 1984). Detailed analysis of the feeding behavior in the pilidium larva of Nemerteans have revealed a complex sequence of behaviors involving muscle contraction and ciliary control that lead to prey capture (von Dassow et al. 2013).

Mechanical forces in the ocean

Mechanical forces in the ocean can be roughly classified into hydrodynamic and hydrostatic forces, although the distinction is not clear-cut. The mechanisms by which small zooplankton detect them are barely explored, but a range of ecologically relevant behaviors driven by mechanical forces have been documented.

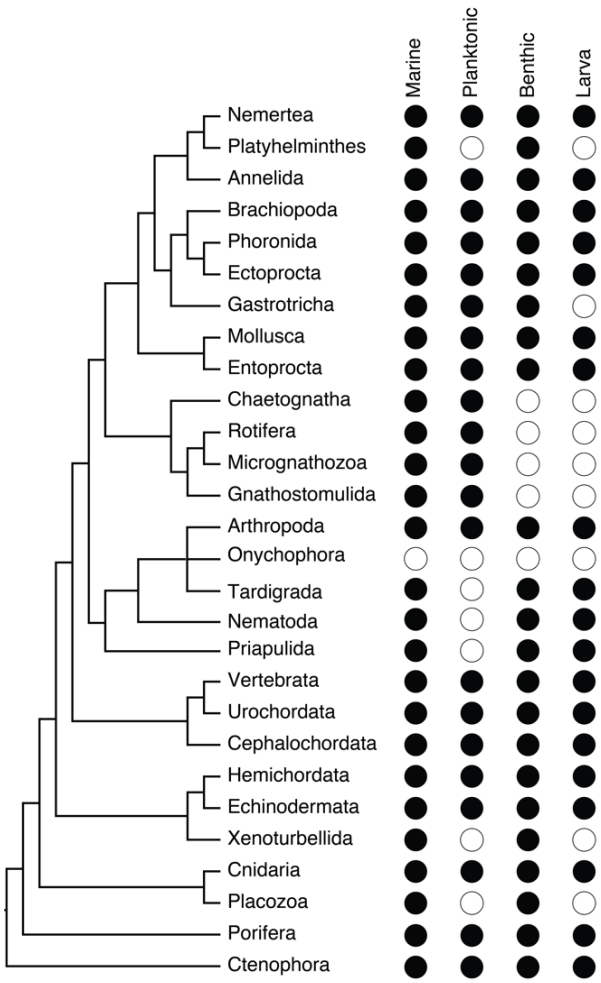


Figure 1-1 Animal phylogeny. Abridged version of the animal phylogeny based on (Laumer et al. 2018; Marlétaz et al. 2019). Filled circles indicate the presence in a given phylum of marine species with either planktonic or benthic life style, or both.

Detection of hydrodynamic signals has a big role in planktonic prey-predator interactions. Chaetognaths sense the near-field vibrations produced by their prey to successfully capture it (Newbury 1972). Predatory copepods are also thought to use mainly near-field mechanoreception to ambush prey (Légier-Visser et al. 1986). Ctenophores localize their copepod prey by the small oscillations they create during swimming (Costello et al. 1999). Additional behaviors related to predator detection will be mentioned at greater length in Chapter 1.

Near-field hydromechanical signals have also a known role in mate detection. Upon chemical detection of males, female copepods jump to create a strong hydrodynamical wake that can be perceived by potential a potential partner (Duren et al. 1998). Male copepods require these hydromechanical signals to precisely locate and capture the female (Seuront 2013; Dur et al. 2011).

Turbulence is another relevant hydromechanical force in the ocean. It can both enhance planktonic encounter rates (Rothschild and Osborn 1988; Evans 1989), but it also interferes with the detection of other hydromechanical signals (Kiørboe and Saiz 1995). Turbulence can be very variable in space and time. Some copepods adapt their swimming behavior to the levels of turbulence they perceive, so they can control and direct their diffusivity properties even in turbulent environments (Michalec et al. 2017). The feeding efficiency of some copepod species is affected by the levels of turbulence (Saiz et al. 2003), thus causing vertical migrations as a function of turbulence levels. Oyster larvae use their statocysts to sense turbulence and adjust their posture, or settle if the turbulence levels are too high (Fuchs et al. 2015). Turbulence is also a cue influencing settlement of sea urchin larvae (Gaylord et al. 2013).

Hydrostatic pressure is another factor potentially affecting the mechanosensory systems of planktonic animals. At steady state, hydrostatic pressure is a scalar measure that increases linearly with depth (Blaxter 1978). Many invertebrate larvae across the animal phylogeny show a locomotor response to changes in pressure (Knight-Jones and Qasim 1955; Rice 1964; Hardy and Bainbridge 1951). This response usually consists in upward or downward locomotion upon pressure increase or decrease, respectively. The sensitivity threshold to changes in pressure can be as low as 10 mbar (equivalent to 10cm change in depth)(Enright 1961). Different sensory mechanisms and structures have been proposed as pressure sensors in planktonic animals (Digby 1961), but no definite answer has been yet reported. The ecological relevance of this behavior has not been fully explored, but it may help planktonic animals retain a certain depth in downwelling and upwelling currents near the shore (Genin et al. 2005). It may also play a role in entraining tidal rhythms (Enright 1965).

Platynereis as model system for neurobiology of sensory systems

Several studies indicate mechanosensory systems drive behaviors relevant for planktonic biology, yet little is known about the underlying sensory mechanisms and neuronal circuitries. Comparing

them to more extensively studied animal groups can help us better understand the function and evolution of the nervous system in general and of the mechanosensory systems in particular.

The planktonic larva of the marine annelid *Platynereis dumerilii* (hereafter *Platynereis*) is an attractive model for answering questions on the neuroethology of zooplankton. The larva of this polychaete worm swims with bands of beating cilia that are controlled by a system of motoneurons expressing acetylcholine and serotonin (Tosches et al. 2014; Verasztó et al. 2017; Starunov et al. 2017). The 2-day old (trochophore stage) and the 3 day-old (nectochaete stage) larva have a main ciliary band called prototroch, and a posterior band called telotroch (**Figure 1-2**A). The region anterior to the prototroch is conventionally called episphere, and that posterior to the telotroch is called the pygidium. The nectochaete larva has additional dorsal and ventral bands in each of the three segments called paratrochs (**Figure 1-2**A) (in this study paratrochs and telotroch will be called bodytrochs). At this stage, the larva has a well-developed musculature that it uses to crawl by moving the extremities called parapodia, each endowed with a set of spiny bristles (technically called chaetae). Muscles are also used during swimming to navigate in the water, for example during phototaxis (Randel et al. 2014). Thus, this larva stage shares the two main modes of locomotion with many other planktonic marine invertebrate larvae (Chia et al. 1984).

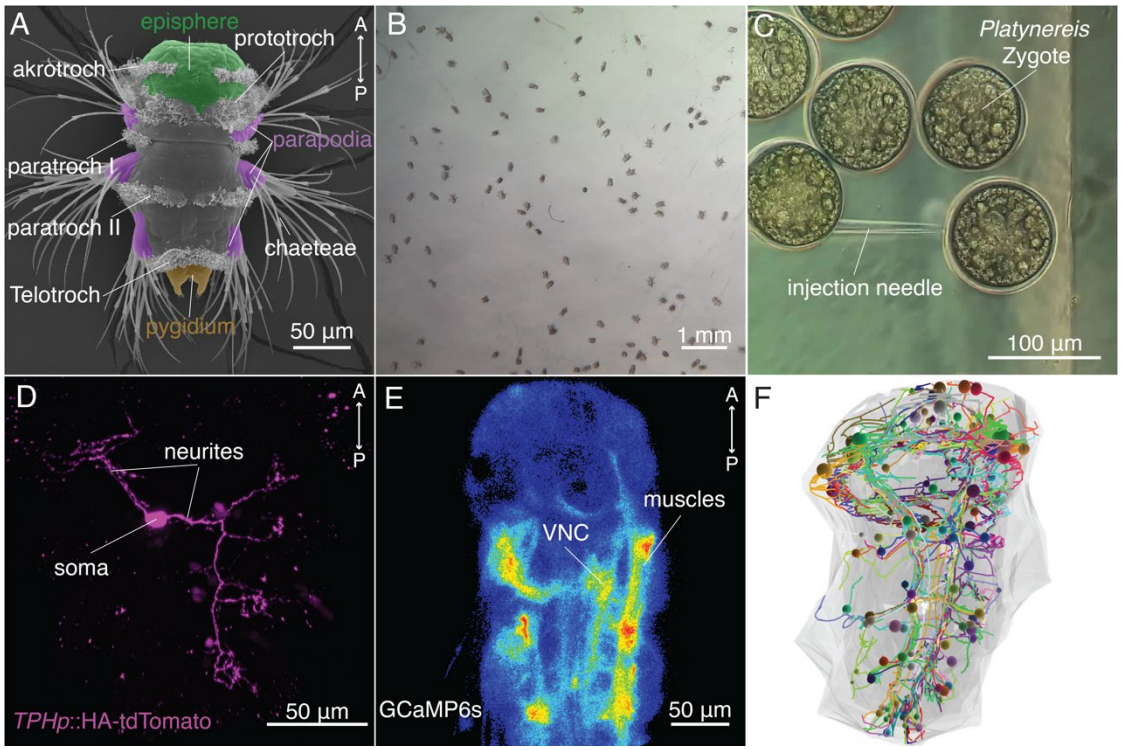


Figure 1-2 The planktonic *Platynereis* larva as a genetically tractable model for studying the neuroethology of zooplankton. (**A**) SEM micrograph of a nectochaete *Platynereis* larva. Different structures relevant to this study are labeled and highlighted in different colors. (**B**) A batch of nectochaete larvae obtained from the in-house culture freely behaving in a culture dish. (**C**) Snapshot of *Platynereis* zygotes during an injection session. Many embryos can be injected in series during a single session. (**D**) The full morphology of a serotonergic neuron can be visualized in the nectochaete larva when injected with the *TPH* promoter construct. (**E**) Neuronal and muscle activity can be imaged in the living larva using genetically encoded indicators such as GCaMP6s. (**F**)Electron microscopy volume reconstruction of neuronal morphologies. Credit SEM micrograph shown in A: Jürgen Berger. VNC: ventral nerve chord. Reconstruction in E taken from the Jékely Lab connectome project.

Detailed molecular, anatomical and functional studies can be performed in *Platynereis* due to key technical features. First, this animal can be easily cultured in the lab and newborn larvae can be obtained in large numbers every day (Hauenschild and Fischer 1969), allowing to perform behavioral experiments and pharmacology treatments with large numbers of larvae (Conzelmann et al. 2011; Gühmann et al. 2015) (**Figure 1-2**B) . The fertilized eggs can be microinjected in large numbers (**Figure 1-2**C), thus enabling to target specific population of cells using transient and stable transgenics (e.g.(Verasztó et al. 2017; Žídek et al. 2018; Veedin-Rajan et al. 2013; Backfisch et al. 2013))(**Figure 1-2**D). This technique also allows to track cell lineages using cell cycle reporters (Özpolat et al. 2017), or to perform calcium imaging using genetically encoded indicators (Randel et al. 2014; Tosches et al. 2014; Verasztó et al. 2017) (**Figure 1-2**E). Microinjection in *Platynereis* also allows to knockdown and knockout genes to analyze their role in behavior and physiology (Conzelmann et al. 2011; Gühmann et al. 2015; Verasztó et al. 2018). Due to its stereotyped development, the molecular fingerprint of single cell types can be obtained at cellular resolution by gene expression atlases and single-cell transcriptome sequencing (Vergara et al. 2017; Achim et al. 2018; Achim et al. 2015; Asadulina et al. 2012). Finally, and also due to the stereotyped morphology and the small size of the larva, synapse-level neuronal circuits (connectomes) can be reconstructed in electron microscopy volumes (Randel et al. 2015; Randel et al. 2014; Verasztó et al. 2017; Shahidi et al. 2015)(**Figure 1-2**F).

Certain biological features in this animal allow for phylum-wide comparisons. The genome of *Platynereis* shows a high conservation of gene structure and gene content when compared to those in vertebrates and other bilaterians (Raible et al. 2005). Many conserved neurotransmitter and neuropeptide signaling systems are present in this animal (Bauknecht and Jékely 2017; Conzelmann, Williams, Krug, et al. 2013; Jékely 2013). A highly conserved molecular toolkit allows to assess the homology of major anatomical structures such as the pallium, the neurosecretory brain centers, or the nerve chord in bilaterian animals (Tomer et al. 2010; Denes et al. 2007; Arendt et al. 2004; Christodoulou et al. 2010; Tessmar-Raible et al. 2007). Finally, behavioral features common to many animals, can be studied in this annelid, such as neuronal control of cilia (Tosches et al. 2014; Verasztó et al. 2017), life-cycle transitions (Conzelmann, Williams, Tunaru, et al. 2013; Conzelmann et al. 2011), or sensory driven behaviors such as phototaxis (Jékely et al. 2008; Randel et al. 2014; Gühmann et al. 2015), light avoidance responses (Ayers et al. 2018; Verasztó et al. 2018), or chemosensory behaviors (Chartier et al. 2018).

Aims of this thesis

The general aim of this study was to characterize the mechanosensory systems of the planktonic larva of *Platynereis dumerilii* at the molecular, circuit and behavioral levels. In particular, I focused on characterizing a seemingly simple response to hydrodynamic stimuli in the nectochaete larva, a startle response that was simple, reproducible and robust. As it has been proven in other organisms, startle responses provide a good model for integral analysis of behavior. Thus, this response was a good opportunity to understand more about the behavior of this animal and about its mechanosensory systems, while in the process develop this system further as a model to answer questions on sensory neurobiology.

The first question regarding the startle response was to know the nature of the mechanical stimulus triggering it. In addition to that, a detailed description of the response was needed in order to understand its neural implementation. The study of startle responses in other animals have shown that seemingly simple startle responses actually have rather complex underlying neuronal circuitry. Conversely, the quantification and recording of the different behavioral features of the startle response was thus required to assess the complexity of the behavior.

A second aim was to identify and characterize the responsible mechanosensory cells triggering the response. In contrast to photosensory cells, which have clearly recognizable sensory structures, and which almost exclusively express marker molecules such as opsins, mechanosensory cells do not have a defining morphology, and few specific markers are known, thus making their identification far from trivial. This challenge, together with the lack of a cellular map of sensory cells in *Platynereis* required the localization and description of all putative mechanosensory cells in the animal based on anatomical information.

In parallel, and towards the general aim of characterizing the mechanosensory systems in the larva, a list of conserved mechanotransduction channels was compiled. The conservation of mechanotransduction channels points to an underlying unity in animal mechanosensation, which is far from understood. The expression of a subset of these molecules was mapped to the putative mechanosensory cells identified by anatomical means. This molecule-anatomical map was carried out with the specific aim of enabling functional access to subsets of sensory neurons. A greater aim to complete this map was to eventually enable comparisons to other mechanosensory systems across animals.

Full characterization of a sensory cell requires physiological evidence. Thus, a further aim towards the study of the startle response was to test the sensitivity of the putative mechanosensory neurons to mechanical stimuli that normally trigger the response. Obtaining definite evidence of

mechanosensory cells in *Platynereis* would be not only an entry point to understand the neuronal response they trigger, but also serve as comparison to mechanosensory cell types in other animals. In genetic model organisms, the use of mutants has contributed to the understanding of the cellular mechanisms behind mechanosensation. As genetic analysis has proven useful in understanding *Platynereis* behavior, this approach was used to learn more about the mechanosensory system behind the startle response. In this process, the effectiveness and utility of the CRISPR genome-editing technology in *Platynereis* was assessed. The final aim was to use molecular information as a link to facilitate comparisons with the genetic systems traditionally studied.

The high temporal and spatial coordination of the startle response and its short latency hinted at a naturally evolved behavior. Startle responses in other animals are usually adapted to respond and deter predators. With the aim of getting closer to the neuroethological study of behavior in *Platynereis*, that is, to a study with an emphasis on behavior, and balance out a purely molecular-neurophysiological approach, the startle response in *Platynereis* was put to test in the context of prey-predator interactions. Neuroethological studies complement studies in more reduced preparations and reveal the relevant behavioral sequences actually implemented in nature.

A final aim of this study was to reconstruct with a connectomics approach the potential neuronal circuitry implementing the startle response. Circuit-level descriptions of other behaviors in *Platynereis* have identified motifs that can be mapped to particular behavioral features. Similarly, reconstruction of startle circuits in other invertebrates have pointed to the key players implementing the main features of the response, but at the same time uncover the nuances and hidden complexities in these behaviors. The cell-level reconstruction of the startle circuit thus promised to open a whole set of hypotheses that could be functionally explored in future studies.

Chapter 1 Kinematics of the startle response in nectochaete larvae

"It is appropriate to define the goal that an organism has to reach before examining the mechanism (and the hardware) in which it is embodied."

-David Marr, Vision 1982

Statement of contributions and publication status

The experiments shown in this chapter as well as the graphs and figures were entirely performed and assembled by the author of this thesis. The results shown in this chapter were published elsewhere in abbreviated form (Bezares-Calderón et al. 2018).

Introduction

Startle responses in animals

A well-thought definition of a startle response can be found in (Bullock 1984). The author defines this type of behaviors as an abrupt response to an unexpected stimulus, which could potentially be a threat to the organism. He emphasizes that this response does not need to be objectively fast, but only fast enough given the natural context of the organism and the nature of the stimulus. Another important feature in this definition is the abruptness of the response. This not only depends on the intensity of the stimulus, but also on the state of attention of the startled organism. The internal nature of this state means it is not always straightforward to determine what constitutes an unexpected stimulus.

Startle responses are widespread behaviors in animals and usually depicted as well-coordinated stereotypic reflexes. These responses usually are the initial set of actions of a more complex and variable escape or deterrence behavior, although the two terms are sometimes used interchangeably. Startle responses and the underlying nerve impulses have been described in a variety of organisms, including annelids (T H Bullock 1945; Nicol 1948; Bullock 1984). A few well characterized examples are presented to illustrate the key features of such responses and the main parameters that can be studied.

One of the best-known examples is the C-start escape response in teleost fish. In this behavior, the animal detects a visual, acoustic or vibrational stimulus and swims away from the potentially threatening stimulus by making a sharp turn (hence the name of a C-start escape) opposite to the side of stimulation (Eaton et al. 1977; Weihs 1973). The initial reaction to the stimulus is of extremely short latency, between 5-10 ms, and the unilateral muscle contraction opposite to the side of stimulation is completed within 20ms. The C-turn is a more stereotyped response than the ensuing set of movements that accelerate the fish away from the threat (Eaton et al. 1977). In fact, variants of the second phase of the response have been proposed (Domenici and Blake 1991).

Another well-studied startle response is the tail-flip escape response in crayfish. Upon detection of a mechanical stimulus from the back, crayfish jump up and forward (a "jack-knife motion) within 30ms (Edwards 2017). Anterior stimulation (either of visual or mechanical nature) triggers a backward jump, with a latency of 25ms (Edwards 2017; Wine and Krasne 1972). In both cases, a fast flexion of the abdominal muscles is required to propel the animal away from the stimulus. This initial response is fairly stereotyped, but like in fish, it is followed by a more variable swimming phase that completes the escape maneuvers (Wine and Krasne 1972).

Analysis of startle responses with high temporal resolution show additional complexities that can be addressed at the neural level. In teleost fish, kinematic analysis of the C-start escape response

have allowed it to divide it into different stages (Weihs 1973). Detailed analysis at high temporal resolution helped reveal another type of escape response called the S-start response triggered by posterior stimulation (Spierts and Leeuwen 1999; Webb 1976; Hale 2002; Liu et al. 2012). Both responses have clearly defined motor patterns that can be mapped to different circuitry dynamics (see (Liu and Hale 2017)for a recent study). In crayfish, a third response of longer latency was identified that involves abdominal extension followed by swimming (Edwards 2017; Wine and Krasne 1972). In each of the three different startle responses the appendages are arranged in different behavioral sequences (Cooke and Macmillan 1985). In the fruit fly *Drosophila*, a detailed high-speed analysis of the escape response upon visual looming stimulation revealed a much more complex set of motor patterns that could not be explained by a simple reflex pathway (Card and Dickinson 2008).

Even though startle responses show a stereotyped component for decreasing reaction times, it is also evident from the studies mentioned above that the exact response also depends on the site of stimulation. Further examples in other animals add to this general view. For instance, in Nereid polychaetes a startle response involving forward pointing of parapodia is observed upon anterior stimulation, while backward pointing of parapodia is observed upon posterior stimulation (Horridge 1959). Similarly, in the nematode *Caenorhabditis elegans* touching the anterior part of the body stops head oscillations and causes a rapid reversal, while touching the tail of the animal causes it to speed up (Chalfie et al. 1985; Alkema et al. 2005).

The nature and intensity of stimulus also determine the type of response observed. In leech, gentle anterior stimulation elicits a localized withdrawal response involving only a few segments, while stronger anterior stimulation induces a symmetric shortening of the whole body (Shaw and Kristan 1995; Kristan et al. 1982). In *C.elegans*, increasing the harshness of the mechanical stimulus triggers a more robust response to that seen with gentle stimulation, as well as novel behavioral components (Li et al. 2011). In *Drosophila* larvae, sound frequencies emulating the buzzing of a wasp trigger a startle response (W. Zhang et al. 2013). Mechanical harsh stimuli also trigger a rolling behavior in this larva that involves a an asymmetric but coordinated contraction of muscles towards the site of stimulation (Hwang et al. 2007). In freely moving crayfish, how "abrupt" the stimulus is determines the type of response observed (Wine and Krasne 1972). In fish and in *Drosophila* adults, fast approach rates of looming stimuli trigger more kinematically stereotyped escape responses (or even a freezing behavior for very fast rates in the case of fish), while at slow approach rates the responses are more variable and slower (Bhattacharyya et al. 2017; von Reyn et al. 2014).

In conclusion, these studies reveal that high temporal synchrony and spatial coordination is a feature of many startle responses, and that the site, intensity and nature of the stimulus shape the type of the startle response.

Startle responses in zooplankton

Although startle behaviors have been studied in numerous groups across the animal phylogeny, for practical reasons this research has concentrated on macroscopic organisms. The study of startle responses in the smaller planktonic animals is lagging behind, although a number of studies in unicellular organisms as well as in animals, mostly on copepods, have explored startle responses in this community.

Now classic studies in the unicellular organism *Paramecium* have shown that when this ciliate encounters an obstacle front-end during swimming, it reverses the direction of ciliary beating to swim backward for a short period, to then swim forward again in a direction at an angle from the original course (Jennings 1900; Jennings 1899). If *Paramecium* is touched on the posterior end, the ciliary beating frequency increases, leading to increased forward swimming (Naitoh and Eckert 1973). Other protists and ciliates show rather a jumping behavior when exposed to a siphon flow (Jakobsen 2001), escaping at non-random angles, and always away from the siphon tip. The stimulus is proposed to be a deformation or a change in speed in the fluid flow. The ciliate *Halteria* also shows a backward jump behavior when contacting an obstacle (Tamar 1979; Tamar 1965). This jump consist a fast contraction of the ciliary mantle and ciliary bristles (Tamar 1974). Contractile appendages found in some ciliates serve as propulsive structures that move them away from a mechanical stimulus with speeds even higher than those shown by the fastest copepods (Gemmell et al. 2015).

The best characterized startle response for planktonic animals can be found in the most abundant group of holozooplankton, the Copepods. This animals display rapid escape responses, achieving 200 to 500 body lengths per second (Strickler 1975). Kinematic studies on tethered copepods, have shown that near-field water disturbances trigger a stereotyped response, involving a fast and highly coordinated sequence of movements of the appendages that was completed with 10ms from stimulus start (Strickler 1975; Lenz and Hartline 1999; Lenz et al. 2004). Experiments in freely swimming specimens have confirmed the locomotor sequences, added new elements to the escape sequence, and revealed in some species alternative responses to similar hydrodynamic stimuli, such as a freeze responses, or a freeze and escape response (Bradley et al. 2013; Buskey et al. 2002).

Startle responses in other planktonic organisms with other forms of locomotion have also been reported. High-speed recordings in free and semi-tethered preparations revealed the kinematics of the escape response in the small rotifer *Polyathra* (Gilbert 1985), an animal that moves by using a set of muscle-controlled paddles on each side of the body. The escape response consists in a brief but sharp increase in swimming speed, basically a jump, that moves the animal 15 body lengths away from its original position. The thrust for this jump is provided by the asynchronously upward movement of all the paddles followed by an equally asynchronous downward movement, all

occurring in less than 30ms. The response is triggered by contact with another specimen (Gilbert 1985), or by hydrodynamic stimuli (Kirk and Gilbert 1988; Gilbert 1987) with a lag time of only 7ms (Gilbert 1985). The medusae of the Hydrozoan jellyfish *Aglantha digitale* shows an escape response to touch stimuli that involves a vigorous contraction of the vellum that displaces the animal five body lengths (Donaldson et al. 1980).

In ciliated larvae, another major component of the zooplankton, some reports exist about startle responses, although not quantified with as much detail as in the previous examples mentioned above. When exposed to low to moderate siphon currents the ciliated larva of spionid polychaete *Polydora ciliata* swims faster, but at higher current velocities it stops swimming and curls up, thereby exposing all its spines (Singarajah 1969). This behavior has been observed elsewhere, but not quantified (Hansen et al. 2010). A similar chaetal extension behavior was observed in a different sabellid larvae, *Sabellaria cementarium* (Pennington and Chia 1984), and in the larva of the brachiopod *Terebratalia transversa* (Thiel et al. 2017). A sinking response in the latter organism was also reported. Finally, in the veliger larva of the gastropod *Mangelia* and in the polytroch larva of the sea snail *Pneumoderma*, both ciliated larvae, mechanical stimulation induced synchronized arrest of ciliary beating and muscle contraction (Mackie et al. 1976). Weak stimuli caused only ciliary arrests while stronger stimulation also induced whole-body contraction.

This by no means exhaustive recount of examples of startle responses, show that these behaviors occur even in the smallest organisms, and although seemingly simple at first glance, startle responses can reveal hidden complexities when analyzed at a finer temporal resolution. In this chapter, the startle response in the *Platynereis dumerilii* nectochaete larva will be presented and show to be triggered by mechanical stimuli. The kinematics of the behavior in tethered animals will be described, dissecting its different elements that show this is a fast and whole-body coordinated response tuned to stimulus intensity and location.

Materials and Methods

Startle assay in freely swimming animals

Experimental setup

4 ml of nectochaete larvae were collected by phototaxis and transferred to a Ø5 cm glass-bottom dish (GWSB-5040, Willco Wells) with a vibrating shaft-less motor (EXP-R25-390, Pololu) glued to its bottom (**Figure 0-2**B). The motor was connected to a circuit and activated for 100 ms via an Arduino microcontroller (Arduino UNO R3, Arduino) with a custom-written script (see `Switch_motor.ino` Appendix for details). The dish was placed in the AxioZoom V.16 (Carl Zeiss GmbH) and videos recorded at 15 frames per second (fps) with a digital CMOS camera ORCA®-Flash-4.0 V2 (Hamamatsu). Experiments were performed in a dark room, and a long-pass filter was placed between the light source and the dish to minimize phototaxis. The behavioral setup was not touched during recording to avoid introducing unwanted vibrations.

Measurement of larvae speed and area

Videos were manually inspected to find larvae that got startled with the stimulus. A custom macro was written in Fiji (Schindelin et al. 2012) to extract the speed and area of the larvae before and after the stimulus (see `FigS1_StartleFreelySwimming.ijm` Appendix for details). Stimulus onset was set as the first frame where any ripple or deformation in the water surface was observed. It was defined for every video before processing it. The measurements obtained from FiJi were used as input to an R script written to calculate, normalize and plot the speed and area (see `FigS1_StartledataFreelySwimming.R`, see Appendix for details).

Kinematics of startle behavior

Experimental setup

Phototactic larvae were relaxed in 50-100 mM $MgCl_2$ 10 min before tethering them to a Ø3.5 cm glass-bottom dish (HBST-3522, Willco Wells) with a non-toxic glue originally developed for *C.elegans* (Wormglu, GluStitch Inc). The dish was filled with 2.5 ml natural sea water (NSW) without $MgCl_2$. To tether the larvae, small drops of glue were smeared against the glass bottom using a 10 µl pipette. Care was taken not to push glue out of the pipette tip before contacting the glass, as the glue solidifies in contact with water. Single paralyzed larvae were gently pushed with a coarse tungsten needle until falling on top of a drop. Both dorsally and ventrally tethered larvae were analysed as there was no discernible difference in the response profiles. Care was taken to minimize or to avoid gluing ciliary bands, sensory cilia, parapodia, head and pygidium.

Once tethered and letting some time for recovery, larvae were assessed for the startle response with a gentle vibration to verify that relevant structures were unhindered, and the animal was healthy and in a swimming mode. 1 µm multi-fluorescent beads (24062-5, Polysciences) were

diluted in 5% BSA to 1:100, sonicated for 1 min and added to the glued larva preparation at a 1:10 dilution. The BSA prevented the beads to bind to the larva, while sonication dissolved any bead cluster that might have formed. Experiments started only after the larva went back to swimming mode.

Recordings were done with an AxioZoom V.16 (Carl Zeiss GmbH, Jena) and responses were recorded with a digital CMOS camera ORCA®-Flash-4.0 V2 (Hamamatsu) at 230 to 350 fps. An HXP 200 fluorescence lamp (Carl Zeiss GmbH) at maximum level (using the Zeiss 45 mCherry filter) was switched on only during each recording (lasting max. 7 sec each) to visualize the beads.

To generate water-borne vibrations, a thin 3-5 cm tungsten needle (RS-6063, Roboz) was glued to the centre of a shaft-less vibration motor (EXP-R25-390, Pololu), using a rubber washer to stabilize the needle (**Figure 0-2**C). The motor was switched on for a defined time interval (1-35 ms) and induced the needle to vibrate. The motor was switched on and off with a custom script via an Arduino microcontroller (Arduino UNO R3, Arduino) (see `Motor-control-script.ino` in Appendix). The probe was positioned in focus at a defined distance from the larva with the manual micromanipulator US-3F (Narishige, Japan). Anterior, posterior and side stimulation were tried. A defined set of stimulation values were used for each of the animals tested, the order of the values was randomized for each larva[3]. The behavioral setup was not touched during recording to avoid introducing unwanted vibrations. Between each stimulation attempt, the larva was left to rest for 1 min. Recordings were discarded if the larva was not in swimming mode while being stimulated. The larvae were still alive and visibly healthy one day after being glued. All wild type larvae tested came from different batches. Experiments were performed in a darkened common microscopy room set at 20°C. Temperature was otherwise not controlled.

Extraction of parameters from startle response recordings

All videos were manually analysed by the author of this study. A range of parameters were extracted and registered in a table (see **Table 0-8** in Appendix). Stimulus start was defined as the onset of probe movement (**Figure 1-1**A, black double-headed arrow). Probe speed was calculated by measuring the change in position of the probe between two consecutive frames using the line tool in Fiji and diving that value by the time per frame (**Figure 1-1**B). The maximum speed prior to ciliary closure and that prior to maximal parapodial elevation were registered. As the probe accelerated in a non-linear fashion, the probe speed triggering ciliary closures and that triggering parapodial elevation were not always the same measurement. In all plots shown in this study only maximum speed prior to maximal parapodial elevation are shown. However, using the probe speed prior to ciliary closure does not change the conclusion of the results (data not shown).

[3] List randomizer accessed at: https://www.random.org/lists/

The onset of prototroch and bodytroch closures was defined by the first frame in which the beads around the corresponding ciliary band stopped moving posteriorly (**Figure 1-1**A, black rectangle). The area around the telotroch was analysed to assess bodytroch closures (**Figure 1-1**A, white rectangle). Parapodia elevation onset was defined as the moment after stimulus start at which one of the parapodia in the 1st segment on the main body side (arbitrarily defined) started to move from its resting position against the body (**Figure 1-1**A, black arrowhead). The parapodial elevation angle was calculated from the same parapodium and was defined as the maximum movement of the tip of the parapodium upon stimulation. To measure it, the resting position and the maximum elevation position were determined (**Figure 1-1**C-D). The angle tool implemented in FiJi (Schindelin et al. 2012) was used to measure the angle thus formed. The angle was normalized to the maximum angle observed for that larva. The elevation onset and elevation angle were always recorded from the same parapodium for a given larva.

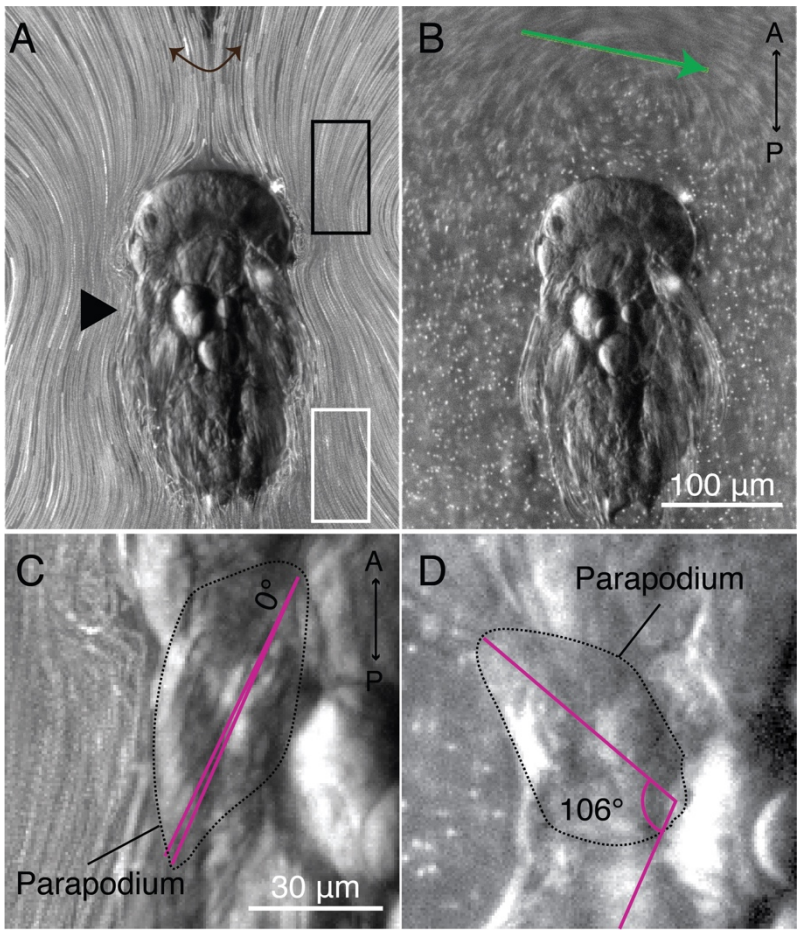


Figure 1-1 Measuring parameters of the startle response. (**A**) Maximum projection of recording used to analyze the kinematics of the startle response. Black and white rectangles indicate the approximate region screened for bead flow to detect prototroch or bodytroch closures, respectively. Black arrowhead points to the parapodium used to measure elevation angle and onset elevation. Brown double head arrow indicates the movement of the filament. (**B**) Snapshot of the frame where the maximum speed probe was observed for that recording. The green arrow was drawn from the location of the filament tip in the previous frame to the position of it in the frame shown. As the probe moved often faster than the temporal resolution, the measurement is only approximate. (**C**-**D**) Measurement of parapodial elevation angle. (**C**) The angle of the parapodium in the resting state is set to 0°. (**D**) In the frame where the parapodia is maximally elevated the lower arm of the angle tool is fixed and the upper arm is aligned to the new position of the parapodium tip.

The temporal coordination between left and right parapodial elevation was analysed by determining the onset of the parapodium elevation on the opposite body side, but in the same segment. For assessing the temporal coordination between parapodia on different segments, the onset of parapodia movement in the 2nd and 3rd segments on the main body side was registered.

The frame rate for each video was extracted from the file metadata. The frame rate could not be set *a priori* as the camera was set to record as fast as it was possible (streaming mode), which is prone to frame rate fluctuations. The dorsal-ventral orientation of the larva, and the distance of the probe from the anterior or posterior side were also registered for each recording. All the measurements were deposited in GitHub[4].

Data analysis

The data were analysed with custom R scripts (see script `FigStartleKynetics.R` in Appendix for details). The latency values, time to maximal elevation, the difference in elevation and closure onset among other variables were computed in this script. The threshold value to split elevation angles into low and wide was estimated with a finite mixture model (Trang et al. 2015) implemented in R (see the script `Startlethresholdestimation.R` in Appendix). All measurements were pooled into the same plot. All plots were generated with these scripts.

[4] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip
(Folder: Kynematics_measurements/)

Results

In their marine environment, *Platynereis* larvae are potentially subjected to mechanical forces as a result of the displacement of water. These hydrodynamic forces could be of big scale such as sound or oceanic currents, which usually are unnoticed by the larvae, or they can be of short range and immediate, such as turbulent flow or water-borne vibrations. I focused in this work on the effect short-range water disturbances have on the behavior of *Platynereis* nectochaete larvae, as they are easier to reproduce in the laboratory and are likely to be ecologically relevant for the larvae.

The startle response in *Platynereis* nectochaete larvae

At the nectochaete stage *Platynereis* move mostly by swimming with ciliary bands with sporadic crawling episodes. Upon producing vibrations above an unknown frequency in the water, larvae stop swimming, and simultaneously elevate all their chaetae-endowed parapodia (**Figure 1-2**A-C). The degree each parapodium is elevated is higher in the first, and lowest in the third segment, but bilaterally symmetrical and segmentally coordinated (**Figure 1-2**C). The decrease in speed and the elevation of parapodia occur in a fast and stereotyped manner (**Figure 1-2**F-G). After this excited stated, the larva usually resumes swimming and start lowering the parapodia (**Figure 1-2**D) until finally reaching the pre-stimulus posture (**Figure 1-2**E). The recovery period, however, is less stereotypic, and slower than the onset phase, with some larvae recovering their posture and swimming speed sooner than others (**Figure 1-2**F-G). This fast, vigorous, and transient response to water-borne vibrations has all the characteristics of a startle response (Bullock 1984), and it will thus be treated as such in the remainder of this work.

Features of the startle response change as a function of stimulation strength and site

To analyze in more detail the startle response, individual larvae were tethered from the trunk region using non-toxic glue and stimulated with a vibrating probe placed at a defined distance from either the anterior, or the posterior side, while recording their responses at fast frame rates (230-350 fps) (**Figure 1-3**A, see Materials and Methods). The speed of the filament was used as a measure of the stimulation strength; the ciliary beating state (i.e. either beating or closed) was indirectly measured with the use of fluorescent beads; and the degree of elevation of the parapodium in the first segment was measured directly from the recorded images (**Figure 1-1**).

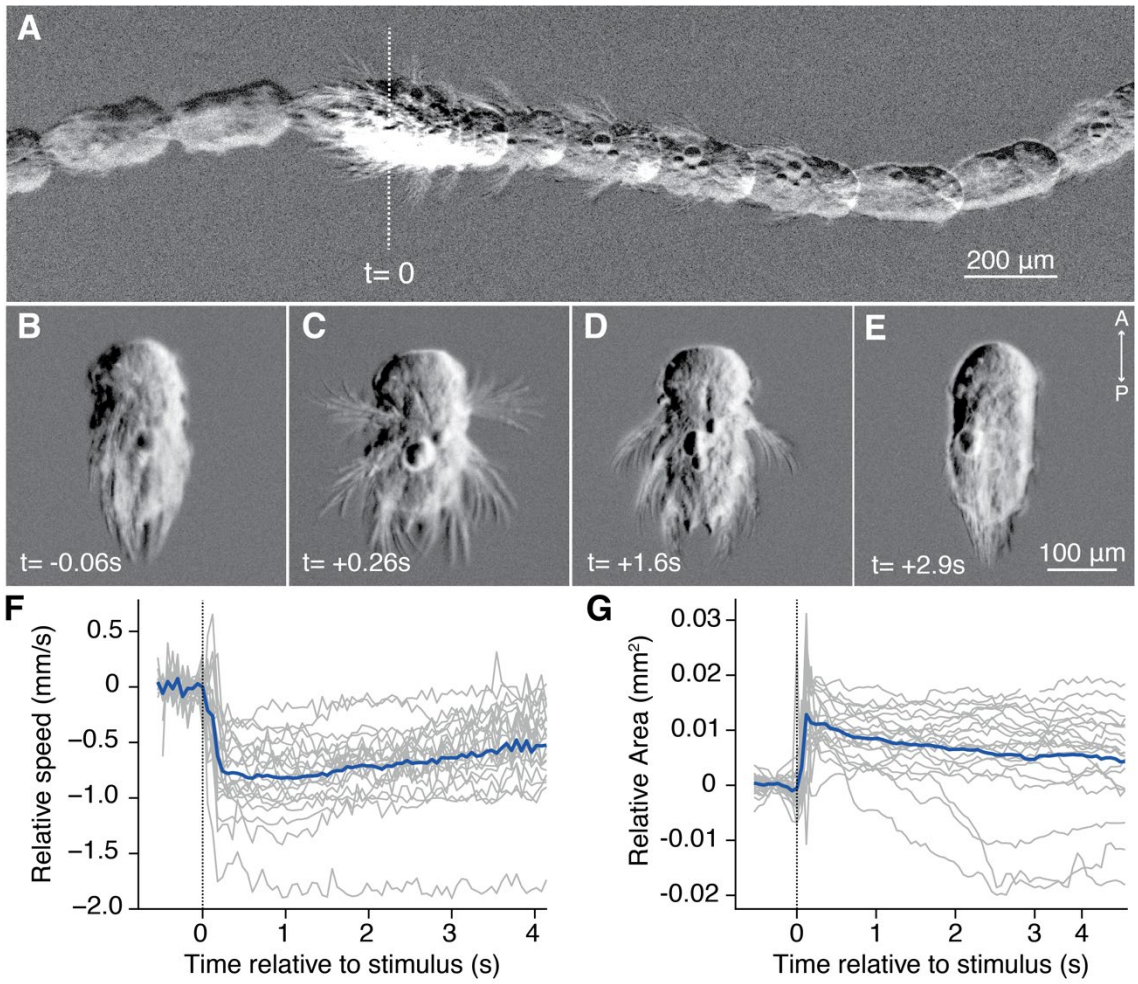


Figure 1-2 The startle response of *Platynereis* nectochaete larvae. (**A**) Summed projection of a time-series showing posture of nectochaete larva before and after vibration stimulus. Larva swimming direction is to the right, and every frame is separated ca. 0.33 seconds from each other. The increased number of overlapping larva profiles right after stimulus indicates the larva stopped swimming. (**B**-**E**) Snapshots of the time series shown in A highlighting different moments in the response. (**B**) Swimming posture ca. 60ms before stimulus started. (**C**) Fully startled posture ca. 260ms after stimulus started. No ciliary beating, all parapodia are lifted and chaetae extended. (**D**) Recovering posture 1.6s after stimulus. Parapodia are already going down and larva started swimming. (**E**) Fully recovered posture ca.2.9s after stimulus started. (**F**, **G**) Plots of change in speed (F) or area (G) relative to average values prior to stimulus of 21 individual larvae that displayed the startle response. Grey traces show the values for individual larva and blue traces show the average value. Time 0 in A, F and G indicates stimulus start (dashed line).

Anterior stimulation

The startle response as a function of the stimulus strength when applied to the anterior region showed a complex but non-random profile (**Figure 1-3**B and E). First, prototroch closures were triggered with filament speeds as little as 4.4 µm ms^{-1} (the minimum speed tested and possible with the current setup was 3.5 µm ms^{-1}). Prototroch closures were invariantly triggered above 9 µm ms^{-1}. Anterior stimulation also triggered the elevation of parapodia with filament speeds as low as 4.4 µm ms^{-1}, but not reliably until above 28 µm ms^{-1} (**Figure 1-3**B and E). The degree at which parapodia were raised varied depending on the stimulation strength. However, contrary to what might have been expected, the degree of parapodial elevation did not increase linearly and uniformly as the stimulus strength increased, but rather formed a bimodal distribution of elevation angles (**Figure 1-3**B).

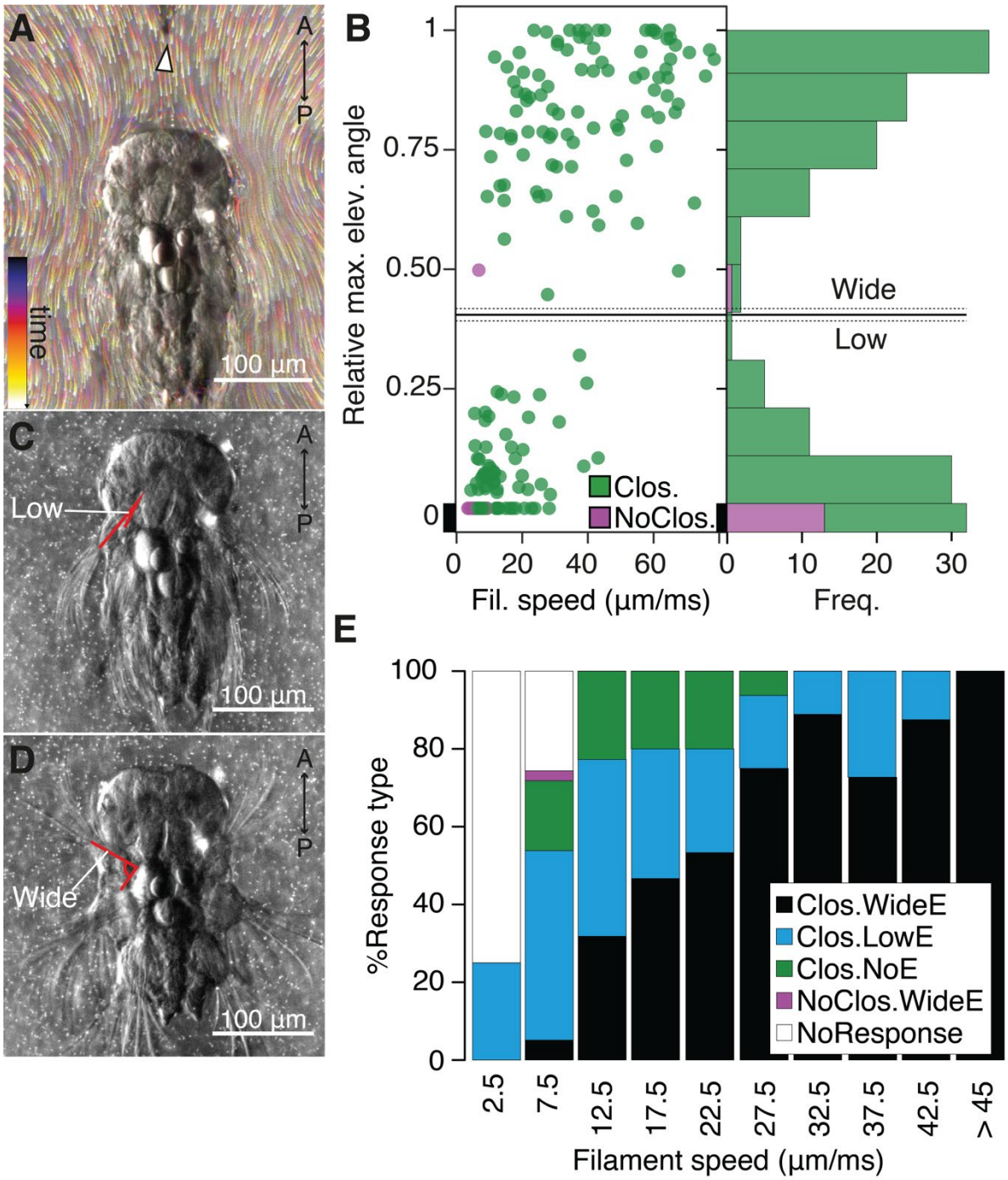


Figure 1-3 The startle response is triggered by anterior stimulation from the head. (**A**) Representative larva tethered glued from the trunk and used for the startle response analysis. Stimulus source is a vibrating tungsten filament placed ~100 µm from the head (arrowhead). 1 µm fluorescent beads were added to visualize the cilia-generated flow around the larva. The larva is in a swimming (non-startled) state prior to stimulation. Red arrow indicates flow direction. (**B**) Scatter plot of the relative maximum parapodial angle as a function of stimulation strength (measured as maximum filament speed). Filament was placed at 100 µm from the head. Stacked histogram of the relative angle frequency is shown to the right and follows the same y-axis scale. Horizontal lines in the scatter plot indicate the threshold estimate (solid line) and 95% confidence interval (dotted lines) that separate Low-angle (LowE) and wide-angle (WideE) elevation responses. The no-elevation bar (i.e. rel. max. angle=0), was expanded for better visualization. (**C**) LowE startled state. Parapodia are slightly elevated, and chaetae are not fully extended (relative maximum angle = 0.16). (**D**) WideE startled state. Parapodia are maximally elevated, and chaetae are completely extended (relative maximum angle = 1). (**E**) Stacked bar plot showing percentages of each parapodial elevation response type upon anterior stimulation. Data presented in A and B were obtained from 175 recordings on 9 larvae, tested multiple times.

To classify the elevation responses into the two apparent categories, a cut-off value of angle elevation was estimated using a finite mixture model (see Materials and Methods) (dashed lines in

Figure 1-3B). Responses below the cut-off value assigned to the low-angle elevation (LowE) category (**Figure 1-3**C), while those above belonged to the wide-angle elevation (WideE) responses (**Figure 1-3**D). LowE responses were observed only at the lower half of the stimulus strength spectrum (up to 43.1 μm ms^{-1}), while WideE responses were observed almost across the whole spectrum (6.8–78.5 μm ms^{-1}) (**Figure 1-3**B, E).

Altogether, anterior stimulation triggered three types of responses: **1**) ciliary closures, but no parapodial elevation (Clos.NoE) **2**) closures and LowE-type parapodial response (Clos.LowE) or **3**) closures and WideE-type parapodial response (Clos.WideE). The proportion of Clos.NoE responses decreased as the probe speed increased, while the opposite was true for Clos.WideE responses (**Figure 1-3**E). The sensitivity threshold was close to the minimum stimulus intensity achievable as only for the lowest values no response was observed.

Timing of the response to anterior stimulation

A number of interesting observations related to the timing of the response were obtained from the recordings. The median latencies of prototroch and bodytroch closures were very similar: 30.8ms (IQR=12.36) and 32.7ms (IQR=14.28), respectively (**Figure 1-4**A). The latencies did not drastically differ when grouped by type of parapodial elevation (**Figure 1-4**B). Both prototroch and bodytroch ciliary arrests always occurred together (i.e. bodytrochs beating was arrested in the 161 times the prototroch beating was arrested upon stimulation). There was a no overall delay between the two events (median delay: 0, IQR=3.6ms, or 1 frame) (**Figure 1-4**D). Likewise, no delay was observed between the timing of either prototroch or bodytrochs closures on the left compared to the right body side (76 measurements, median: delay:0, IQR=0; **Figure 1-4**D).

The median latency for the two parapodial elevation types was 66.8ms (IQR=61.75) and 31ms (IQR=18.21) for LowE and WideE, respectively (**Figure 1-4**A). The variation in LowE latencies, however, resides partly in the difficulty to determine with good temporal precision the onset of the parapodial elevation. LowE responses were characterized by weak and slow elevation (median speed: 253.2 ms, IQR=134ms; (**Figure 1-4**C) of mostly the first segment parapodial group, and without extension of chaetae ((**Figure 1-4**C). WideE responses recapitulated the response in freely swimming larvae, that is, a fast (median speed:93.9ms, IQR=20.91ms; (**Figure 1-4**C) and bilaterally symmetric elevation of all parapodia (and glands) and a wide extension of all the chaetae (**Figure 1-3**D).

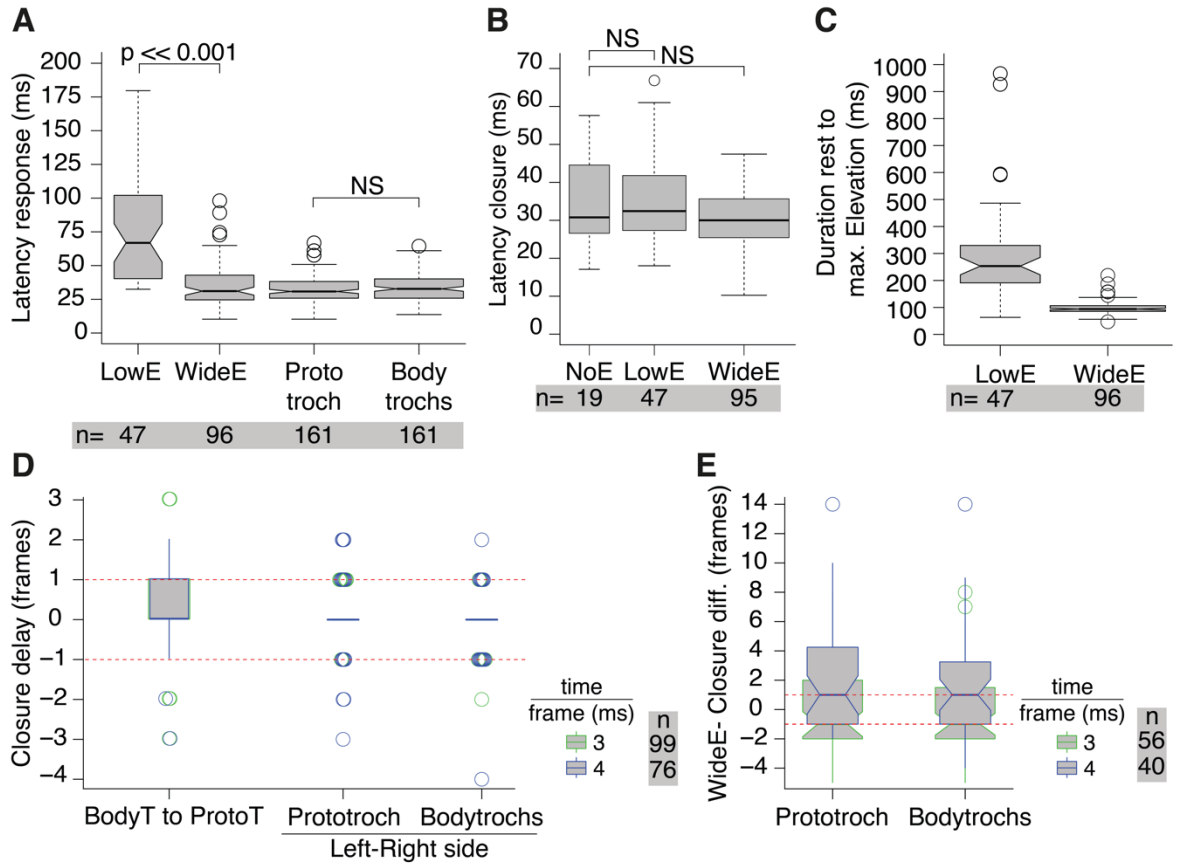


Figure 1-4 Latency and delay statistics of startle response elicited upon anterior stimulation. (**A**) Notched Tukey boxplot showing latency distribution of ciliary band closures, low-angle (LowE) and wide angle (WideE) elevation responses. (**B**) Distribution of latency of prototroch closure sorted by the type of parapodial response. NoE: no elevation. (**C**) Notched Tukey boxplot of time from rest to maximal parapodial elevation for LowE and WideE responses. (**D**) Tukey boxplots of closure delay of bodytrochs relative to prototroch (BodyT to ProtoT, left); closure delay of prototroch (middle) or bodytroch (right) on one side of the body, relative to the other. (**E**) Notched Tukey boxplots of onset delay of WideE relative to ciliary band closures, Measurements are sorted by recording speed in D and E. In all panels the width of each box is proportional to $\sqrt{n}$, and notches display the 95% confidence interval around the median. Notches are not shown in C and E as they extended beyond hinges in at least one case. A two-sided Kolmogorov-Smirnov statistic was used to compare distributions in A and B. NS: not statistically significant.

The similar latency median value of the onset of ciliary closures and that seen in the WideE responses raised the possibility that the delay between both events in individual larvae was actually very small. Indeed, the distribution of the delay in the onset of WideE elevation and the onset of ciliary closures falls within the limits of the recording resolution (median delay=0 for both ciliary band groups, IQR=5 and 5.5 frames for prototrochs and bodytrochs, respectively (**Figure 1-4**E).

Posterior stimulation

The profile obtained upon posterior stimulation with the vibrating filament markedly differed from that obtained from stimulating the head (**Figure 1-5**A-B). Firstly, the minimum speed at which both closures *and* parapodial elevation were triggered was 19 µm ms^{-1}, while only speeds above 56.6 µm ms^{-1} reliably triggered a response. Prototroch and bodytroch closures without parapodial elevation were virtually absent (occurred only once). Five cases were observed of parapodial elevation *without* prototroch closures (only one case was observed upon anterior stimulation).

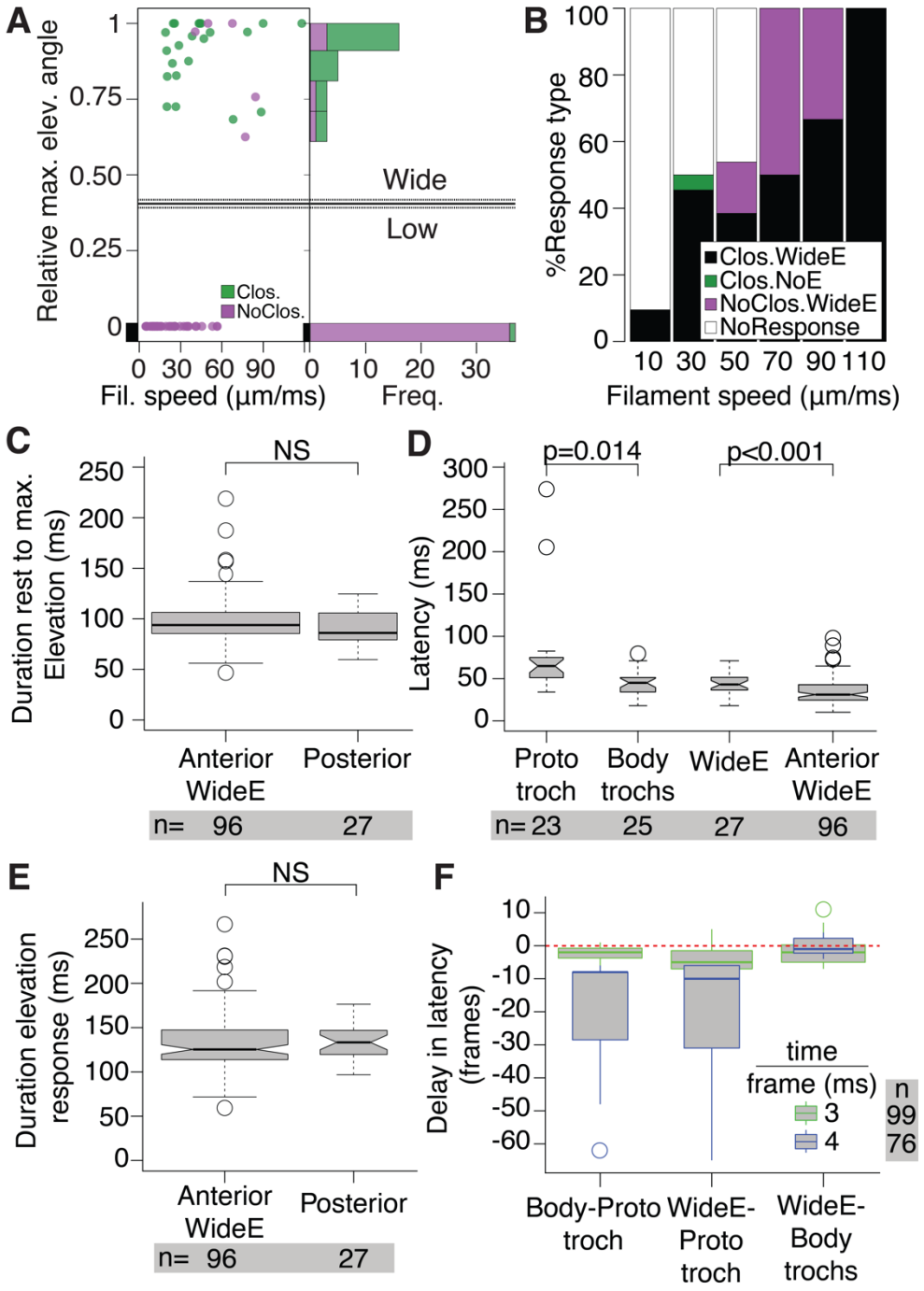


Figure 1-5 The startle response is triggered by posterior stimulation. (**A-F**) Startle response profile and descriptive statistics obtained upon posterior stimulation. Stimulus source placed ~100 µm from the pygidium (**A**) Scatter plot of the relative maximum parapodial angle as a function of stimulation strength (measured as maximum filament speed). Stacked histogram of the relative angle frequency is shown to the right and follows the same y-axis scale. The no-elevation bar (i.e. rel. max. angle=0), was expanded for better visualization. Horizontal lines in the scatter plot indicate the LowE and WideE threshold estimate as defined in **Figure 1-3**. (**B**) Stacked bar plot showing percentages of each parapodial elevation response type upon anterior stimulation**.** (**C**) Tukey boxplots of time from rest to maximal parapodial elevation for WideE responses triggered by anterior (AnteriorWideA) or by posterior (Posterior) stimulation. (**D**) Tukey boxplot of latency distributions for prototroch and bodytrochs ciliary arrests, as well as for parapodial elevation responses upon posterior stimulation. (**E**) Tukey boxplots of the duration of the full parapodial elevation response upon anterior or posterior stimulation. (**F**) Tukey boxplots of delay in latencies of bodytrochs relative to prototroch closure onset (Body-Prototroch), and of parapodial elevation relative to either prototroch (Elevation-Prototroch), or to bodytrochs closure onset. Observations were sorted by recording speed. Red dashed line indicates the no delay point. Width and notches in C-F as defined in **Figure 1-4**. Notches not displayed in cases where they extended beyond the box boundaries. Data presented in A and B were obtained from 64 recordings on 9 larvae, tested multiple times. A two-sided Kolmogorov-Smirnov statistic was used to compare the distributions in C to E. NS: not statistically significant.

Interestingly, parapodial elevations triggered by posterior stimulation were exclusively of the WideE type. Such WideE responses were undistinguishable from those triggered by anterior stimulation in terms of the speed to maximum elevation and bilateral symmetry of the elevation (median values: 93.88 ms and 86.1ms for anterior and posterior stimulation, respectively; **Figure 1-5**C and data not shown). Thus, the startle response profile upon posterior stimulation showed a steep increase in the proportion of Clos.WideE responses as stimulus intensity increased, and along with it, a bigger fraction of NoClos.WideE responses than in the case of anterior stimulation.

Timing of the response to posterior stimulation

The timing and coordination of the responses upon posterior stimulation also showed important differences, specifically in the profile of ciliary closures. The median latency for prototroch closures induced upon posterior stimulation was 65ms (IQR=23.9ms), twice as much as that seen upon anterior stimulation (**Figure 1-5**D). In contrast, the bodytrochs closed with a median latency of 45 ms (IQR=17.1ms). This implies that the prototroch and bodytrochs closed slightly out of phase, with the bodytrochs closing in most cases earlier than the prototroch (median bodytroch to prototroch delay: -9.1ms or 3 frames, IQR=25.7ms or 7 frames; **Figure 1-5**E).

In addition to the higher threshold, the latency of parapodial elevation upon posterior stimulation was slightly but nonetheless significantly longer that seen with anterior stimuli (median latency: 43.1 ms, IQR: 15ms) (**Figure 1-5**D). However, this difference in timing disappeared when the latency and the speed of parapodial elevation were added up (**Figure 1-5**E). The full elevation response is completed within a median duration of 125.3 ms (IQR: 33.5ms) upon anterior stimulation or 133ms (IQR:27.3 ms) upon posterior stimulation. In other words, the WideE response is initiated faster upon anterior stimulation, but the two responses reach peak elevation with similar time efficiency.

Finally, parapodial elevation was more coordinated with bodytrochs than with prototroch closures at the individual level (**Figure 1-5**F). As in the case of the delays measured for anterior stimulation, these delay distributions are sensitive to the recording speed used, as the values are close to the recording speed limit.

A hydrodynamic sensory system is responsible for triggering the startle response

The fact that the startle response was triggered by creating water disturbances near the animal without direct contact, and that the site of stimulation influences the characteristics of the startle response suggests that the local water disturbances created with the filament were directly stimulating some sort of hydrodynamic receptor in the larva and triggering the response. If this were the case, then varying the distance between the probe and the animal would also influence the type of response observed, especially if the response is a result of near-field hydrodynamic receptors.

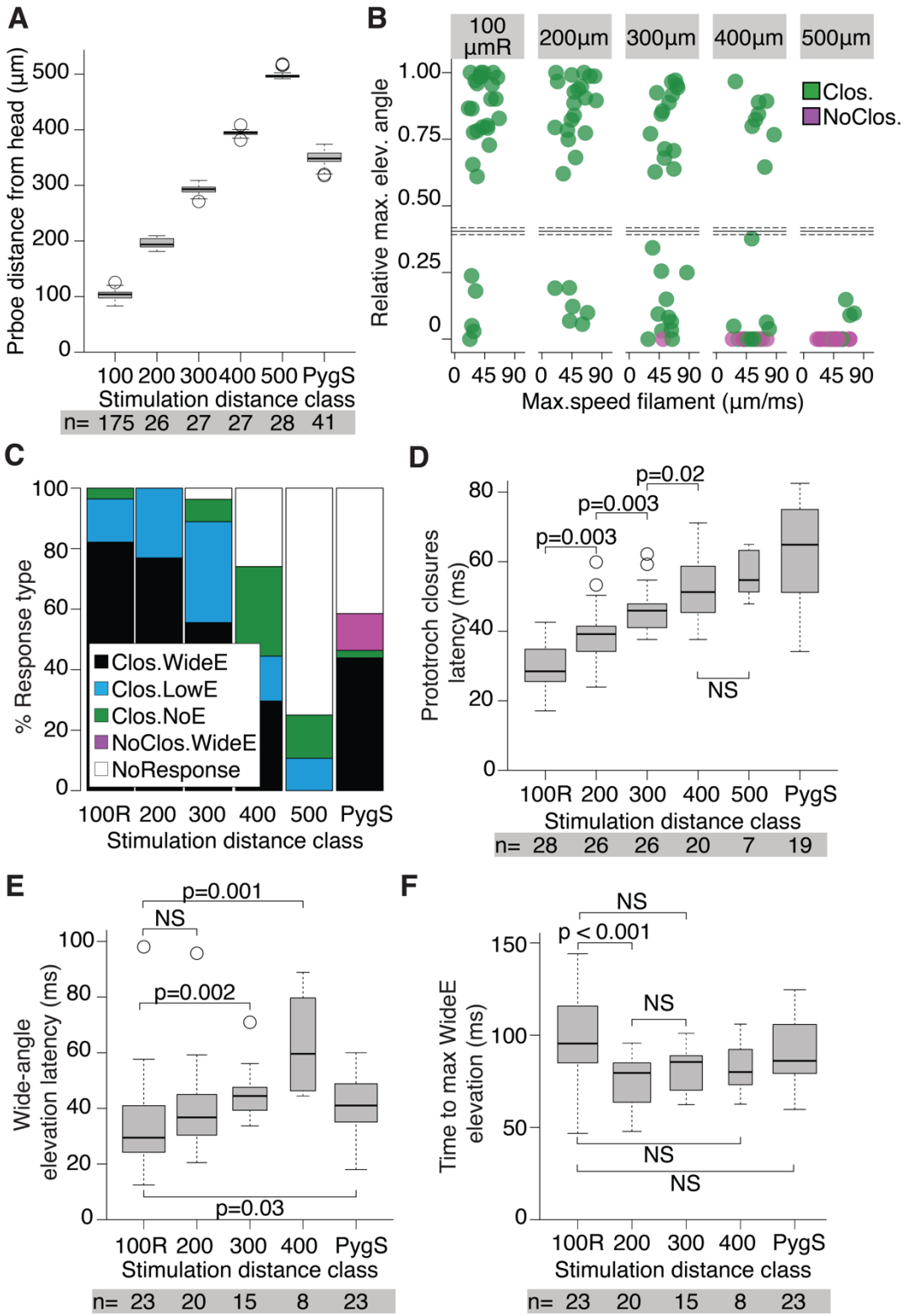


Figure 1-6 Startle response profile at varying stimulation distances. (**A**) Tukey boxplots showing the distribution of probe distances from the head in each stimulation class. (**B**) Scatter plot series of startle response profiles upon anterior stimulation at increasing probe distances. The 100 µmar category is a randomly-chosen subset of the 100 µm data (shown in **Figure 1-3**B) that matches the other categories in number of observations and in stimulus intensity range. (**C**) Stacked bar plot showing the percentages of each response type for each stimulation distance category. The data for posterior stimulation experiments in the equivalent stimulation range (PygS) (shown in **Figure 1-5**) are shown for comparison. (**D**-**F**) Tukey boxplots of latency of prototroch ciliary arrests (D), of latency of WideE responses (E), or of time from rest to maximal parapodial elevation for WideE responses (F) for each stimulation distance category. Boxplot width in C-F is as defined in **Figure 1-4**. Data presented in B and C were obtained from 136 and 177 recordings, respectively, on 9 larvae, tested multiple times. A two-sided Kolmogorov-Smirnov statistic was used to compare distributions in D to F. NS: not statistically significant.

To test this possibility, the distance between head and filament was varied in 100 μm steps, from 200 to 500 μm (**Figure 1-6**A), and the resulting responses to a range of speeds (19.5–86.7 μm ms^{-1}) analyzed as before. Since the range of speeds and number of data points differed between these experiments and that with a 100 μm distance, a subset of the latter spanning a comparable sample size and range was randomly selected for comparison (the 100 μmR category in **Figure 1-6**).

As predicted from a hydrodynamically driven response, the response profile shifted to weaker responses with increasing distance from head to probe (**Figure 1-6**B-C). While the proportion of Clos.WideE responses decreased, the proportion of cases where there was no response increased at longer distances (**Figure 1-6**B). At 500 μm most of the stimulation attempts failed to trigger any response. To address whether the change in the response profile seen upon posterior stimulation (see **Figure 1-5**) was the result of an increase in distance from the anterior end, the distance between the anterior end of the larva and the stimulus filament was measured in the posterior stimulation experiments. Although the median distance was 351 μm (**Figure 1-6**A), the response profile was unlike that seen upon anterior stimulation from distances in the posterior stimulation experiments (300- 400 μm range, **Figure 1-6**B-C). Specifically, in these ranges stimulation still triggered ciliary closures without elevation or closures with LowE responses, which were not seen at all upon posterior stimulation (**Figure 1-6**C).

The latency in initiation of prototroch closures and parapodial elevation also showed an upward change in profile as a function of stimulation distance (**Figure 1-6** D and E). In contrast, the time from rest to maximum elevation in WideE responses did not show a similar shift in the duration (**Figure 1-6**F). Comparisons with the WideE responses upon posterior stimulation revealed that the latency in prototroch closures at increasing distances approach the distribution observed upon pygidium stimulation.

Parapodia are elevated in a synchronous manner during the startle response

At first glance, the elevation of all parapodia during the startle response occurs at the same time in both sides of the body and across segments. To quantify these behavioral features, I analyzed in the recordings of startle response in tethered larvae the timing difference between elevation of parapodia on different body sides and between parapodia on different segments. The delay between elevation of the 1st parapodia on the left and right body sides either upon anterior or posterior stimulation was within the recording speed limits (**Figure 1-7**A-B, right column). A similar result was obtained in the few instances when a lateralized stimulus was used to trigger the startle response (**Figure 1-7**C, left column). Thus, the left and right parapodia on the first segment are elevated within 3 or 4 ms from each other, thus hinting at a mechanism to ensure a bilateral temporal coordination.

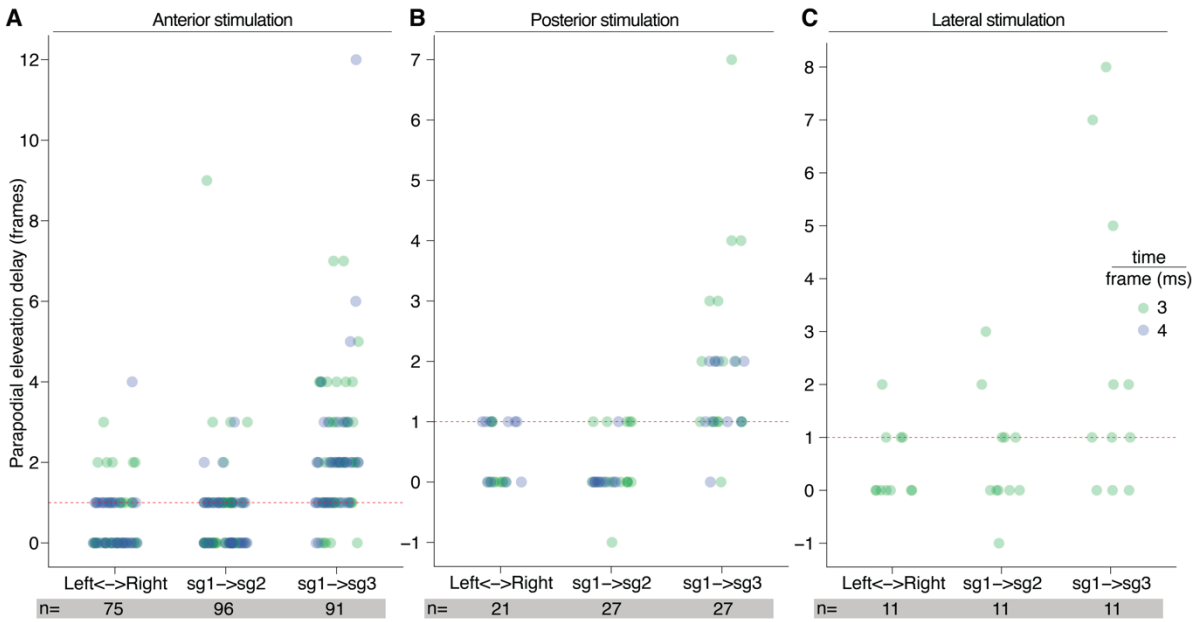


Figure 1-7 Bilateral and intersegmental synchrony of parapodial elevation. (**A**-**C**) Dot plots showing the delay in parapodial elevation across segments and sides of the body upon (A) anterior, (B) posterior, or (C) lateral stimulation. On the left column of each plot the number of frames separating the elevation onset of parapodia in the 1st segment on the left and right body sides is shown. The middle column of each plot shows the delay in number of frames between elevation of the parapodia in the 2nd segment (sg2) relative to the 1st segment (sg1) on the same side of the body. The right column of each plot shows the delay in number of frames between elevation of the parapodia in the 3rd segment (sg3) relative to the 1st segment (sg1) on the same side of the body. Stimulus filament was set at a 100 µm distance from the animal. Red dotted line is placed in all cases at the 1 frame difference. Only WideE type responses are included in the plots. Observations are colored according to recording speed. Data presented were obtained from recordings, on 9 larvae, tested multiple times.

This coordination in time was also expected across elevation of parapodia in different segments. To assess this, the same dataset was used to calculate the delay in the elevation onset of the second and third parapodia relative to the first parapodia on the same side. Indeed, most of the data points of the difference in timing between the elevation of the parapodia on the second and first segments fall between one and cero frames independently of stimulus direction (**Figure 1-7**, middle column). The distribution of the delay in elevation onset of parapodia in the third segment relative to the first was also skewed to positive delay ((**Figure 1-7**, right column). It is important to note that in most of the recordings here analyzed the third parapodium was not in focus. Thus, the delay may include a confounding effect based on the difficulty of precisely assigning the frame at which elevation of parapodia starts in the 3rd segment.

In conclusion, the response is in general temporally coordinated not only in the timing of parapodial elevation and ciliary arrests, but also in the moment at which each set of parapodia is elevated, thus being a good example of fast whole-body coordination task.

Correlation of stimulus strength and stimulus duration may affect measurements of response duration

To understand the neuronal mechanisms controlling the startle response, not only the initial but also the subsequent events are relevant. In the present study, only the events close to the stimulus start were recorded, and thus only some potential parameters could be reliably measured.

The distribution in the duration of prototroch closures triggered upon anterior stimulation showed a similar shape between those occurring without parapodial elevation (median:102 ms,IQR:62.5ms) and those accompanied with a low-angle (median:111.2ms, IQR:95.9ms) parapodial elevation (**Figure 1-8**A). The distribution of closures co-occurring with wide-angle parapodial elevation showed a similar median duration but much more widely spread (median: 151.5ms, IQR:835.9ms; A). A similarly wide distribution in the duration of prototroch closures was observed upon posterior stimulation, as well as a much higher median duration (1020ms; IQR:1278.2ms, A). The duration of closures of bodytroch ciliary bands showed a similar pattern (i.e. narrower distributions for closures not occurring with wide-angle elevations). Interestingly, the duration of bodytroch closures was overall longer than the prototroch closures (median duration for anterior stimulation recordings: 244.8ms; IQR:167ms), irrespective of the type of parapodial elevation (**Figure 1-8**A). In contrast, the delay between the resumption of ciliary beating of the bodytrochs relative to the prototroch was not significantly different among the different parapodial elevation categories, but only between stimulation site categories (**Figure 1-8**B).

The time the larvae maintained the wide-angle elevation of the parapodia was quantified for both anterior and posterior stimulation. The quantification revealed a much narrower distribution than that for the duration of ciliary band closures: a median duration of 161.2ms, (IQR:177.9ms) and of 84ms (IQR:95.7ms) for anterior and posterior stimulation, respectively (**Figure 1-8**C). Videos were in general too short to quantify the time it took the larvae to bring the parapodia to the resting position (i.e. for full parapodial adduction), but it is a feature that can be quantified in future studies.

Given that the stimulus used in this study was a resonant probe, the duration of the stimulus was not uniform across the range of values used. Indeed, the stimulus duration is positively correlated with the maximal filament speed (**Figure 1-8**D). This unwanted effect could be influencing the observed profile in the duration of elevation and of ciliary band closures. Although the median of the distribution of the stimulus duration (167.5ms) is in the same order of magnitude as that of the prototroch closures and maximal parapodial elevation (see values above), there is no statistically significant correlation between the stimulus duration and either of the two measurements (**Figure 1-8**E-F), at least in the case of anterior stimulation experiments. The duration of prototroch closures is relatively constant across the range of stimulus durations, but a subset of observations is an order of magnitude higher. This cloud of observations is mainly composed of closures co-

occurring with wide elevation of parapodia (**Figure 1-8**E). The maximal parapodial elevation is constant across the range of stimulus duration (**Figure 1-8**F).

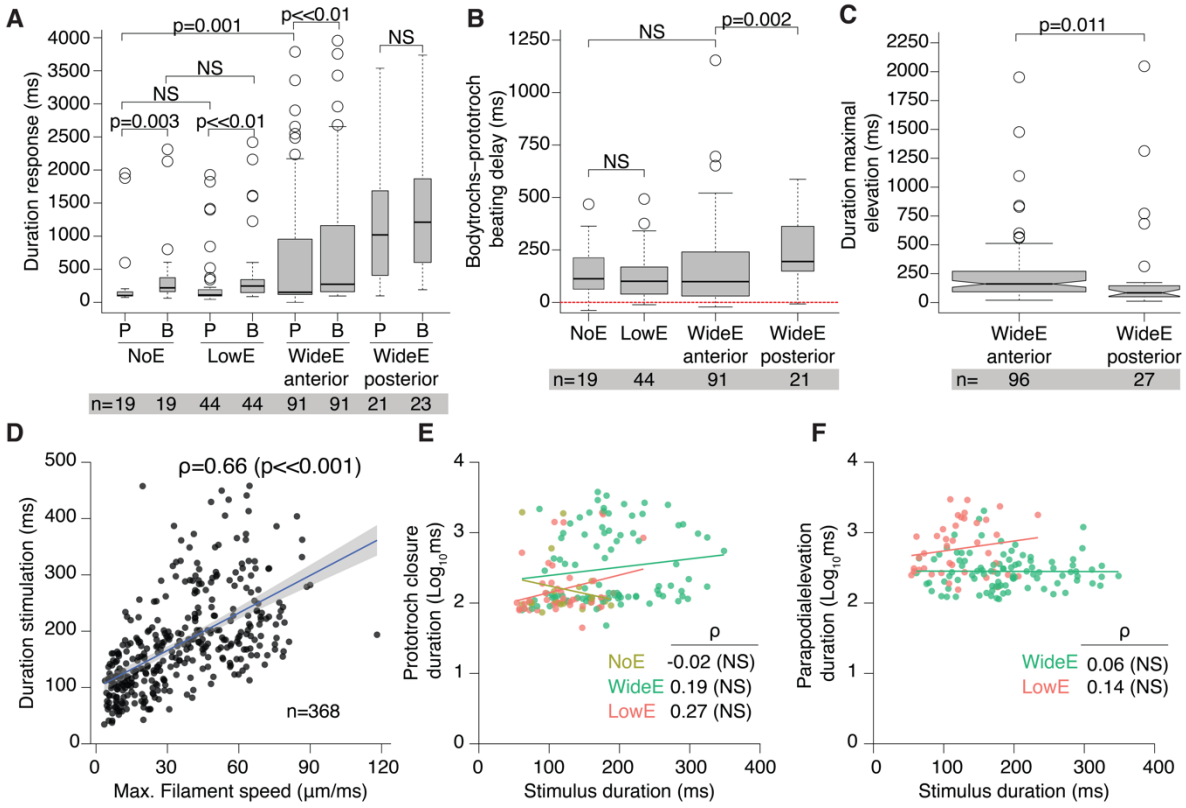


Figure 1-8 Duration of prototroch closure and of parapodial elevation as a function of stimulus duration. (**A**)Tukey boxplots of duration of prototroch closures (P) or bodytrochs closures (B) grouped by type of parapodial elevation response. (**B**) Tukey boxplots of the delay of ciliary beating resumption of the bodytroch relative to the prototroch. Red dashed line is set at 0. (**C**) Tukey boxplots of duration of maximal parapodial elevation. (**D**) Scatter plot of duration of stimulation as a function of maximal filament speed. A regression line is shown in blue with a 95% confidence interval shown as a grey shaded area. Spearman's ρ correlation coefficient is statistically greater than 0. (**E-F**) Scatter plots showing duration of prototroch closure (E) or duration of maximal parapodial elevation (F) as a function of stimulus duration. Regression lines and Spearman's ρ are shown for each elevation type. A two-sided Kolmogorov-Smirnov statistic was used to compare distributions in A to C. NS: not statistically significant ($p>0.05$). Box widths and notches in A to C are defined as in **Figure 1-4**. Notches were not displayed in cases where they extended beyond the box boundaries.

Discussion

In this chapter the startle response in the *Platynereis* nectochaete larva was described in freely swimming and tethered specimens and showed to be triggered by hydrodynamic stimuli. The analysis revealed that it is a fast, coordinated, and stereotyped behavior that involves arrest of all the ciliary bands and the elevation of parapodia. This response has been observed in other annelid and brachiopod larvae (Wilson 1929; Thiel et al. 2017; Pennington and Chia 1984; Okuda 1946), but only reported in anecdotal terms. The fast and coordinated elevation of the parapodia as observed in *Platynereis* is probably also characteristic of those other reported responses, as they were also considered startle responses. It is less clear if any effect on ciliary beating is also observed. The startle response in *Platynereis* could provide a reference to which future studies in other planktonic animals can be compared in quantitative terms, to elucidate commonalities and different startle mechanisms, as it has been done for fish or copepods (Burdick et al. 2007; Hale et al. 2002) .

Overall comparison to other startle responses

Stimulus direction

Most startle and escape responses are modulated according to the location of the stimulus (Bierman et al. 2004; Horridge 1959; Chalfie et al. 1985; Edwards 2017). The analysis of the startle response in tethered *Platynereis* larvae allowed to uncover both qualitative and quantitative differences in the startle response depending on the site of stimulation. Firstly, a lower stimulus threshold was needed to trigger both ciliary closures and parapodial elevation when the stimulus was directed from the anterior than when it was generated from the posterior side. The different sensitivity could be in part due to the number and/or to the intrinsic sensitivity of the sensory receptors on each side. However, not only the threshold sensitivity, but also the overall response profile was different between anterior and posterior stimulation experiments. Although it is not clear what is the ecological relevance of such differences (discussed at greater length in Chapter 4), in other organisms the variation in the response according to the site of stimulation has an adaptive significance, as it usually helps a more effective escape from the threatening stimuli.

Ciliary closures

The arrest or closure of the ciliary bands is one of the main features of the startle response in *Platynereis*. Ciliary band closures in this animal occur spontaneously under the control of an internal rhythm (Verasztó et al. 2017; Tosches et al. 2014). Ciliary beating is altered by neuroendocrine factors (Conzelmann et al. 2011), or in response to light stimuli (Jékely et al. 2008). This report shows that mechanical stimulation also alters the beating state of the ciliary bands in *Platynereis* as it does so in the unicellular protist *Paramecium*(Jennings 1900; Jennings 1904), in other polychaete

(Lacalli 1986), mollusk (Mackie et al. 1976) and echinoderm larvae(Strathmann 1971; Lacalli and Gilmour 1990).

Ciliary closures caused by the vibration stimulus were almost always observed in both prototroch and bodytrochs. The degree of synchrony varied depending on the site of stimulation. Upon anterior stimulation, the delay between closure of bodytrochs to prototroch was at the limit of the temporal resolution of the recordings (3-4 ms per frame). Such synchrony has also been seen in spontaneous closures (although it was not quantified, a small, but noticeable delay between prototroch and bodytrochs is still present in Figure1E of (Verasztó et al. 2017)). The order of ciliary band closures induced from anterior stimulation is also the same as in spontaneous events: prototroch closing earlier than bodytrochs in most cases. This suggests that ciliary arrests during the startle response are also neuronally controlled, the signal probably converging onto the neurons proposed to be responsible for ciliary band closures: the cholinergic ciliomotor neurons MC and Loop (Verasztó et al. 2017). A pacemaker circuit is hypothesized to synchronize these neurons' activity during spontaneous closures, so the signal could be converging onto these neurons. A more careful comparison between the delay observed upon mechanical stimulation and that seen in spontaneous closures could help determine the degree of similarity between these two types of closures.

On the other hand, posterior stimulation revealed that ciliary closures in prototroch and bodytrochs can also be temporally uncoupled. This may be due to a differential control of the MC, which only innervates the prototroch, and of the Loop neurons, which mainly target the bodytrochs (although they also synapse onto some cells of the prototroch) (Verasztó et al. 2017). Interestingly, the closure delay changed sign compared to anterior stimulation: bodytrochs closing earlier than prototrochs in most cases. In two cases, bodytroch beating was arrested while the prototroch kept beating (data not shown). Although its physiological significance of this is not clear, this observation highlights the possibility of finer control of the activity of the different ciliary bands. It also implies that the induction of closures in this case could have been due to the activation of different neuronal pathways acting separately on the bodytrochs and on the prototroch.

Tight synchrony in ciliary closures was also observed between left and right body sides. Two non-mutually exclusive reasons may explain the bilaterally synchronous arrest of the prototroch. First, all the cells in this ciliary band are innervated by the MC cell (Verasztó et al. 2017).And second, electrical coupling between prototroch cells may enhance coordination of ciliary arrests. Such coupling mechanism has been reported in the ciliary bands of the gastropod veliger (Arkett et al. 1987). The bilaterally synchronous arrest of the bodytrochs has to be explained by different mechanisms, as the bodytroch bands are not continuous across both body sides, and there is not a single neuron innervating all the bodytrochs. One possible mechanism could be through the

proposed pacemaker circuit acting on the ipsilateral Loop neurons (Verasztó et al. 2017). A bilaterally synchronous closure of the bodytrochs might be also simply a consequence of a symmetric stimulus. Lateral stimulation experiments could be used to assess if this is the case.

Ciliary arrests have been shown to be due in some cases to the intrinsic mechanosensitivity of ciliary band cells (Stommel 1986). The arrests during the startle response are unlikely to be solely due to this for a couple of reasons. First, the synchrony of prototroch and bodytroch closures was preserved independently of stimulus strength. Due to the sharp decline in the propagation of vibrations at small scales (Visser 2001), it would be rather expected that at least at the lowest intensities those ciliary band cells closer to the stimulus would be closed first than those farther away. And second, the closures were dependent on the site of stimulation: weak stimuli readily trigger ciliary arrests when stimulus was anterior, while the same stimuli did not trigger any closures when it was located posteriorly. Intrinsic mechanosensitivity of ciliary band cells would rather trigger closures at similar stimulus strength only dependent on the distance to the stimulus source.

Elevation of parapodia

The other main behavioral feature of the startle response is the elevation of the parapodia, a behavior likely driven by the simultaneous contraction of striated muscles, which are already well-developed by this larval stage (Fischer et al. 2010; Brunet et al. 2016). This synchronous locomotor behavior is quite different to parapodia crawling or swimming in adult polychaetes, where a consecutive muscle control is in action (Kristan 2018). The parapodial elevation response shows some features seen in other startle responses.

One initial hypothesis was that the degree of parapodial elevation would be a continuous function of the stimulus strength. Instead, the parapodial elevation responses upon anterior stimulation fell into a low-angle elevation (LowE) and a wide-angle elevation (WideE). In addition to the degree of elevation, the two responses differed in the variability in the latency and in the time to maximal parapodial elevation, with LowE responses being overall much slower and variable. Such bimodality is a common feature in escape behaviors in other organisms such as *Drosophila*, fish or crayfish (von Reyn et al. 2014; Edwards et al. 1999; Troconis et al. 2017). As in the fly or in the fish, the probability of the faster response, in this case the WideE response, increased as a function of stimulus strength. A further similarity to faster responses such as the WideE response is the more stereotyped (i.e. less variable) temporal parameters such as latency and time of execution. Such temporal precision in response initiation is thought to reflect the adaptive importance of startle responses for survival, where a few milliseconds could make the difference between life and death.

In cases where bimodal responses are observed, semi-independent neuronal pathways are thought to control the different responses (von Reyn et al. 2014; Bhattacharyya et al. 2017). Posterior

stimulation elicited WideE but not LowE responses. All WideE responses had similar dynamics, thus suggesting that they may be controlled by the same neuronal mechanism irrespective of the site of stimulation. Such mechanism may be distinct from the neuronal pathway responsible for LowE responses.

Segmental and bilateral coordination

Wide-angle parapodial elevation responses were remarkably synchronous across segments, independently of the site of stimulation. Parapodial movement in other nereids are controlled by muscles intrinsic to the parapodia, as well as by muscles attached to the aciculae (Tzetlin and Filippova 2005; Mettam 1967), a set of modified chaetae that provide support to each parapodium (Hausen 2005). Since parapodial musculature is segmental in nature (Tzetlin and Filippova 2005; Mettam 1967), a neuronal mechanism of intersegmental activation is needed to drive the simultaneous elevation of all parapodia on both sides of the body. Intersegmental muscle activation is a common feature of startle responses in vertebrates and invertebrates, as it is required to coordinately translate the body away from the stimulus. The parapodial elevation in *Platynereis* is a further behavior that could be compared at the neuronal level with other mechanisms of whole-body muscle control.

Another interesting feature of the WideE responses is the bilateral symmetry of muscle contraction in both timing and extent. Bilaterally synchronous contraction has been observed in other startle responses such as the acoustic startle response in mammals (Grosse and Brown 2003; Yeomans et al. 2002). Bilateral symmetry in time and amplitude of muscle contraction may also occur in those responses where the vector of motion is along the anteroposterior axis, such as the crayfish tail-flip response, or the whole-body shortening in the leech, although no measurement could be found in the literature. Bilateral muscle contraction may be a more widespread feature of startle responses, as even in lateralized responses, such as the escape response in fish, synchronous bilateral muscle contraction is observed, although the extent of contraction is often not symmetric (Tytell and Lauder 2002; Westneat et al. 1998; Hale et al. 2002). As there is no report of extensive musculature crossing the midline in polychaetes (Tzetlin and Filippova 2005), the symmetry in the extent of parapodial elevation must be due to an intrinsic symmetry on the neuronal pathway controlling the elevation. Since the bilateral synchrony in parapodial elevation was preserved even upon lateral stimulation, a mechanism that monitors and compensates for any resulting asymmetry in the stimulus must be in place (Heckscher et al. 2015; Murchison et al. 1993).

Future kinematic analyzes with better temporal resolution could refine the measurements of the degree of bilateral and intersegmental coordination. For example, in the present study it was noted that WideE parapodial responses tended to start from anterior to posterior, but the lack of good temporal resolution prevented a more conclusive result. Additional parameters to be measured in

future experiments include the intra-segmental delay of parapodial elevation (in this study only the 1st segment was analyzed). An explicit measure of the degree of symmetry in the extent of bilateral parapodial elevation is also desirable for future experiments.

Coordination between ciliary band closures and parapodial elevation

As discussed above, the startle response involves on one side the tight coordination of all the ciliary bands to ensure an immediate pause in swimming and on the other the simultaneous muscle contraction that causes parapodial elevation. Characterization of the control dynamics of both systems indicates that the two systems can act independently of each other, but in some cases coordination between the two systems may also be required. For instance, coincident onset of ciliary arrests and WideE responses was observed upon anterior elevation, but the latency of ciliary closures was not affected by the type of parapodial elevation response. The fact that ciliary closures elicited upon posterior stimulation only occurred when parapodia were elevated suggests a common control mechanism. Coincident activation of ciliary closure and muscles upon mechanical stimulation has been reported in mollusk larvae (Mackie et al. 1976), although no quantification of the latencies was reported. Coordinated (but not entirely coincident) ciliary closure and muscle contraction is also observed in the siphon of the Ascidian *Corella* (Mackie et al. 1974). Coordination of muscle contraction and ciliary beating has been observed in other behaviors relevant for planktonic life.

On the other hand, anterior stimulation experiments clearly showed that the two locomotor systems can be independently activated during the startle response. Moreover, the stimulus threshold for activation seems to be different for each system. This observation does not totally rule out a common pathway controlling cilia and muscles, but it speaks of a type of grading response that selects the response based on the site of stimulation and nature or intensity of the stimulus. Such graded activation was also seen in the experiments on mollusk larvae (Mackie et al. 1976). The discussion on the neuronal pathways will be continued in Chapter 5, once the putative startle circuit is presented.

Stimuli triggering the startle response

Hydrodynamic stimuli

The startle response in *Platynereis* larvae was shown in this study to be elicited by near-field hydrodynamic stimuli. First, the use of a vibrating filament near but not touching the larva was sufficient to elicit the complete response. And second, a decrease in the probability of the response was observed with increasing distances between the stimulus source and the larva. This is consistent with the decay of near-field hydrodynamic forces following a as a power law function of distance (Visser 2001). An increase in response threshold with an increase of the distance of the stimulus source is seen in other behaviors driven by near-field water disturbances (Lenz and Yen 1993).

Hydrodynamic disturbances of some kind induce most, if not all, startle responses reported in other planktonic animals (see the Introduction to this chapter). These stimuli are perhaps the fastest and most precise way at the small scales typical for zooplankton to locate an approaching predator (Martens et al. 2015)—probably the main reason for displaying a fast startle reaction.

The fast decay in the hydrodynamic signal indicates that even such small animals need to have responses adapted to the location of the stimulus. The difference in responses to anterior and posterior stimulation could not be solely explained by an increase in distances of the stimulus source to the anterior side of the animal. Thus, the larvae may be able to detect the location of the stimulus. Other small plankton such as copepods are able to locate the direction of the hydrodynamic signal and display an appropriate escape response (Buskey et al. 2002; Buskey et al. 2017). In copepods, this ability is partly assigned to the long antennae that can detect minute changes in fluid deformation at their distal tips (Takagi and Hartline 2018). How do Platynereis larva achieve directional specificity is still unclear.

Also interesting was to see an upward shift in the latency distribution for both closures and elevation responses as the probe was moved farther away from the larval head. The shift in the latency could be seen as a delay in the signal to arrive to the receptors. Related to this, it is worth noting that the response latency of closures upon anterior stimulation had a similar distribution to that observed by posterior stimulation when the stimulus source was at similar distances from the anterior end of the animal. Although perhaps only superficial, this similarity leaves open the possibility that at least prototroch closures seen in posterior stimulation experiments are a result of indirect anterior stimulation.

Other stimuli

Although the focus of the study was on the effect of hydrodynamic stimuli on the startle response, it does not mean that other stimuli cannot trigger it. In fact, touch stimulation with the tungsten filament on tethered animals also triggers a similar response (data not shown), although it was not assessed if it had the same profile as hydrodynamically triggered ones. Touch may enhance the response to vibration. Both touch and flow stimuli are reported to trigger startle or escape responses in other planktonic organisms (Kirk and Gilbert 1988; Gilbert 1985). It can be conceivable that touch may be needed to maintain the parapodia in an elevated position in case of a continued attack.

Although there is no evidence that changes in light intensity elicit the light response, it is known that in adult polychaetes, abrupt decreases in light intensity trigger a fast muscle contraction. The eye circuit previously reported (Randel et al. 2014; Randel et al. 2015) does not suggest a clear pathway that could explain all the behavioral features of the startle response. Nonetheless, other small planktonic animals show startle responses to sudden changes in light intensity (Buskey and

Hartline 2003), thus leaving open the possibility that the same or a different startle response in *Platynereis* could be elicited by photic stimuli.

A few reflections on the method of stimulation

Although hydrodynamic stimuli elicit the startle response, the imprecision of the method used for generating the stimulus makes it hard to determine what parameter or set of parameters is most relevant for triggering the response. One of the caveats in the stimulation method was that the movement pattern of the filament could not be directly controlled. This produced a complex and unpredictable mix of fluid patterns, as well as variable rates of peak velocity that often occurred after the response was initiated. Second, the duration of the stimulus could not be precisely controlled. In fact, the probe speed was positively correlated with the stimulus duration, when it would have been better to keep stimulus duration constant and below the latency of the response. The vibration frequency could likewise not be controlled, but this parameter is likely to be less relevant for small freely swimming animals (Visser 2001). Finally, maximal filament speed was used as a measure of the abruptness of the stimulus. However, as this method depended on the recording speed, the position of the filament could not always be located with precision from frame to frame. Determination of the parameter triggering the response is relevant as an approaching predator can in principle be detected by a combination of vorticity, fluid deformation, fluid velocity or fluid acceleration parameters (Kiørboe and Visser 1999). While some startle responses are triggered by a defined level of fluid deformation (such as in unicellular organisms (Jakobsen 2001)), others require a change in the rate of fluid deformation (such as in copepods) (Fields and Yen 1992; Kiørboe and Visser 1999; Kiørboe et al. 1999). Controlling the stimulation parameters can also help determine sensitivity thresholds and allow comparisons between species and across developmental stages.

To solve the stimulation issues mentioned above, future experiments should make use of a piezo-controlled mechanical dipoles that generate step stimulus with predictable velocities and flow deformations (e.g. in (Lenz and Hartline 1999)). A different stimulation method encountered in the literature is the use of a suctioning siphon (Kirk and Gilbert 1988; Jakobsen 2001; Singarajah 1969). The flow in this system is laminar, adjustable, symmetric and irrotational (Visser 2001), and it can be used on freely swimming larvae. In the long run, naturalistic stimuli would be key to place the different properties of the behavior in an ethological context. The parameters of such stimuli could be derived from measurements of the hydrodynamic disturbances created by natural interactions.

Future studies of the startle response in *Platynereis*

There are many future lines of inquiry that could be followed to better understand the startle response in *Platynereis*. Some of them are discussed below.

Additional phases of the startle response

After the initial startle phase, the larva eventually needs to restart ciliary beating and bring the parapodia back to its resting position. Although the analysis presented here focused on the initiation of the startle response, the ensuing events that determine its length and conclusion are likely under neuronal control, and most of them are probably quantifiable.

One of such features quantified was the duration of ciliary arrests. Previous reports have found spontaneous prototroch closures in trochophore larvae to be bimodally distributed, short closures having a median of 200 ms and long closures around 1 second (Tosches et al. 2014). The prototroch closures measured in this study have a median of 128.5ms, but as the larval stage and the experimental setups are different, it is not possible to know the source for such difference. The duration of closures was affected by the type of parapodial elevation observed, in particular by wide-angle elevation. Although the median duration was in the same order of magnitude, drastically longer closures were observed more often when wide-angle elevation of parapodia occurred. Long closures were also observed upon posterior stimulation. No additional variable was found that could explain why in some cases longer closures were observed (although one possibility may be the internal state of individual larvae). In other ciliated larvae, prolonged ciliary arrests are due to repeated depolarization of the ciliary band cells (Mackie et al. 1976). Thus, the maintenance of the ciliary arrest during WideE responses may be an actively controlled process. In theory, longer arrests imply longer sinking distances, but the elevation of parapodia may modify the sinking rate of the larva, as previously suggested (Pennington and Chia 1984).

In most cases bodytrochs resumed beating only after the prototroch (a median of 100ms after). This pattern was also seen in spontaneous closures, although the duration of each was not quantified (Verasztó et al. 2017). This may reflect that resumption of ciliary beating in both cases is controlled by a common mechanism. The order of beating may not have an adaptive reason but only obey design principles (in an evolutionary sense).

The termination of the response was not analyzed in detail due to the short length of the recordings. The only observation made of this phase was on freely swimming animals. A wide variation was observed in the time to recover pre-stimulus speed and posture. However, a more controlled assay is needed to rule out unwanted secondary stimulation artifacts.

Ontogeny of the response

The startle response observed in the larva may be further developed as an adult. Indeed, atokous *Nereis* worms also extend the chaetae and parapodia if stimulated from the front (Horridge 1959) . The same author also noted that the response also differs if the worm is stimulated from the back, as the chaetae are projected backwards. The response may be gradually developed starting from the time larvae hatch. At this stage, larvae do not have chaetae, but it is conceivable that they will

be still able to deploy ciliary arrests upon hydrodynamic stimuli, as other polychaete larvae do upon mechanical stimulation (Lacalli 1986). The behavioral features of the response could be compared across development to understand the stage-specific adaptations of the response.

Dependence of the startle response on behavioral and internal states

The startle response may not only change across developmental stages, but it may also do so depending on the internal and behavioral state of the animal. The current experiments were performed on larvae undergoing 'fictive' swimming prior to the stimulus. It is not known if larvae crawling or sitting on the bottom get startled at all. If they do, the type of stimulus parameter detected has to be different. While a passive animal needs to detect absolute velocity of the fluid, a swimming animal needs to detect a differential speed (Visser 2001). The crawling state will also imply the parapodia will need to be elevated from different resting positions. Ciliary beating persists during crawling (Chartier 2017), and thus it may also need to be controlled in this behavioral state.

The sensitivity to the hydrodynamic stimuli may also depend on the internal state of the animal. *C.elegans* shows changes in the sensitivity to touch according to its experience (Chen and Chalfie 2014). The stimulation history affects the internal state of the animal leading to varying responses to the same mechanical stimulus. A similar approach could be followed on *Platynereis*, as it is amenable for neuropeptide and neurotransmitter treatment, and it has a conserved battery of neuromodulatory molecules but with unknown functions (Williams et al. 2017; Conzelmann, Williams, Krug, et al. 2013).

Chapter 2 Identification of hydrodynamic receptors by morphology and calcium imaging

"Observation sets the problem,

experiment solves it"

-Hugh Raffles

Statement of contributions and publication status

The TEM data was acquired by Reza Shahidi. Reconstruction of neurons and muscles in the EM volume was done in part by Martin Gühmann, Sanja Jasek, Gáspár Jékely, Reza Shahidi and Csaba Verasztó. All the SEM data shown in this chapter including sample preparation was acquired by Jürgen Berger. Some of the cloning steps leading to the construct used for calcium imaging were done by Sara Mendes. The rest of the work presented in this Chapter was carried out by the author of this thesis including the acquisition of snapshots and the assembly of figures.

The electron microscopy volume was previously reported (Randel et al. 2015). Part of the anatomical data and of the calcium imaging data shown in this chapter were published elsewhere in abbreviated form (Bezares-Calderón et al. 2018).

Introduction

The diversity of mechanosensory neurons

A single organism is endowed with a variety of mechanosensory systems, each specialized in a particular function that extracts information from external and internal mechanical forces—oftentimes with high sensitivity and specificity—to drive behavior or to regulate physiological processes. The diversity in mechanosensory systems is reflected in the diversity in mechanosensory structures. For instance, both mammals and insects have at least three morphologically and spatially distinct proprioceptor systems in the limbs specialized in detecting different force parameters, such as the speed of limb movement, or the mechanical load on a limb (Tuthill and Wilson 2016).

The morphology of a mechanosensory cell together with the material properties of the tissue in which it is embedded dictates its final mechanosensory modality and selectivity (Katta et al. 2015). One prime example is the labyrinth in the internal ear of terrestrial vertebrates, where the mechanosensitive cells, the hair cells, have a directional selectivity in part conferred by the asymmetrical shape of their sensory dendrite and their relative orientation to the stimulus, while the tissue morphology of the organ of Corti and of the vestibular organ determine that the hair cells in each specialize to detect sound, or acceleration, respectively (Hudspeth 2014).

Hydrodynamic sensory cells in animals

The oscillation of vibrating structures in the water lead to the propagation of a mechanical force that can be detected by specialized mechanosensory cells called hydrodynamic receptors(Bleckmann 1994). Water-borne vibrations can be divided into near-field, if water flow oscillates around an oscillatory source with high amplitude but fast decay with distance, or far field if a medium oscillation is accompanied by a pressure wave and travel for longer distances. Sound belongs to the latter type and involves an alternating change in pressure, but in aquatic animals these changes in pressure are transformed to displacement that can be in turn detected by hydrodynamic receptors (Markl 1978; Markl 1973). However, far field stimuli can only be effectively produced by relatively large structures, and thus small plankton only are able to generate near-field displacements (Gallager 1993). Thus, this type of vibrations is the most relevant for interactions between plankton.

It is usually conceived that touch and vibration receptors are the simplest of mechanoreceptors, serving as the evolutionary basis for more complex mechanosensory cells, or organs (Markl 1978). Vibration receptors (hereafter hydrodynamic sensory cells) are present in a wide range of organisms, from vertebrates to cnidarians, and they usually have a deformable structure such as a cilium or a cuticular hair exposed to the local hydrodynamic fields (Budelmann, 1989). Like touch

sensory cells, hydrodynamic receptors are usually found distributed across an animal's body, but they may extract additional parameters from local hydrodynamic events (Bleckmann 1994).

The vertebrate hair cell is a well-known class of hydrodynamic receptor. In teleost fish this secondary sensory cell is composed of a single non-motile cilium—usually called kinocilium—with a 9x2+2 microtubule pattern, and rows of microvilli—somewhat misleadingly called stereocilia—that form a staircase-like structure, tallest at the end closest to the cilium. Stereocilia are linked to each other by tip links, which are thought to serve as gate springs opening the mechanotransduction channels (Pickles et al. 1984; Corey and Hudspeth 1983), and to kinociliary links, which have a role in mechanosensation only during hair cell development (Kindt et al. 2012).

One of the few physiological characterizations of a hydrodynamic system in invertebrates is found in the epidermal lines found in cephalopods. Summed electric potentials—also called microphonic potentials—were recorded from these lines upon near-field water displacements (Budelmann and Bleckmann 1988). Its response depends on the amplitude and frequency of the stimulus (Bleckmann et al. 1991). Although this organ is considered an analog organ to the lateral line in vertebrates, it is composed of multiciliated sensory neurons without microvilli collars (Sundermann 1983). Thus, it does not bear a direct morphological resemblance to vertebrate hair cells. Direct recordings from these cells have not been reported, yet.

Arthropod hydrodynamic receptors have also been characterized in some detail. These receptors show a great variety of morphologies, but the basic pattern consists of cuticular hairs innervated by sensory cells (Budelmann 1989; Tautz and Sandeman 1980; Laverack 1962). In crayfish, near-field hydrodynamic receptors located in the telson have a single soft cuticular hair innervated by two sensory cells. This arrangement confers them with directional sensitivity to hydrodynamic stimuli (Wiese 1976). Crayfish have two more types of hydrodynamic receptors in the antennae, called upright and procumbent hairs, one directly sensing water vibrations, while the other indirectly doing it so by detecting the movement of the antennae by water disturbances (Tautz et al. 1981; Masters et al. 1982). In Paenid shrimps, antennae covered with numerous feather-like hairs form a structure proposed to have an analogous function to the lateral line in vertebrates (Denton and Gray 1985).

The cnidarian sea anemone also has hydrodynamic receptors strikingly similar to hair cells in vertebrates. A hair bundle complex formed by a cell with a single cilium and a collar of five to seven microvilli is flanked by cells with numerous microvilli (Mire-Thibodeaux and Watson 1994). When stimulated, this complex shows electrical currents in response to water puffs (Mire and Watson 1997). These cells show asymmetry in the responses to mechanical stimuli and similar pharmacological properties to vertebrate hair cells (Mire and Watson 1997; Watson et al. 1997). Hair cells in another Cnidarian, the hydrozoan medusa *Aglantha digitale* have a single non-motile

cilium surrounded by a collar of microvilli (Singla 1983). These cells have been shown to respond to tactile stimulation, but not conclusively shown to be also sensitive to water vibrations (Arkett et al. 1988). In demosponges, a group of biciliated cells lining the epithelium of a flow-sensitive organ have a pharmacology similar to vertebrate hair cells, and loss of cilia is correlated with loss of flow detection (Ludeman et al. 2014).

Collar cells as hydrodynamic receptors

A cilium associated to a group of microvilli is a set of traits shared among many sensory cells in animals across the animal phylogeny. This morphology bears resemblance to the choanocyte in sponges, which in turn gets its name from the similarity to the choanoflagellates, the extant unicellular group of organisms most closely related to animals. As exemplified above, this choanocyte-like sensory morphology is also seen in cells associated to organs with putative hydrodynamic sensory function in multiple animal phyla, and collectively called as collar cells (Budelmann 1989). In few cases other than hair cells, however, this sensory modality has been directly confirmed.

Some examples illustrate the different degree of evidence supporting the function of such collar cells have a hydrodynamic sensory modality. In the sea squirt *Ciona*, a non-vertebrate chordate, cells with a cilium and a collar of microvilli are found in the cupular organ and in the coronal organ, which are two putative hydrodynamic sensory organs at the base of each siphon (Bone and Ryan 1978; Burighel et al. 2003). In trochid gastropods, cells with a long cilium and a collar of 9 thick microvilli are found in the epipodial sense organ (Crisp 1981). In leech, a putative hydrodynamic receptor called S-hair has a single cilium surrounded by a collar of 10 microvilli (Phillips and Friesen 1982). This cell type is located in the annular organ that has been shown to respond to near-field vibrations (Friesen 1981; Young et al. 1981). Similarly in the scallop *Mizhuopecten*, the abdominal sense organ has been shown to respond to water-borne vibrations (Zhadan and Semen'kov 1984). The putative hydrodynamic receptors in this organ have a single cilium and a collar of 9 microvilli (Zhadan et al. 2004). Direct confirmation at the single cell level awaits in these and other invertebrate systems.

Hydrodynamic sensory systems in planktonic organisms

Lower vibration frequencies in water than in air produce the same kinetic energy, thus even tiny short-range vibration sources such as locomotor cilia or small legs can become relevant. But also due to this, underwater vibration sensors can be smaller than a terrestrial counterpart and detect the same vibration amplitudes ((Tautz and Sandeman 1980; Markl 1973). Near-field hydrodynamic signals generated by small objects such as a passively sinking particle or by self-propelling small planktonic animals such as copepods or ciliated larva decay exponentially with distance (Visser 2001; Jiang and Kiørboe 2011). The range of detection for a small planktonic animal is estimated

to be the millimeter range (Martens et al. 2015), but the hydromechanical signals generated by self-propelling ciliated organisms can be considerable (Gallager 1993).

The hydrodynamic sensory systems in small planktonic animals have focused on copepods. The sensory receptor underlying the fast escape response in copepods was localized to the antennule, and more precisely to the setal mechanoreceptors (a sort of cuticular specializations) by using a tethered preparation and a simple stimulation setting (Gill 1985). Extracellular recordings later showed that these receptors are exquisitely sensitive to water movements (as small as 10 nm) and quite broad frequency reception (Fields and Yen 1992). The stimulus actually causing the escape response is thought to be velocity difference between the flow sensed at the setal mechanoreceptors and that sensed by the rest of the body.

Identification of mechanosensory neurons

Besides the research on copepod antennae mentioned above, few other studies have identified the hydrodynamic receptors driving planktonic behavior. This may partially stem from the difficulty in recording neuronal activity from the cells of such small animals. Modern neurobiology is increasingly relying on optical methods to assess neuronal activity (Yang and Yuste 2017), as it can be a less invasive, higher in throughput and more targeted method than electrophysiology. Together with the feasibility in invertebrates of identifying cells from animal to animal using morphological and molecular markers (Marder 2007), optical methods to record neuronal activity could allow to experimentally access putative sensory cells in small planktonic animals.

Calcium imaging with GCaMP

The most direct method to record neuronal activity in a cell is to measure the flow of ions and voltage changes using electrodes inserted into or placed next to the cell of interest. Changes of voltage in the cell can also be translated into an optical signal by using fluorophores sensitive to voltage (Peterka et al. 2011). An indirect method to measure neuronal activity is to image changes in intracellular calcium using a fluorescent reporter sensitive to such changes. Calcium imaging has become a powerful tool as the activity of larger ensembles of neurons can be recorded, and using genetically encoded calcium indicators (GECIs), specific cell types can be imaged (Grienberger and Konnerth 2012). Genetically-encoded voltage indicators currently lack a comparable signal to noise ratio to GECIs, and require high excitation levels that lead to photobleaching (Lin and Schnitzer 2016). Thus, GECIs have become the standard tool of choice to report neuronal activity *in vivo*.

A popular GECI is a protein construct called GCaMP. Originally described in 2001(Nakai et al. 2001), this protein is a fusion of cyclically permutated green fluorescent protein (cpGFP), the calcium-binding protein Calmodulin, and a peptide from the myosin light chain kinase M13. In low-calcium state, the conformation of GCaMP is such that the chromophore in the cpGFP is protonated, resulting in poor absorbance of photons. In high-calcium state, the Calmodulin domain

binds free Ca^{2+} and changes to a conformation that is able to bind the M13 peptide. This binding in turn changes the conformation of cpGFP to a closed state where the chromophore is protected from the outside aqueous environment, allowing it to absorb photons more efficiently at the excitation wavelength range characteristic of GFP (Akerboom et al. 2009; Wang et al. 2008). Different variants have been developed since the original study that vary in baseline fluorescence, kinetic response, brightness, fluorophore, etc.(e.g.(Chen et al. 2013; Dana et al. 2016)).

In *Platynereis* GCaMP has been successfully used to image neuronal (Chartier et al. 2018; Verasztó et al. 2017; Tosches et al. 2014; Verasztó et al. 2018) and muscle activity (Randel et al. 2014; Williams et al. 2015). In this chapter this tool will be used to physiologically identify the mechanosensory cells responding to the hydrodynamic stimulus eliciting the startle response in *Platynereis*. A morphological description is presented first that narrows down the possible candidate neurons to analyze.

Materials and Methods

Scanning Electron Microscopy (SEM) of nectochaete larvae

Nectochaete larvae to be processed for SEM were collected by phototaxis and rinsed several times in NSW. Unless otherwise indicated, nectochaete larvae were fixed in 2.5-3% glutaraldehyde/PBS for several days at 4°C. Samples were washed three times with PBS at RT to remove the fixative and incubated overnight at 4°C. The next day, samples were washed again twice in PBS at RT and incubated at 4°C for one hour. Samples were post-fixed with 1% OsO_4 for 2h on ice in the dark and washed thrice afterwards with PBS at RT. Each wash was done for more than 1h at RT. Samples were stored overnight in fresh PBS at 4°C. The larvae were then dehydrated in ethanol. A serial dilution of ethanol in PBS was used to minimize a drastic change in osmolarity: 30%, 60%, 75%, 95%, 100% ethanol washes were made; every wash was left overnight at 4°C. Animals were finally transferred to 100% Ethanol pre-adsorbed with a molecular sieve. Dehydrated samples were dried by critical point using a Polaron E3100. Samples were sputter coated with an 8 nm layer of platinum (Bal-Tec MED010) and imaged with a Hitachi S-800 field emission scanning electron microscope at an accelerating voltage of 15KV. All samples were processed and imaged by Jürgen Berger (Max Planck Institute for Developmental Biology, Tübingen). The different body regions were screened and at least two micrographs per structure was recorded. Some images were pseudo colored using Photoshop CS5 (Adobe) to highlight particular structures.

A fraction of the samples were fixed with Osmium tetroxide (OsO_4) to better preserve the ultrastructure of certain sensory cilia (Satir 1963). After collecting larvae into a tube, a few drops of 1% OsO_4 were added and tubes put on ice for 6h in the dark. Samples were washed out three times with distilled water and incubated overnight at 4°C. The following day, samples were washed again with water and then dehydrated with ethanol and processed further as indicated above.

Sensory neuron reconstruction in serial Transmission Electron Microscopy (sTEM)

Localization of penetrating ciliated sensory neurons

The location of penetrating cilia identified in SEM micrographs was used as a guide to find the corresponding sensory neurons in the electron microscopy volume previously reported (Randel et al. 2015). The collaborative annotation tool CATMAID (Schneider-Mizell et al. 2016; Saalfeld et al. 2009) was used to navigate and reconstruct these neurons. Penetrating ciliated neurons were identified by cilia crossing the cuticular layer. Some neurons not meeting this criterion were also annotated as penetrating ciliated cells if they fulfilled any of the following: 1) if they had a penetrating ciliated bilateral pair; 2) if the cilium could be detected in SEM micrographs; or 3) if they had a penetrating cilium projecting directly outwards and coming close to the cuticle. Sensory

endings in the left pygidial cirrus could not be imaged and thus sensory cell type identity could not be assigned in this region.

Identification of microvilli collars

Microvilli collars were defined as groups of small membrane profiles surrounding the cilium from its base and closely associated with it over a portion of its length. In some cases, the collar was easily identifiable, like in CR neurons, but in many cases the array was not clearly defined (due to the cutting plane, to loss of material or to the intrinsic structure). The number of microvilli per collar was not constant across the sections it spanned, but the maximum number observed was always registered (see **Table 0-6** in Appendix).

Neuron reconstruction

Penetrating ciliated cells identified in the EM volume were reconstructed to the extent possible and the main neurite projection and sensory dendrite fully reviewed using the Review widget in CATMAID. A custom-made nomenclature was followed to name the different cells (see **Table 0-6** in Appendix

Definition of segmental boundaries

Segmental boundaries were defined based on muscular anatomy, as these structures are segmentally repeated, and cover most of the larval trunk. A cell-level description of the musculature is part of an ongoing project and was used to define muscle subtypes (S. Jasek and G. Jékely, unpublished data). The animal was subdivided in prostomium or head, segment 0-3, and pygidium. Segment 0 is a cryptic segment giving rise to head tentacular cirri (Hempelmann 1911), and supported by molecular evidence (Steinmetz et al. 2011). The boundaries of segment 1-3 were drawn above the anterior oblique muscles and below the posterior oblique muscles (

Figure 2-1). The anterior boundary of the cryptic segment was set slightly below the prototroch, and its posterior boundary above the anterior oblique muscles to the oblique muscles. Even though only the oblique muscles were used to draw the boundaries, other muscle types, such as the ventrolateral longitudinal muscles and the transverse muscles, were also used as a reference to determine the segmental blocks. Some prominent cells with neurite projections in the parapodial nerve were also used as references to assess the boundaries of segmented blocks (

Figure 2-1).

Determination of other sensory cell properties

Measurement and imaging of cilia on the episphere

The median length of WT cilia as well as representative DIC images were obtained from wildtype larvae used for comparisons with mutants (see Materials and Methods in Chapter 4).

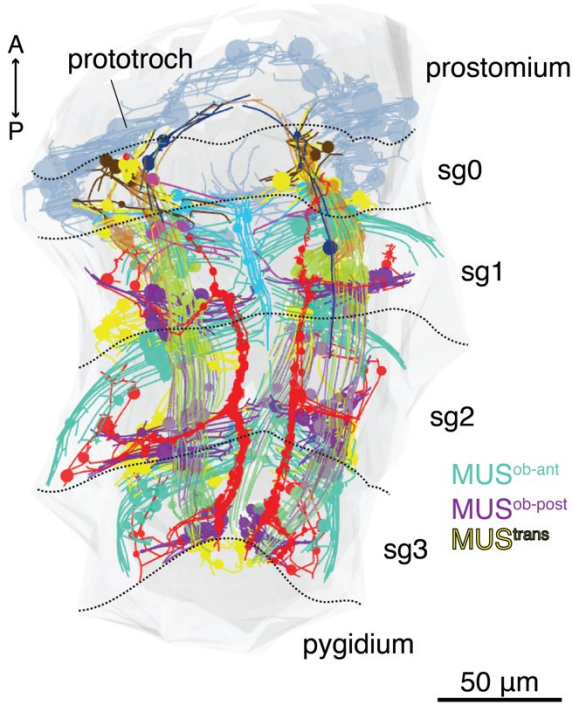


Figure 2-1 Definition of segmental boundaries in the electron microscopy volume. Electron microscopy volume reconstruction of several muscles of different classes. This snapshot was used to draw the boundaries shown in black dotted lines. Parapodial nerve cells are shown in red. Longitudinal muscles are shown and are colored by segment. Muscles were reconstructed and annotated primarily by Sanja Jasek. Other neurons shown were collaboratively reconstructed by members of the Jékely Lab.

MS ciliary beating frequency

Sample preparation

Single larvae were mounted using the same protocol indicated for measuring cilia length (see Materials and Methods in Chapter 4). In this case an inverted AxioObserver microscope (Carl Zeiss GmbH) equipped with a sCMOS camera (Axiocam 702 mono, Carl Zeiss GmbH) was used to record ciliary beating. Recordings were done using DIC microscopy. In some cases, the cilia were absent or shorter than normal. Only fully grown cilia were chosen for the analysis. To increase recording speed, the FOV was cropped to the region contain the cilium only, and the illumination was set to the maximum possible (12.2V) in order to decrease exposure time (0.6 ms). A frame rate of ~670 fps was achieved in this way. Around 6 sec-long recordings were acquired in each case. Two recordings for MS1 and two for MS2 were acquired.

Data analysis

Time-series stacks were rotated manually to orient the cilium vertically. The plugin FeatureJ:Edges (by Erik Meijering[5]) was run on the stack using default parameters (smoothing scale:1.0) to detect the edges of the cilium. An XZ orthogonal view was acquired at the base of the cilium in order to capture the lateral oscillation (**Figure 2-3**J-O). Snapshots of the level at which the X-time plane was taken were stored for reference. The beating frequency was manually calculated for 100 ms.

[5] Full plugin description at: https://imagescience.org/meijering/software/featurej/

Calcium imaging of hCRs and pygPB^{unp}

Plasmid construct for calcium imaging

To generate the construct used to synthetize *Palmi-3xHA-tdTomato-P2A-GCaMP6s* mRNA a PUC-57 plasmid with two ORF insertion sites was used to insert *GCaMP6s* (kindly provided by Douglas Kim, Addgene plasmid #40753) and *Palmi-3xHA tdTomato* (*tdTomato* kindly provided by Martin Bayer). The *tdTomato* ORF was flanked by AscI-AgeI restriction enzymes sites and the *GCaMP6s* ORF by NotI-PacI sites. The region between *tdTomato* and *GCaMP6s* was replaced by the sequence coding for the self-cleaving viral peptide P2A (Kim et al. 2011), using a standard oligo annealing procedure (see in the Appendix for protocol, and **Table 0-2** for P2A oligo sequences). Care was taken to maintain a common reading frame for all coding regions. The resulting translational fusion was inserted into the *Pdum*IVT plasmid, a previously reported plasmid used for mRNA synthesis for *Platynereis* work (Randel et al. 2014). A similar version of the plasmid using histone-tagged *GFP* (H2B-GFP) instead of *GCaMP6s* was used to verify efficient self-cleavage of the P2A peptide (not shown, in collaboration with Sara Mendes). The plasmid map and sequence are shown in the Appendix.

Synthesis and injection of GCaMP6s mRNA

The mRNA was synthesized and injected as described in the Appendix. The *Palmi-3xHA-tdTomato-P2A-GCaMP6s* mRNA was injected at a concentration of 2 μg/μl to increase the SNR of GCaMP6s.

Experimental setup

For calcium imaging, nectochaete larvae injected with *Palmi-3xHA-tdTomato-P2A-GCaMP6s* were tethered with Wormglu (GluStitch Inc, Canada) in the same way as for the kinematics experiments (see Materials and Methods in Chapter 1), but in a Ø5 cm glass-bottom dish like that used for the analysis of the startle response in freely-swimming larvae . The larvae were additionally anesthetized with 500 μM Mecamylamine (M9020, Sigma), a known nicotinic acetylcholine receptor antagonist that decreases muscle contractions in *Platynereis* larvae (Denes et al. 2007). This was required to minimize undesired changes in the Z-plane recorded. The anesthetic was left for ca. 30 min to take action before the actual experiment started.

For the experiments where GCaMP6s was quantified, the motor-controlled probe used for the kinematics experiments (Chapter 1) was placed at 40-60 μm from the anterior end or at 100 μm from the posterior end. This was needed to have both the cells of interest and the vibrating filament in the field of view, and thus be able to record the stimulus start. The stimulus start was determined from the bright-field channel as the first frame where a movement of the probe was evident (a movement artifact in the scanned frame was the telltale sign for stimulus start). The probe was controlled using the same Arduino hardware and script used for the kinematics experiments in

Appendix (see Arduino script `Motor-control-script.ino` in Appendix). Probe speed could not be calculated in these experiments due to the slow frame rate used for the recordings, but the motor ON delay (which indirectly dictates probe speed) was set to a value that reliably triggered a startle response in non-anaesthetized animals. Larvae were stimulated multiple times with resting periods longer than 1 min.

The experiments were performed under an LSM 780 NLO confocal microscope (Carl Zeiss GmbH) at RT. The 488 nm and 551 nm laser lines were simultaneously used to excite GCaMP6s and tdTomato. The recording rate for most videos was 4 Hz; 3 recordings were obtained at 2.6 Hz.

Data analysis

The recordings were processed in Fiji (Schindelin et al. 2012), First, each time-series file was shortened to include only stimulation events without drastic changes in the Z-plane. The time-series was then aligned by applying descriptor-based series registration (Preibisch et al. 2010) on the bright-field channel stack to correct shifts in the XY plane. The parameters thus obtained were used to align the green and red channels. Intensity values in both channels were measured in user-defined ROIs using a custom-written script (see `Fig3_FigS3_Measure-intensityvaluechanges.ijm` in Appendix). The signal from the membrane-localized tdTomato was used to draw the ROIs. A background intensity value for each channel using the same ROIs was also acquired with this script.

The intensity values were then fed into a custom-written R script (`Fig2_Analysis_CaimagingCR.R`). This script calculates the $\Delta R/R$ metric as defined by (Böhm et al. 2016), and plots the results (see Appendix for details). The metric was preferred over other metrics such as $\Delta F/F$ as it accounts for shifts in the Z plane, and the initial fluorescence values were easily defined. Recordings were relatively short, and thus a correction for bleaching artifacts was not needed.

Image acquisition

Higher resolution TEM micrographs were acquired from the same electron microscopy volume (by Reza Shahidi). Snapshots of the neuron reconstructions were acquired with CATMAID and scale bars added assuming the larva was 260 μm long. A smoothing gamma filter set at 6000 was used for snapshot acquisition to avoid displaying artefacts due to section misalignment. Plots showing $\Delta R/R$ plots were generated with R.

Results

A first step to investigate the underlying neuronal mechanism for the startle response is to identify the cellular sensors of the water disturbances that lead to the activation of the downstream circuit responsible for the startle behavior. In this chapter, these hydrodynamic receptors are identified in nectochaete larvae using anatomical and physiological information. In Chapter 3, the molecular analysis of these mechanosensory cells is reported.

Anatomical characterization of penetrating uniciliated neurons

Hydrodynamic receptors in other animals consist of structures protruding out to the environment, in form of modified cuticle (as in arthropods) or more commonly, as ciliary structures across the animal body that can be deflected and deformed relative to the body (Budelmann 1989). In annelids, and more specifically, in polychaetes, these sensory cells—the so-called penetrating ciliated cells—are classified in terms of the number of cilia and their sensory morphology (Purschke 2005). Following this classification scheme and taking advantage of the stereotypic development of *Platynereis* larvae (Fischer et al. 2010), penetrating ciliated sensory cells were mapped in nectochaete larvae using scanning and serial transmission electron microscopy (SEM and sTEM, respectively) data (Randel et al. 2015; Shahidi et al. 2015).

Penetrating ciliated sensory neurons in the episphere of *Platynereis* nectochaete larvae

The primary region to look for the putative hydrodynamic receptors mediating the startle response is the episphere of the nectochaete larva, as the animal was shown to be the more sensitive in this region than elsewhere tested (see Chapter 1).

Excluding the nuchal organ and crescent-shaped ciliated cells, which have been the subject of recent studies and are thought to be chemosensory cells (Verasztó et al. 2017; Shahidi et al. 2015), almost all penetrating ciliated sensory neurons in the episphere and parts of the prostomium have a single cilium (i.e. they are uniciliated). Most of these cells form bilaterally symmetrical pairs and are located in known nascent organs (such as the antennae or the head cirri), but also isolated or forming clusters (**Figure 2-2**A-D and **Figure 2-4**). In sTEM data, many uniciliated cells in the antennae had a cilium not entirely penetrating the cuticle. However, SEM data suggests at least some of them eventually will puncture the cuticle (see **Table 0-6** in Appendix for the full cell complement).

Collar receptor (CR) neurons

Most uniciliated sensory neurons in the episphere belong to the collar receptor (CR) neuron class (**Figure 2-2**A-F). This name was given to these cells due to their striking resemblance in sensory dendrite morphology to collar receptors in other polychaetes (Purschke et al. 2017; Schlawny et al.

1991; Windoffer and Westheide 1988; Purschke 2005), as well as to the S hairs in leeches (Phillips and Friesen 1982) and earthworms (Knapp and Mill 1971). CRs are located in the frontal tip of the episphere, and in the nascent antennae (**Figure 2-2**A,Ab,C-D). They are also found on the dorsal side of the episphere, specifically forming part of the dorsal sensory organ (DSO; **Figure 2-4**).

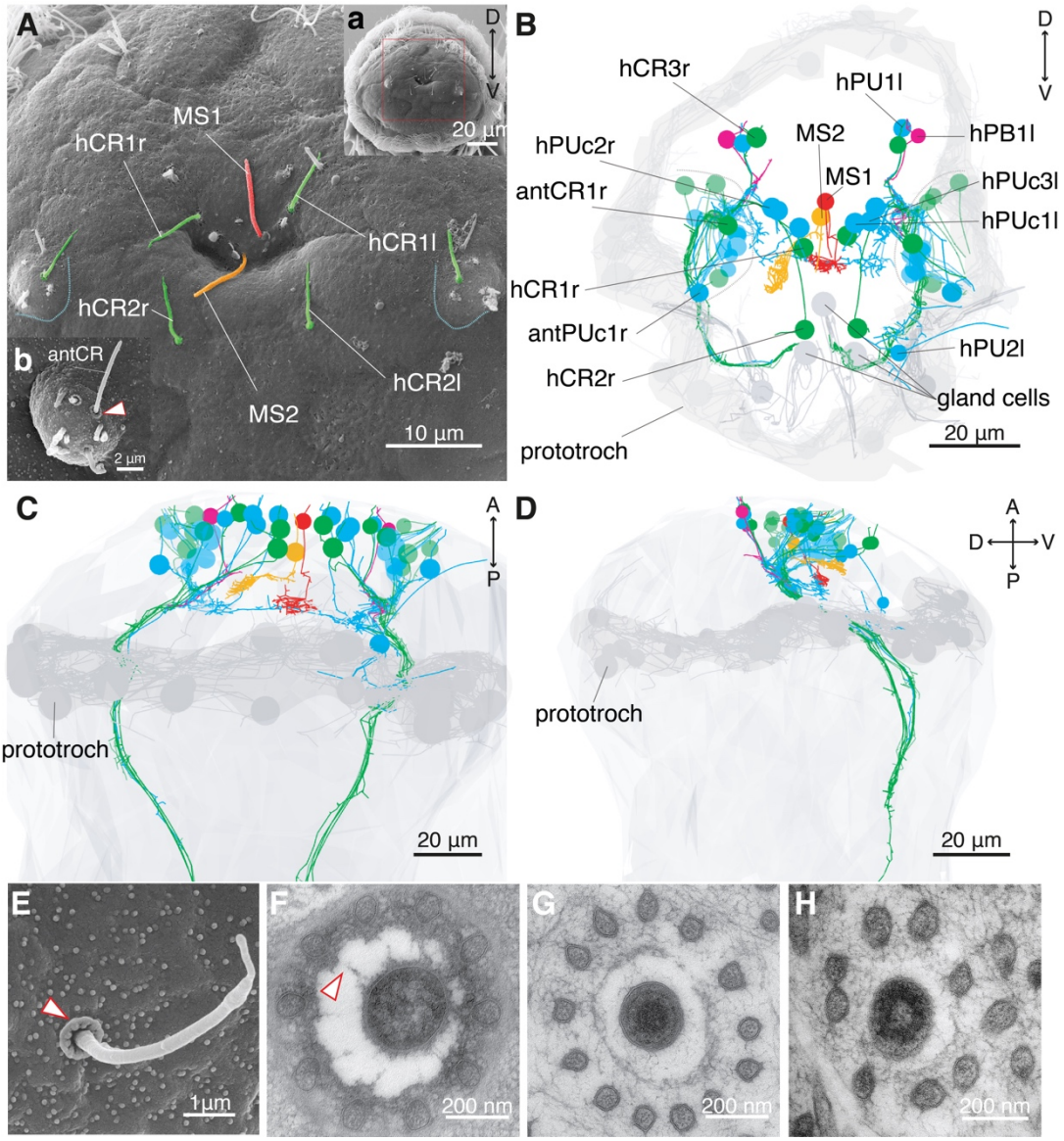


Figure 2-2 Penetrating ciliated sensory neurons in the episphere of nectochaete larvae. (**A, Aa**) SEM micrographs of nectochaete *Platynereis* larva, apical view. (**Aa**) Overview of nectochaete larva viewed from the anterior. Close up shown in A was taken from the central part of this specimen (indicated by dashed red rectangle). (**A**) SEM micrograph of nectochaete *Platynereis* larva showing a close up view of most apical group of cilia in the episphere. Colors and labels match those in B-D. (**Ab**) Detailed view of an antenna of a different larva. Red arrowhead points to the collar of a putative antCR. (**B**-**D**) ssTEM reconstruction of penetrating ciliated neurons found in the episphere. CR neurons are colored in green, PU/PUc neurons in blue and MS neurons in tones of red. PB cells are colored in magenta. Cells with fully penetrating cilia are displayed in solid colors while cells with cilia not fully crossing the cuticle in the TEM data (but otherwise annotated as penetrating ciliated cells) are shown in semi-transparent colors. Prototroch and head glands are shown for spatial landmarks. For clarity, only one of the bilateral pairs of fully penetrating ciliated cells is labeled. Apical (B), ventral (C) or lateral (D) views. (**E**) SEM micrograph of a CR neuron cilium. Arrowhead points to the microvilli collar covered by a thin cuticular layer. (**F**) TEM micrograph showing the typical CR collar morphology. Arrowhead points to fibril connecting a microvillus with the cilium. (**G**) TEM micrograph showing MS1 collar morphology. (**H**) TEM micrograph showing hPUc1 collar morphology. PU: penetrating uniciliated, PB: penetrating biciliated. Sample in A was fixed with 1% OsO_4. Antennal regions are demarcated by dashed lines in A and B.

Like collar receptors in other annelids, CRs have a long non-motile cilium (median length: 8 μm and 8.8 μm for hCR1 and hCR2, respectively) with a 9x2+2 axoneme (**Figure 2-2**F). The cilium is surrounded by a symmetric (or slightly oval) collar of 10 thick microvilli that entirely crosses the cuticle (**Figure 2-2**E, F).

As a consequence, microvilli collars are also visible using SEM, appearing as a cuticular bulging at the base of the cilium, and in some cases the 10 microvilli could be identified in these images (arrowhead in **Figure 2-2**E). Interestingly, CR neurons show filamentous connections between cilium and microvilli (arrowhead in **Figure 2-2**F). In live specimens, head CR neurons appear to be ensheathed by a thin pellicle, but this is not visible (or not present) in all specimens analyzed (**Figure 2-3**G). The appearance of this sheath may be an artefact caused by the hypotonic conditions in the medium (Short and Tamm 1991).

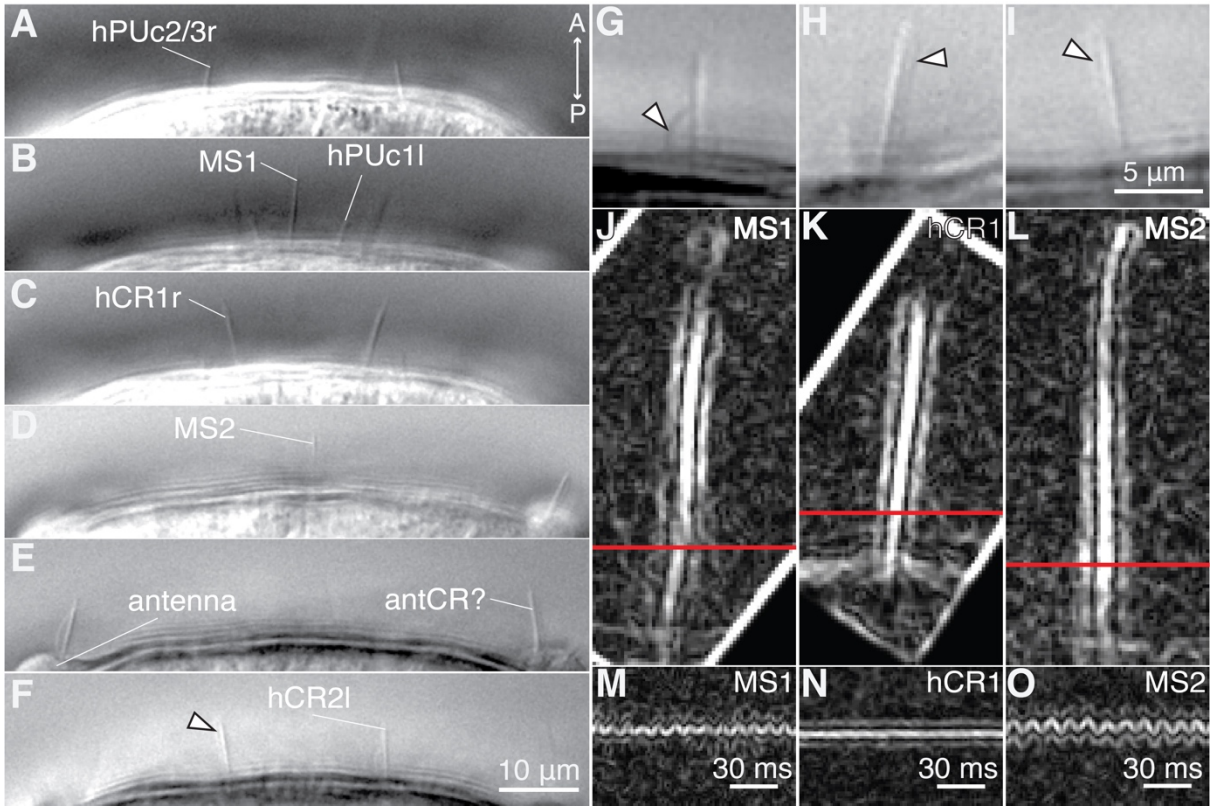


Figure 2-3 Features of episphere cilia as seen using DIC microscopy. (A-F) Optical sections highlighting the different penetrating sensory cilia found across the central episphere of a single nectochaete larva. Images are arranged from dorsal to ventral. See **Figure 2-2**A for a spatial reference. (A) On the dorsal side of the central episphere, a pair of short cilia is visible, belonging to either hPUc2 or hPUc3 cells. (B) In the very centre of the episphere, the long cilium of MS1 is visible. To either side the much shorter cilia of hPUc1 cells is also visible. (C) The hCR1 cilia are located slightly more ventral. (D) The tip of the MS2 cilium is visible. As the cilium has a tilted orientation it was not possible to see the complete structure in focus. (E) Long cilia on each of the nascent antenna are likely to belong to one of the CRs in these sensory organs. (F) The most ventral cilia belong to hCR2 neurons. (G-H) Close-up images showing examples of the bulging observed in the ciliary membrane of hCR cells (arrowheads).(J-L)Images of MS1(J), hCR1 (K), and MS2 (L) processed as detailed in the Material and Methods. Red line indicates the site at which the X-time plane shown in M-O was taken in each case. (**M-O**) X-time projections of the cilium base of MS1 (J), hCR1 (K), and MS2 (L).

MS neurons

A second group of uniciliated collared cells found in the episphere are the MechanoSensor (MS) neurons. MS neurons have a cilium of similar length to that seen in CR neurons (median length: 9.1 μm and 8.7 μm for MS1 and MS2, respectively). Each cilium has 12-13 thin microvilli regularly arranged around it (**Figure 2-2**G). In these cells, the collar of microvilli does not fully cross the cuticle (see **Table 0-6** in Appendix). The cilium of MS neurons also has a 9x2+2 axoneme structure (**Figure 2-2**G). Interestingly, fast-frame recordings of MS1 and MS2 cilia revealed an intrinsic motility with a low-amplitude frequency of ~100Hz (**Figure 2-3**J, L,M,O). In comparison, the cilium of hCR1 was completely immotile in recordings acquired with the same method (**Figure 2-3**K, N). The cilia of MS3 and MS4 cells are also motile (data not shown) but due to their position and orientation, the beating frequency could not be calculated. MS neurons have an asymmetric location and do not form part of either the antennae or the DSO (**Figure 2-2**A-D and **Figure 2-4**).

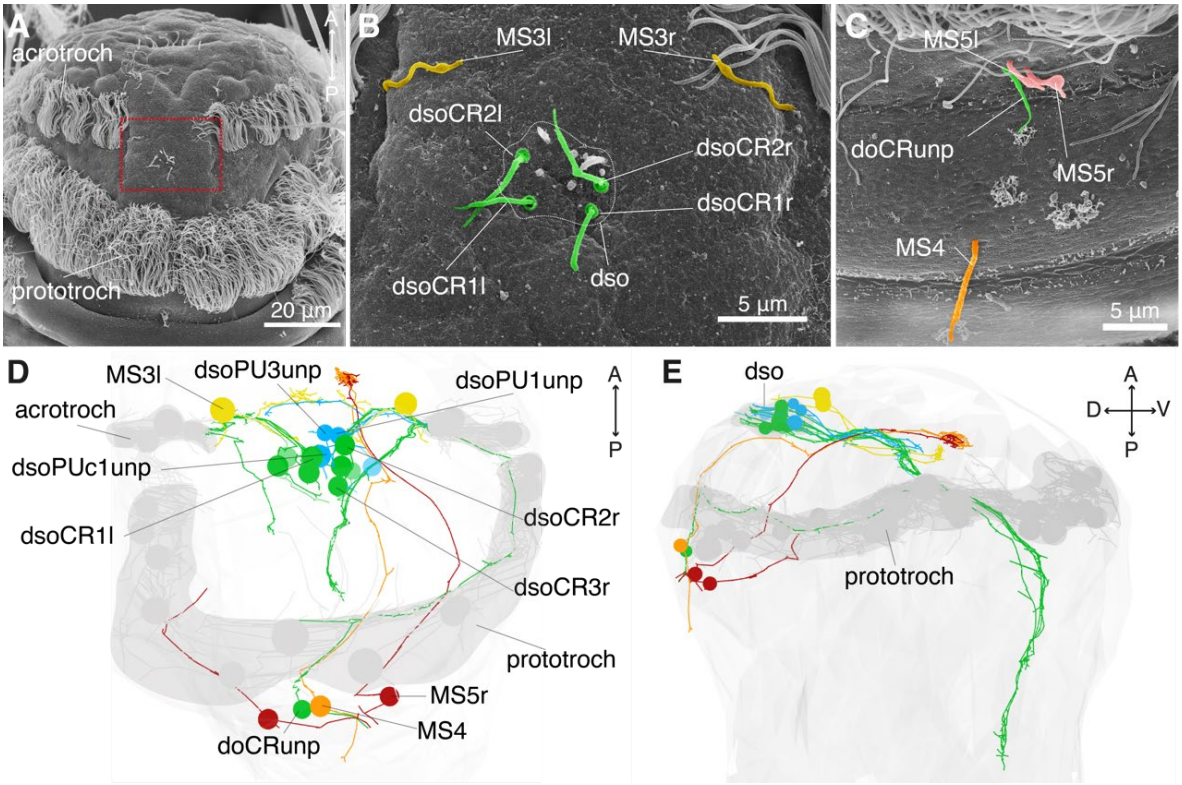


Figure 2-4 Penetrating ciliated sensory neurons in the dorsal episphere of nectochaete larvae. (**A**) SEM micrograph showing an overview of cilia on the dorsal side of the larva. Dashed red rectangle highlights blown up region shown in B. (**B**-**C**) SEM micrograph highlighting penetrating cilia between the acrotrochs and the prototroch (B), or posteriorly adjacent to the prototroch (C). Colors and labels match those in D-E. (**D**-**E**) ssTEM reconstruction of penetrating ciliated neurons found in the dorsal side of the head. Only cells with fully penetrating cilia are displayed. Prototroch and acrotrochs are shown as spatial landmarks. CR neurons are colored in green, PU neurons in blue and MS neurons in tones of red. Cells with fully penetrating cilia are displayed in solid colors while cells with cilia not fully crossing the cuticle in the TEM data are shown in semi-transparent colors. Prototroch and acrotrochs are shown as spatial landmarks. For clarity only one of the bilateral pairs of fully penetrating cells is labeled. Dorsal (A-D), or lateral (E) view. Samples in A-C were fixed with 1% OsO_4. See Appendix for the complete list of cells.

hPUc neurons

A third group of uniciliated collared cells was recognized solely on the basis of more heterogeneous features and may in reality comprise a mix of cell types. These cells were called head penetrating uniciliated collared (hPUc) neurons. In general, these cells have shorter cilia than those of either the MS or CR neurons. The collars of these cells have between 7 to 9 slender and often branching microvilli that do not cross the cuticle (**Figure 2-2**H and **Figure 2-3**A,B). Three bilateral pairs of hPUc cells were found in the center of the episphere. A number of PUc neurons were found in the antennae and DSO (antPUc, and dsoPUc cells), but the cilium of most of these cells did not cross the cuticle in the sTEM data (faded out in **Figure 2-2**A-D and **Figure 2-4**).

PU/PB neurons

The remainder of the penetrating ciliated cells that were found in the episphere was composed of penetrating uniciliated cells without a collar (PU neurons), and of a single pair of biciliated neurons (hPB,1 cells) in the dorsal side of the episphere. These set of cells were far less numerous than CRs and MS cells, and their cilia were shorter, thus not showing any additional feature that could indicate a hydrodynamic sensory modality (**Figure 2-2**B-D).

In summary, based on sensory dendrite morphology, the episphere of the *Platynereis* nectochaete larva is endowed with three main types of penetrating uniciliated neurons: CR, MS, and PUc neurons, among which the first two have long cilia and thus represent good candidates to detect hydrodynamic disturbances.

Numerous penetrating ciliated sensory neurons populate trunk and pygidium of nectochaete larvae

Penetrating ciliated sensory neurons were also mapped to the trunk and to the organs in this region following the same criteria as for the episphere (**Figure 2-5**). Most of these cells were either uniciliated or biciliated, with some cilia having a collar of microvilli. No additional morphological specializations were evident among these cells with the exceptions noted below. These cells were more easily distinguished by their anatomical position, as they were present in the ventral and dorsal sides of the trunk, in the parapodia, in the glands and in the pygidium of the larva. Many of these cells did not entirely cross the cuticle in the TEM data, but evidence from SEM micrographs indicate the cilia of at least a subset of these cells do eventually puncture the cuticle. A brief description of the spatially different cell groups will be given below (see **Table 0-6** in Appendix).

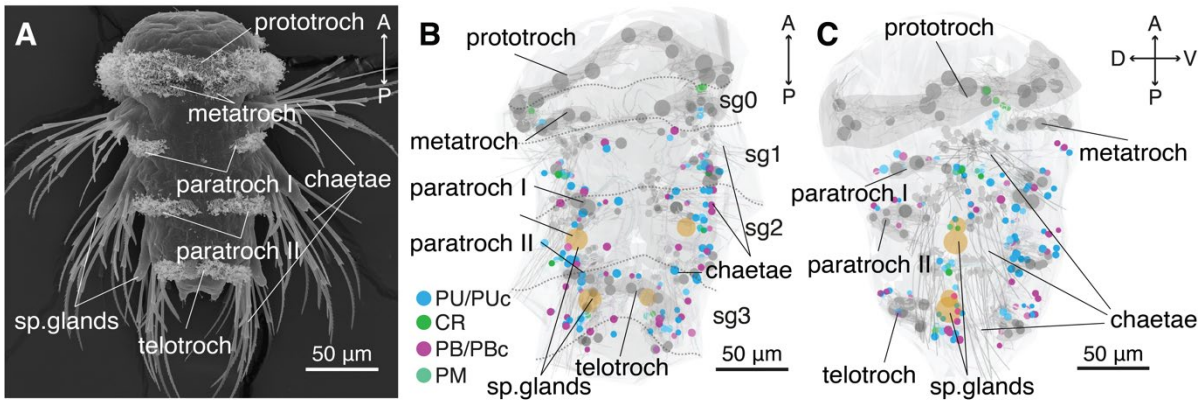


Figure 2-5 Diverse types of penetrating ciliated sensory cells line trunk and parapodia of the nectochaete larva. (**A**) SEM micrograph showing a nectochaete larva. Major ciliary bands and morphological features are indicated. Ventral view. Sp. glands: spinning glands. (**B**-**C**) Penetrating ciliated sensory neurons found below the prototroch and above the telotroch mapped in the ssTEM stack. Only cell body positions are highlighted (see full morphology in **Figure 2-6** and **Figure 2-7**). Cells with fully penetrating cilia are displayed in solid colors while cells with cilia not fully crossing the cuticle in the TEM data are shown in semi-transparent colors. Ventral (B), or lateral (C) view. PU, PB, and PM stand for penetrating uniciliated, biciliated and multiciliated neurons, respectively. Segment boundaries are marked with dashed lines.

Ventral and dorsal trunk penetrating ciliated sensory neurons

Several penetrating uniciliated (PU) and biciliated (PB) sensory cells were found lining the ventral and dorsal side of the larval trunk (**Figure 2-5**). The majority of these cells form pairs of bilaterally symmetrical biciliated and uniciliated neurons located immediately above each paratroch and telotroch on both the dorsal and ventral sides (**Figure 2-6**). All of the dorsal penetrating ciliated sensory cells belong to this class (namedventralParaPU/PB or dorsalParaPU/PB; see **Table 0-6** in Appendix).

A few biciliated cells are rather located in the center of the trunk with no association to the paratrochs (called ventraltrunkPU/PB) (**Figure 2-6**A-D). Most of the cells in the ventral and dorsal sides of the trunk do not have a collar of microvilli around the cilium, and the few that have them are not of the CR or MS cell type (see **Table 0-6** in Appendix). In addition, the SEM images reveal that most of these cilia are rather short. Therefore, these cells probably belong to a different cell type to those found in the head of the larva.

Parapodial penetrating ciliated sensory neurons

Small sensory tufts of penetrating ciliated sensory neurons were found by SEM and serial TEM in the area between the notopodia and neuropodia on both sides of each segment (prefix: interparpod, **Figure 2-7**A). Four types of cells gave rise these ciliary tufts: a single CR neuron (recognized by the dendrite morphology), two uniciliated cells of the PU type, either with or without a microvilli collar, and a penetrating multiciliated (PM) neuron. PM neurons have either 3 or 4 cilia (if located in the first or in the second segment, respectively) and are only found in this region of the larva. PM cells in the third segment have multiple basal bodies, but not fully formed cilia (**Figure 2-7**B-C).

A class of single collared uniciliated (PUc) neuron was found to be associated to each of the parapodia in the animal. Some of them had a 10 microvilli collar but otherwise not similar to the CR collar. I called these cells notopodPUc or neuropodPUc to reflect the parapodia to which they were associated with.

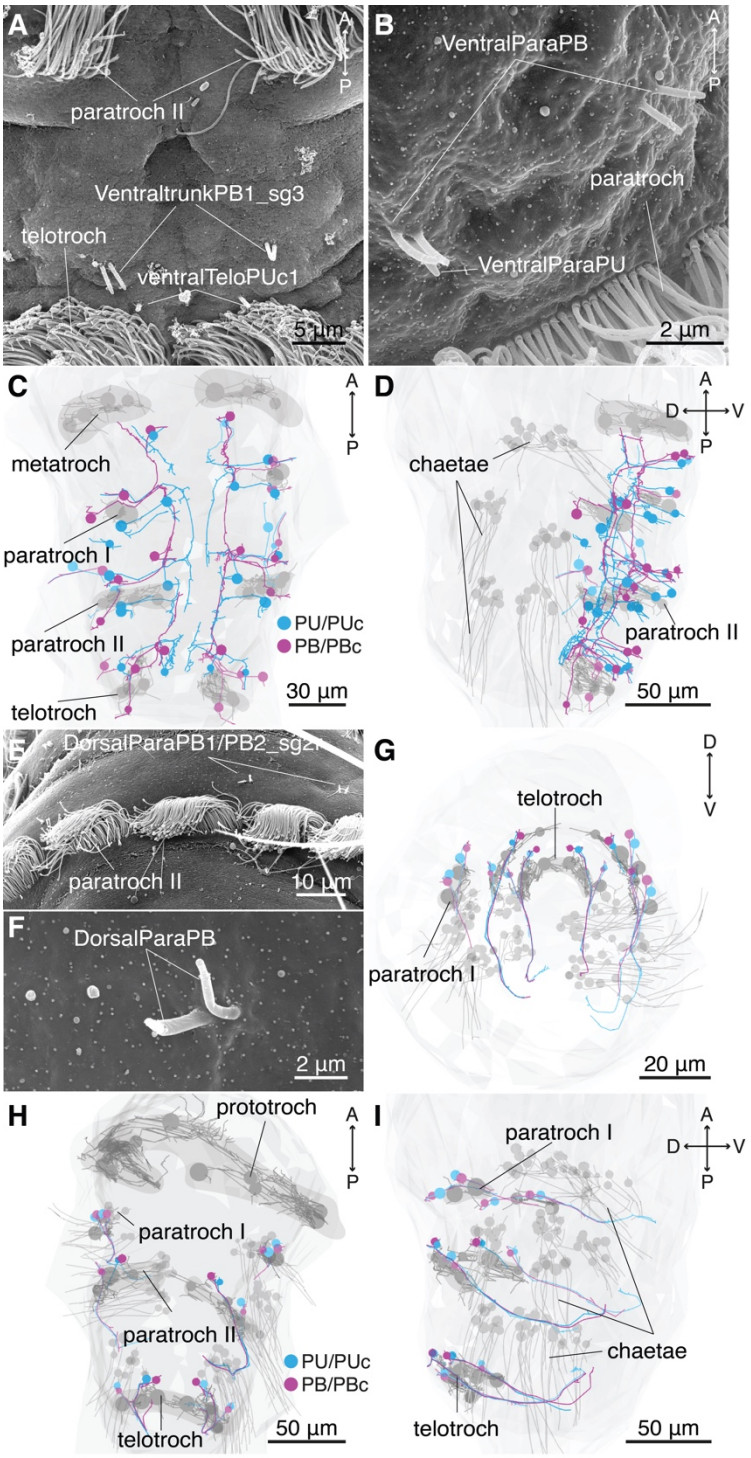


Figure 2-6 Bilateral pairs of penetrating uniciliated and biciliated sensory neurons found in ventral and dorsal trunk of the nectochaete larva. (**A**-**B**) SEM micrographs showing examples of cilia found on the surface of the ventral side of the trunk. Note their proximity to the trunk ciliary bands (paratrochs and telotroch). (**C**-**D**) ssTEM reconstruction of penetrating uniciliated (PU/PUc) and biciliated (PB/PBc) ciliated cells in the ventral side of the trunk, ventral (C) or lateral (D) view. (**E**-**F**) SEM micrographs showing examples of cilia found on the surface of the dorsal side of the trunk. (**G**-**I**) ssTEM reconstruction of uniciliated (PU/PUc) and biciliated (PB/PBc) penetrating ciliated cells in the dorsal side of the trunk, apical (G), dorsal (H) or lateral (I) view.

Organ-associated penetrating ciliated sensory neurons

Spinning glands

Additional penetrating ciliated sensory neurons were found in the side organs of the larva, including the glands, and the cirri primordia (**Figure 2-7**B-F). Each of the spinning glands in the second and third chaetiger segments had 3 biciliated (PB) and 3 uniciliated (PU) cells some of them endowed with thin microvilli collar that did not cross the cuticle (see **Table 0-6** in Appendix). The long cilia of these cells can be seen by SEM, but they cannot be assigned to individual cells. Some PB neurons in this organ had what appear to be a collar of cilia and are in that respect different to the PB cells found in the trunk (see **Table 0-6** in Appendix).

Head and segmental cirri

The nascent cirri in the region immediately below the prototroch, the head cirri, have numerous immature sensory neurons (data not shown). Some of these cells belong to the CR class, while the remaining 2 cells are uniciliated PUc cells, but do not fully cross the cuticle (**Figure 2-7**B-C). Likewise, cirri on the ventral side in each segment are just starting to emerge at this stage. Only a single PB neuron in each cirrus was found in this organ (see **Table 0-6** in Appendix).

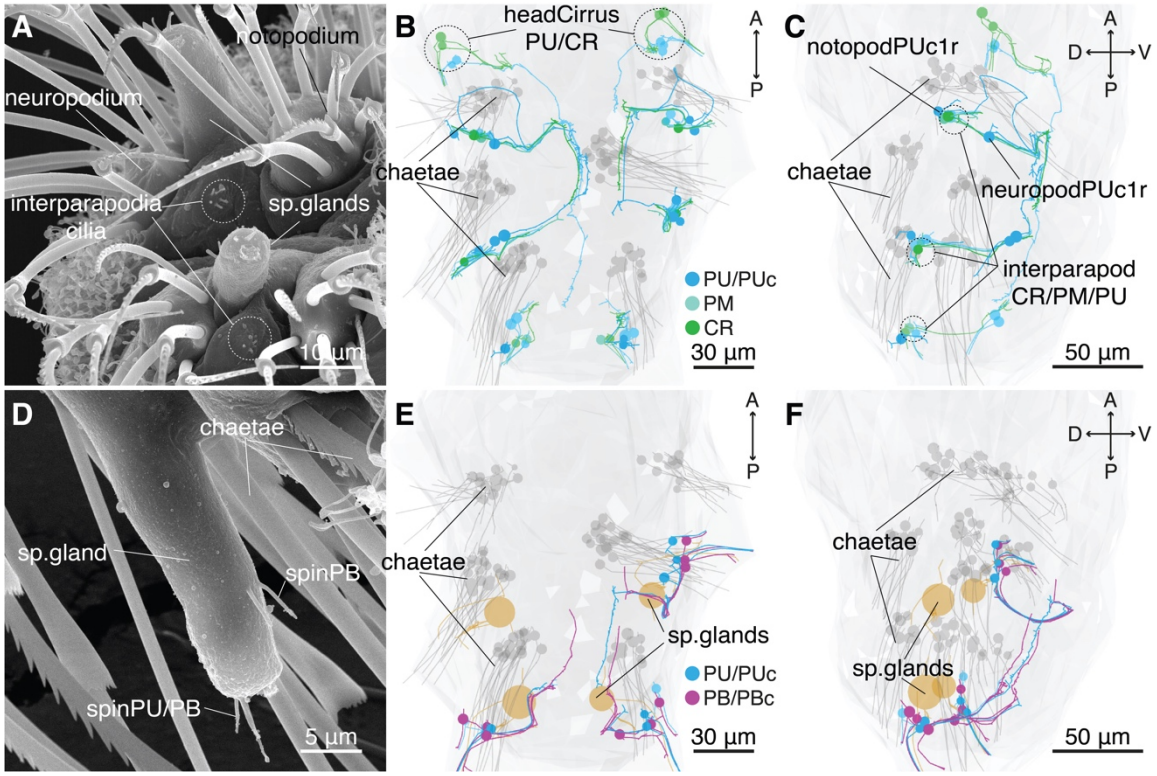


Figure 2-7 Multiple types of penetrating ciliated sensory neurons in the parapodial region and in the spinning glands. (**A**) SEM micrograph showing cilia of penetrating ciliated neurons in the interparapodial region. Tilted view from posterior. (**B**-**C**) ssTEM reconstruction of penetrating uniciliated (PU/PUc/CR) and multiciliated (PM) neurons found above, below or between the notopodia and neuropodia. Ventral (A) or lateral (B) view. (**D**) SEM micrograph showing penetrating cilia (from either PB or PU cells) on a spinning gland. (**E**-**F**) ssTEM reconstruction of penetrating uniciliated (PU/PUc/CR) and biciliated (PB/PBc) neurons found in the spinning glands (sp. Glands) in the second and third trunk segments. Cells on the second segment gland on the right body side were not reconstructed due to the loss of the gland tip.

Penetrating ciliated neurons in the pygidium of the nectochaete larva

The tail or pygidium has a big proportion of the ciliated cells found in the larva, most of them forming part of the nascent pygidial cirri (**Figure 2-8**A-D). The right side pygidial cirrus is formed by several PU cells, three of which (pygCirrus_MSPUc1r-c3r) have actually a collar with a similar of microvilli as that found in MS neurons (see **Table 0-6** in Appendix). The pygidial cirrus also contained a number of CR neurons, and a single biciliated collared cell (**Figure 2-8**A-B).

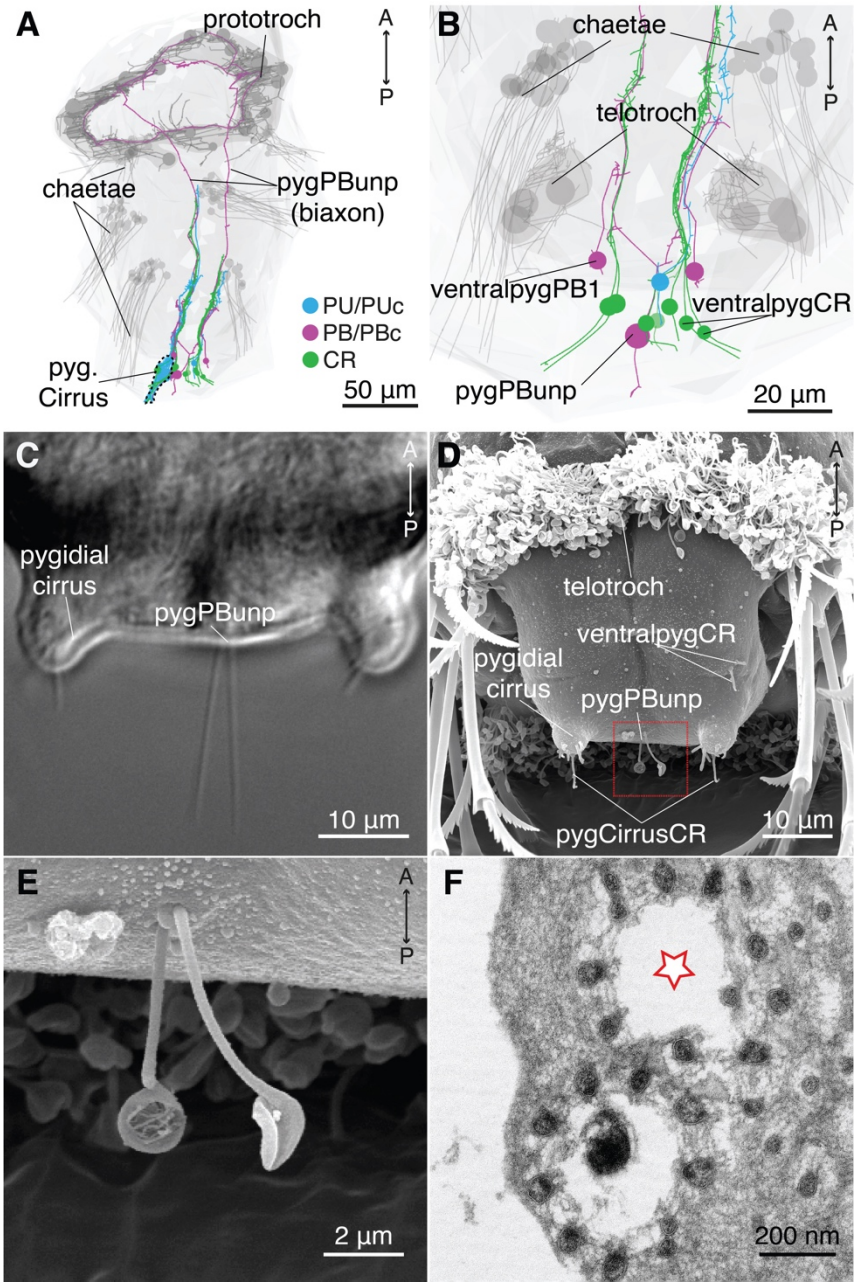


Figure 2-8 Penetrating ciliated neurons in the pygidium and nascent pygidial cirri of nectochaete larvae. (**A-B**) ssTEM reconstruction of uniciliated and biciliated neurons located in the pygidium and in the nascent pygidial cirri. Sensory ending of cells on the left cirrus were lost from the stack and are thus not included. The cell innervating prototroch is pygPB^{unp}. For clarity pygidial cirrus cells are not shown in B. (**C**) DIC microscopy image showing the pygidium of a living nectochaete larva. (**D**) SEM micrograph showing close up of pygidium at the nectochaete stage. (**E**) Blown-up image of region in D (red dashed square) showing pygPB^{unp} cilia. The 'paddle' cilia morphology is a fixation artifact (Short and Tamm 1991). (**F**) TEM micrograph of pygPB^{unp} sensory morphology showing microvilli

collar around each of the two cilia. One of the cilia was lost from this section (red star is at its expected location). Ventral views in A-E.

Additional CR neurons were found on the ventral pygidium as well as another bilateral pair of PB cells (**Figure 2-8**B). Finally, a giant biaxonal neuron previously reported (pygPB^{bicil} ,(Shahidi et al. 2015; Verasztó et al. 2017)) had two long non-motile cilia protruding out between each nascent pygidial cirrus (**Figure 2-8**C-E). This cell was renamed to pygPB^{unp} to keep a consistent nomenclature. Each cilium of pygPB^{unp} is on average 21.9 μm (4 measurements) long and has a collar of 10 microvilli (**Figure 2-8**F), and together with CR neurons, it is the only cell type found in the nectochaete larva to have a collar of microvilli clearly crossing the cuticle (see **Table 0-6** in Appendix).

In summary, an anatomical approach alone revealed multiple types of penetrating ciliated cells, mainly with one or two cilia (see **Table 0-6** in Appendix). Some of these cells had a collar of microvilli around the cilia. Chiefly among these collared cells was the CR neuron cell type; this class was more numerous in the episphere and pygidium of the nectochaete larvae and had characteristic collar morphology already described in many other annelid taxa. The collared-cilium morphology is widespread among hydrodynamic receptors (Budelmann 1989), thus making this group of cells an especially attractive candidate for mediating the startle response.

Testing CRs as hydrodynamic receptors by calcium imaging

The localization of candidate hydrodynamic receptors in *Platynereis* is a necessary first step towards understanding how the larva detects and eventually reacts to water disturbances. However, a definitive step to determine if at least a subset of the sensory cells found reacts to water-borne vibrations, is to measure their neuronal activity upon stimulation. To tackle this challenge, mRNA encoding the calcium indicator GCaMP6s (Chen et al. 2013) was injected into *Platynereis* eggs and the changes in GCaMP fluorescence in selected penetrating ciliated cells upon stimulation analyzed at the nectochaete stage. Calcium indicators provide an indirect measure of neuronal activity with good signal-to-noise ratio, albeit with lack in temporal resolution (Grienberger and Konnerth 2012; Lin and Schnitzer 2016).

CR neurons are hydrodynamic receptors

As previously mentioned, the head is likely to contain at least some of the receptors involved in the startle response. Among the putative hydrodynamic receptors in this region, CR and the MS neurons have all the characteristics for being sensitive to water disturbances, namely, a long cilium, and a collar of microvilli. Therefore, these cells were analyzed in GCaMP-injected larva for evidence of activation upon hydrodynamic disturbances. Larvae were tethered with glue, in a similar fashion as performed for the kinematics experiments detailed in Chapter 1. To avoid

movement artifacts recordings were performed on animals treated with the nicotinic receptor antagonist mecamylamine (see Materials and Methods).

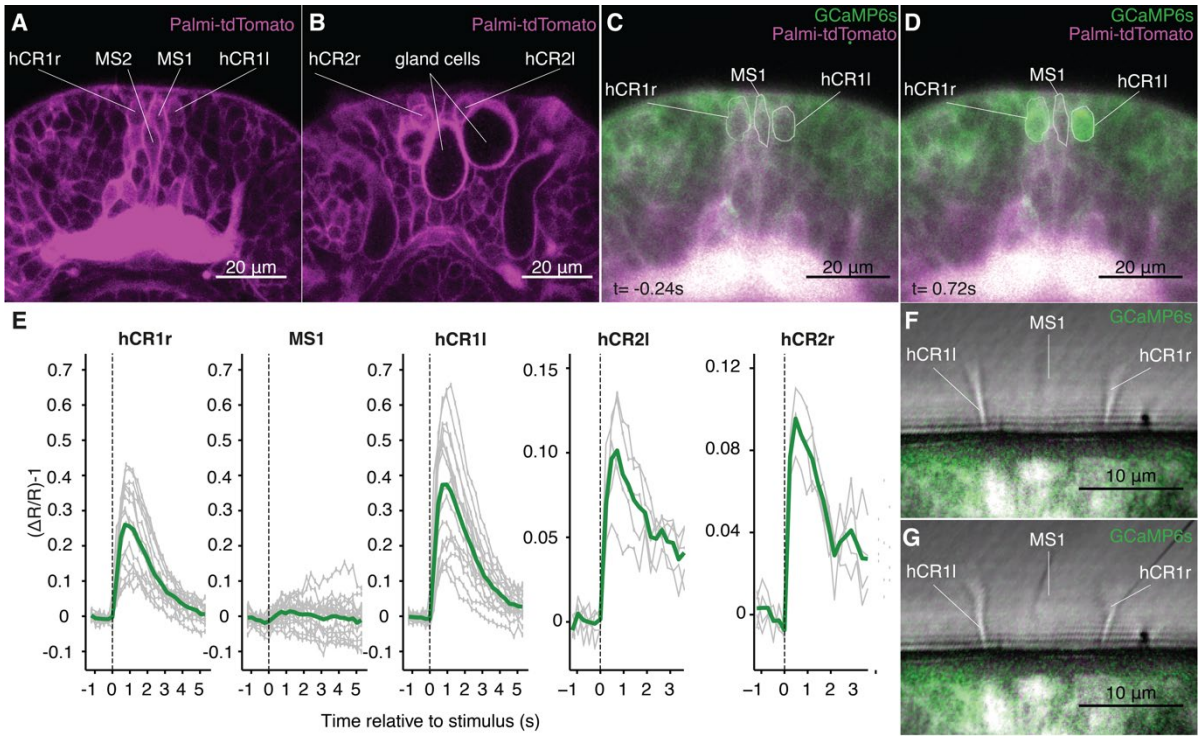


Figure 2-9 hCR neurons, but not MS1, respond to hydrodynamic disturbances. (**A**-**D**) Larva used for calcium imaging expressing Palmi-tdTomato-P2A-GCaMP6s. (**A**-**B**) Palmi-tdTomato fluorescence is localized to the membrane (magenta) and allows to identify (A) hCR1, MS1, MS2 and (B) hCR2 neurons. (**C**-**D**) Merged GCaMP6s and Palmi-tdTomato signals. ROIs used for fluorescence measurements (hCR1s and MS1) are indicated. (**C**) Snapshot taken 0.24 seconds before stimulus. (**D**) Snapshot 0.75 seconds after start of stimulus; note increase of GCaMP6s fluorescence in hCR1 neurons. (**E**) Quantification of GCaMP6s fluorescence in somata of hCR1s, MS1 and hCR2s before and after stimulus (stimulus start t=0, dashed line). The ΔR/R metric was used to correct for motion artifacts. Green line is the mean of individual measurements (gray dotted lines). Number of measurements in E: hCR1r:15, MS1:18 hCRl:16 (MS1 activity was always measured together with either of the hCR1s, 5 larvae were tested), and hCR2l: 5, hCR2r: 4, 2 larvae tested. (**F**-**G**) Closeup view of larval episphere showing the hCR1 and MS1 cilia (F) before (top panel) or (G) immediately after stimulus start (bottom panel). Note the deformation of all three cilia.

The group of CR and MS neurons located at the frontal end of the episphere were particularly suited for this experiment. First, the cell bodies of the head CR neurons hCR1, and of the MS neuron MS1 lie adjacent to each other and thus give the possibility to image them in the same plane and perpendicularly to their long axis (**Figure 2-9**A). Moreover, these cells have unique cell body morphologies: hCRs have a round shape and MS1 an elongated shape, and this allows their reliable identification in living animals injected with a membrane-tagged fluorescence reporter (**Figure 2-9**A). And finally, the cilia of these cells can be confidently and repeatedly identified (**Figure 2-3**). Upon stimulation with a vibrating filament placed at 50 to 60 µm from the head, the fluorescence increased in left and right hCR1, but not in MS1 neurons (**Figure 2-9**C-E).The stimulus deflected both MS1 and CR1 cilia, and thus the lack of signal in MS1 is not due to a lack of stimulation (**Figure 2-9**F-G). A minute increase in MS1 may be actually due to signal from the adjacent hCR1 cells.

The slightly more ventral hCR2 cell pair, which is readily identifiable as their cell bodies juxtapose the head gland cells (**Figure 2-9**B) showed a weak but reproducible increase in fluorescence upon vibrational stimulation (**Figure 2-9**E). These results suggest that at least head CR neurons are hydrodynamic receptors. More importantly, since the same stimulus that triggered the startle response was sufficient to activate these sensory cells, they may mediate the activation of the downstream circuit controlling the response.

PygPB^{unp} responses to water-borne vibrations

The same experiment as detailed in the previous section was attempted on penetrating ciliated cells in the posterior end. However, the only cell that could be unambiguously identified in this region with the membrane-tagged reporter was pygPB^{unp}, the giant biciliated cell at the center of the pygidium (**Figure 2-10**A-B). The stimulus used was stronger in order to match the increased threshold needed to elicit the startle response from the posterior. A weak, but reproducible response in three animals to a strong vibration stimulus was observed in this cell (**Figure 2-10**C). The mean response had the same dynamics as that seen for hCRs. If these results are confirmed in future experiments, it would mean that pygPB^{unp} is a second cell type tuned to hydrodynamic stimuli, and possibly involved in the startle response.

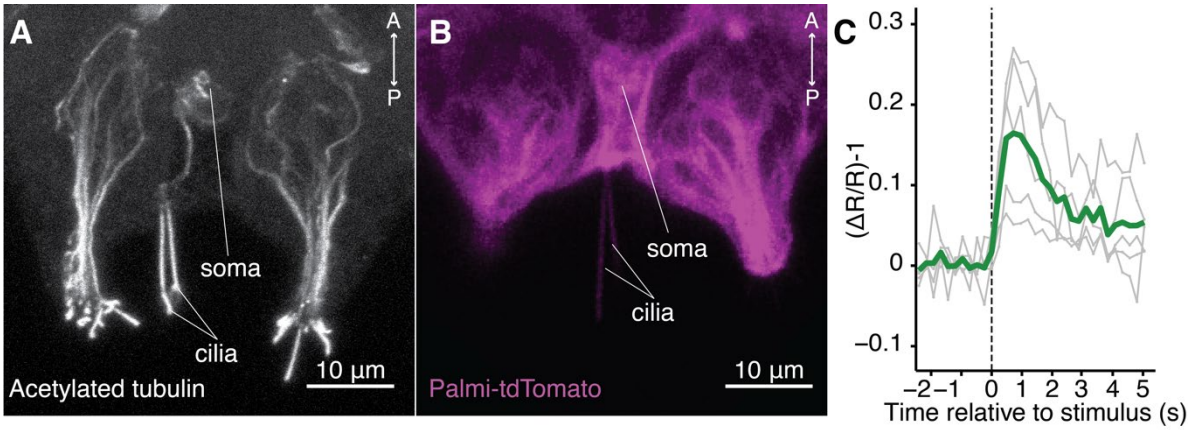


Figure 2-10 The giant penetrating biciliated sensory/ciliomotor neuron pygPB^{unp} responds to hydrodynamic disturbances. (**A**) pygPB^{unp} is recognizable in immunostainings against acetylated tubulin by the prominent soma and long cilia in the middle of the pygidium. (**B**) The large soma and long cilia of pygPB^{unp} are also distinguishable in living larva expressing Palmi-tdTomato-P2A-GCaMP6s mRNA (only the Palmi-tdTomato signal is shown). (**C**) Quantification of GCaMP6s fluorescence in pygPB^{unp} soma before and after stimulus (stimulus started at t=0, dashed line). The ΔR/R metric was used to correct for motion artifacts. Green line is the mean of individual measurements (gray dotted lines). 5 measurements from 3 larvae. Dorsal view in A and B.

Discussion

Diversity of penetrating ciliated neurons

A diverse set of penetrating ciliated primary sensory neurons was found in the nectochaete larva, showing variation in the number of cilia, as well as in the structure of the sensory dendrites including the length of the cilium and the presence or absence of collars of microvilli. Ultrastructural studies carried out across the animal phylogeny have similarly revealed different types of penetrating ciliated cells (Westfall et al. 1998; Bedini et al. 1973; Lacalli and Hou 1999; Eakin and Kuda 1971; Cantell et al. 1982; Schlawny et al. 1991). For illustration, one recent study found in the dorsal cirri of the polychaete *Eurythoe* two types of penetrating uniciliated collared cells with long cilia, one type of non-collared uniciliated cell, and one type of multiciliated cell (Purschke et al. 2017). A whole-body serial EM reconstruction in the dwarf male of the Polychaete *Dinophilus* likewise identified collared uniciliated and non-collared uniciliated cells and multiciliated cells (Windoffer and Westheide 1988). Similar to those studies, the work presented here leverages the stereotypy of the *Platynereis* larva to generate a full body map of penetrating ciliated neurons that could be used as a reference for future developmental, anatomical and functional studies.

The location of at least some of the ciliated neurons may already give a hint about their function. For instance, many of the ciliated neurons were associated with nascent organs such as the antennae, or cirri. These sensory appendices have been recently shown to respond to chemicals at 6 days post-fertilization (Chartier et al. 2018), and thus at least a subset of these cells may be important for this sensory modality. Uniciliated and biciliated cells were also found in other structures such as those at the tip of the spinning glands in the second and third segments. Sensory cells have been found associated with secretory cells in other polychaetes (Schlawny et al. 1991). The function of the spinning glands is not known (despite what the name suggests), but the ciliated cells on them may help sense mechanical or chemical stimuli that regulate gland secretion. Such functional association between sensation and secretion has been proposed for similar structures found in mollusks (Owen and McCrae 1979; Bickell-Page 1991). Many uniciliated and biciliated ciliated cells were located slightly above or below the paratroch. Whether this association has a functional connection is not clear. Other planktonic larvae have ciliated sensory cells tightly associated to the ciliary band (von Dassow et al. 2013; Lacalli 1982; Lacalli et al. 1990). Their function is not totally understood.

The neurite projections of most of the penetrating ciliated neurons could be reconstructed, which tells us more about the putative function of the neurons. Such is the case of the biciliated cell type pygPB^{unp} located at the tip of the pygidium. This neuron has two neurites projecting upward along the VNC until reaching the prototroch, where it innervates non-ciliated and ciliated cells of the

prototroch . Due to this innervation pattern, pygPB^{unp} is thought to be part of the ciliomotor system regulating ciliary beating (Verasztó et al. 2017). All MS neurons projected to the brain plexus region, and at least MS1 and MS2 synapse onto the head ventral motoneurons and interneurons in the phototaxis circuit (Verasztó et al. 2018; Randel et al. 2014), thus suggesting a role in this behavior. In contrast, CRs in the anterior and in the posterior side projected along the VNC without major branching events. As in the case of MS cells, the common projection pattern of CRs adds support to their unity as a cell type.

Collared ciliated neurons

CR neurons

A subset of the ciliated cells identified in the larva had a collar of microvilli around each individual cilium. Difference in the number, and arrangement of collar microvilli combined with the cilium length allowed to distinguish between different types. One of these subtypes, the CR neurons, showed striking similarity to collared receptors in other annelids. For instance, among the two types of collar receptors found on *Eurythoe* cirri (Purschke et al. 2017), one of them had a long cilium surrounded by exactly ten thick microvilli, as it is the case for all CRs in the *Platynereis* larva. In *Dinophilus*, a uniciliated cell type found at the posterior end of the animal also has a long cilium and a collar of 10 microvilli (Windoffer and Westheide 1988). Additional ultrastructural similarities between these ciliated cells include the presence of electron dense material on the inner side of each microvillus, and a collar crossing the cuticle. In the Eunicid polychaete *Ophryotrocha* these features as well as the fibrous material connecting the microvilli and cilium was identified (Schlawny et al. 1991). Not only polychaetes, but also other more distantly related annelids in the Clitellata and Hirudinea groups have collared cells sharing the same sensory dendrite features of CR neurons (Phillips and Friesen 1982; Knapp and Mill 1971; Moritz and Storch 1971). Whether such ultrastructural similarities are due to a common ancestor cell type is an open question.

MS neurons

A second type of well- defined collared cell was the MS cell type. In contrast to the numerous reports of CR-like neurons in annelids, few reports of cells with some similarity MS neurons could be found. For instance, male *Dinophilus* have in addition to the CR-like collar cell, a second collared cell type in the anterior end of the animal with 15 microvilli arranged in a funnel shape (Windoffer and Westheide 1988). Likewise, *Ophryotrocha* has a second collar receptor type with "several" microvilli arranged in a funnel shape (Schlawny et al. 1991). These cells have only in common with MS cells the larger number of microvilli in their collars. Moreover, the motility of MS cilia—a probably important feature for their function—cannot be contrasted to these and other studies focusing on fixed specimens. Thus MS neurons seem to be a previously unreported anatomical

type of penetrating uniciliated cell not figuring in major reviews of the sense organs in polychaetes (e.g. (Purschke 2005)).

Of note, the motility of MS cilia is unlike that seen for the rest of the motile sensory cilia in the animal such as the nuchal organs, or those in the crescent cell—which have a beating pattern with power and recovery stroke phases (Schmidtberg and Dorresteijn 2010)—and a mechanism of control more similar to those of the ciliary bands (Verasztó et al. 2017). The MS cilia beating pattern is symmetric, which indicates that the beating does not have as purpose to mobilize fluid around it (Purcell 1976). That said, the beating pattern will need to be characterized in freely behaving animals to rule out potential artifacts, and to analyze other parameters, such as whether MS1 and MS2 beating in or out of phase. The high beating frequency of MS cilia needs also further confirmation, as other cilia beat at much lower frequencies. For instance, in the Kupffer vesicle cilia beat at around 30-40Hz. Whether such fast oscillations have a role in MS sensory function, and whether these cells are tuned to mechanical cues are open questions that will be discussed further in Chapter 3.

Other uniciliated collared neurons

Numerous uniciliated cells in the *Platynereis* nectochaete were found to have microvilli collars (grouped under the term PUc), which suggests that the use of this trait alone is not specific enough to define mechanosensory cells. Fewer and thinner microvilli and in a less ordered arrangement than those found in CRs were common traits of the collars found in PUc cells. A second type of collar receptor cell reported by Purschke *et al* is similar to the general PUc type reported in this study. Based on their location and on their rather immature morphology (based on the incomplete penetration of cilia and short neurite projections), some of the PUc cells may become part of sensory organs not present or hardly distinguishable at this stage such as dorsal and ventral cirri, ligulae, etc.

pygPB^{unp} neurons

The pygPB^{unp} is one of the few collared biciliated neurons found in the nectochaete larva. Some reports in other invertebrate groups exist of collar receptors with more than one cilium (Jouin et al. 1985; Cantell et al. 1982), but none of those is clearly similar to that seen in pygPB^{unp}. Nonetheless, developmental and anatomical studies in other annelid larvae suggest that this cell type might be conserved.

For instance, in the trochophore larva of the closely related polychaete *Phyllodoce maculate,* an almost identical cell (called *sp1*) has been reported (Voronezhskaya et al. 2003; Nezlin and Voronezhskaya 2017). The *sp1* neuron is a biciliated neuron with a biaxonal morphology located in the tip of the pygidium, all features shared with pygPB^{unp}. In both *Platynereis* and *Phyllodoce* these neurons stained positive for serotonin (Starunov et al. 2017; Voronezhskaya et al. 2003; Verasztó et al. 2017). A

posterior biaxonal serotonergic-positive cell was also found in the larva of the sedentary polychaete *Pomatoceros lamarckii* (McDougall et al. 2006). It is unclear if this cell is sensory ciliated cell as ultrastructural analyzes of this region have not been carried out in this species. However, the *Pomatoceros triqueter* trochophore larva was described to have two long cilia at the posterior end (Segrove 1941), thus leaving open the possibility that the cell found by McDougall et al is biciliated as well. Depictions of other polychaete larvae show 2 cilia at the posterior end (Okuda 1946; Lacalli 1984), which suggests that pygPB^{unp}-like cells are conserved, at least in this group of annelids.

Note on multiciliated neurons

Multiciliated cells were the least numerous in the larva, only located bilaterally between notopodia and neuropodia in each segment. Mutliciliated neurons are also quite common in other annelids (Jouin et al. 1985; Knapp and Mill 1971; Purschke et al. 2017; Schlawny et al. 1991), often occurring in sensory tufts mixed with uniciliated cells in other animals (Crisp 1981; Purschke et al. 2017; Knapp and Mill 1971; Phillips and Friesen 1982), as it was the case in *Platynereis*. The rather undeveloped morphology towards the posterior segments suggests that this cell type is still developing and may have a major function only at a later stage. The absence of collars of microvilli may indicate these multiciliated cells function as a type of chemoreceptor (Jouin et al. 1985).

CR neurons as confirmed hydrodynamic receptors

Collar receptors found in many invertebrate groups such as Hemichordates (Nørrevang 1964), Priapulids (Moritz and Storch 1971), Nemerteans (Montalvo et al. 1996), Cnidarians (Lyons 1973), Platyhelminthes (Ehlers and Ehlers 1977), and in almost any major group of animals have been hypothesized to be hydrodynamic receptors of some sort. The same is true for the collar receptors in Annelids, and indirect evidence in the leech suggests that CR-like neurons (called S-hairs) are sensitive to near-field vibrations (Friesen 1981; Young et al. 1981).

The easy identification of head CRs in the alive *Platynereis* larva made possible to find a temporal correlation between the onset of a hydrodynamic stimulus and an increase in GCaMP6s signal in these cells. Such correlation more directly suggests that collar cells in general, and head CRs in particular are some sort of near-field hydrodynamic receptors. Calcium imaging has been used in other systems as a tool for assessing mechanosensitivity of sensory cells (Kindt et al. 2012; Zhang et al. 2016). However, it will be ultimately required to record mechanoelectrical currents in the CRs to directly prove that mechanosensation is occurring at these cells (Ranade et al. 2015), as opposed to secondary activation.

Are CRs the sensory cells initiating the startle response?

One important implication of the calcium imaging experiments in CRs is that the stimulus that activated these cells was also sufficient to trigger the startle response. These results thus support the hypothesis that CRs detect disturbances in the water and trigger the startle response. Additional

work will be needed to prove the causal link between the activation of these cells and the initiation of the response. Genetic-based manipulation of neuronal function such as optogenetics or thermogenetics (Luo et al. 2008) would be ideal methods to show that acute activation or inactivation of CRs lead to the induction or abolishment of the response, respectively. To date, only genetically-mediated cell ablation has been used with reported success in *Platynereis* (Veedin-Rajan et al. 2013), thus urging the need for adapting other cell inactivation tools to this system.

Mechanism of CR activation

Despite the lack of electrophysiological evidence, it is already worth discussing the mechanism of activation of CR neurons. The most straightforward mechanism to couple the hydrodynamic signal to the cell response would be through the deflection of the CR cilium in a magnitude proportional to the strength of the hydrodynamic stimulus. In support of this is the fact that the stimulus used was strong enough to cause the deflection of the hCR1 cilia. In kidney tubular cells, such deflection is thought to be required to initiate the calcium-dependent mechanotransduction cascade activating these cells (Nauli et al. 2003; Praetorius and Spring 2001; Singla and Reiter 2006). The source of mechanosensitivity of unicellular organisms such as *Chlamydomonas* and *Paramecium* has been suggested to reside also in the cilia (Ogura and Takahashi 1976; Yoshimura 1996).

Although conceivable, it is not granted that cilia are the direct sensors of the mechanical stimulus in CRs. Lack of cilia does not completely inhibit the mechanically driven activation of kidney tubular cells (Rodat-Despoix et al. 2013). In hair cells the site of mechanotransduction is in the microvilli bundle and not in the cilium, the latter thought to only transmit force to the former (Hudspeth and Jacobs 1979). Moreover, no correlation of calcium influx and mechanical stimulation was seen in the cilium of either cell type when signal was recorded at higher temporal resolution (Delling et al. 2016). Thus, the increase in GCaMP6s fluorescence in the CR cytoplasm is not necessarily a consequence of an earlier rise in calcium at the cilium. In *Paramecium*, deciliation does not abolish mechanoreceptor currents, which may be actually due to channels localized to the cell membrane (Machemer and Ogura 1979; Ogura and Machemer 1980) . It is perhaps surprising that despite their prevalence, few other ciliated mechanosensory cells besides those mentioned above have been studied in detail. The open question about the cilium as mechanosensory organelle could be addressed using the CR cilium as a model.

Additional structural features informing the mechanism of CR activation

The bendability of the cilium may not be enough to explain the mechanism of CR activation. Notably, the cilium of the MS1 neuron also was bent, but it did not get activated, thus suggesting that further specific structural features of CRs play a role in their sensitivity to hydrodynamic disturbances.

The collar of microvilli

The collar of microvilli was a defining feature of CR neurons, always in the same number and arranged in a symmetric pattern. Thus, as in vertebrate hair cells (Hudspeth 1989), CR microvilli could actually be the locus of mechanotransduction and their structure, key to this cells' function. In vertebrate hair cells, their directional sensitivity is provided by the difference in length and ladder-like arrangement of microvilli. Thus, the apparently similar length and symmetric arrangement of CR microvilli would suggest by comparison that these cells do not have an intrinsic directional sensitivity. The regular arrangement of the microvilli may actually ensure that the tension needed for mechanical gating increases in all directions perpendicular to the sensory dendrite, perhaps by contact of the cilium with the microvilli or by some deflection of the microvilli themselves. As previously mentioned, the regular arrangement of CRs is conserved in annelids, but it has been also noted in reports in other marine invertebrates (Lacalli 1982; Moritz and Storch 1971; Ehlers and Ehlers 1977).

Fibrous connections

Another feature that may be important for CR function is the set of fibrils found connecting microvilli and the cilium. It is tempting to speculate they have an analogous function to kinociliary and tip links seen in vertebrate hair cells (Kindt et al. 2012). These links have also been observed in confirmed and in putative hydrodynamic receptors in Cnidarians (Watson et al. 1997), annelids (Schlawny et al. 1991), and in bivalve Mollusks (Zhadan et al. 2004).

Future work regarding CR mechanosensation

Additional questions remain regarding CR mechanosensation. As previously discussed for the behavioral experiments (see Chapter 1), the hydrodynamic stimulus used lacked precision. It is thus not possible to know what parameter of the stimulus activates the CRs. Preliminary experiments using sustained laminar flow to deflect the CR cilium did not lead to an increase in GCaMP activity (data not shown). Thus, the cells must be tuned to a particular type of hydrodynamic stimulus, for instance to the rate of increase in flow rate. The CR cilia seem to be responsive to deformations of the flow created by the ciliary bands. It is not clear if they need this flow to better detect the hydrodynamic signal, or if in the absence of any ciliary band flow they would still be equally sensitive. In fish, the sensitivity of the lateral line in fish decreases when it is swimming partly due to the increase in hydrodynamic noise, but also to active efferent inhibition (Russell and Roberts 1974). Inhibiting ciliary beating while stimulating the CRs could address any change in sensitivity in the CR activation.

Hydrodynamic receptors in the pygidium

The larvae display the startle response to both anterior and posterior hydrodynamic stimulations (see Chapter 1). The calcium imaging suggests head CRs may be sensing signals from the anterior.

Although it is reasonable to assume that CRs in the pygidium will also be sensitive to vibrations, this still has to be shown.

pygPB^{unp} response to calcium

The preliminary results shown here indicated that the biciliated cell pygPB^{unp} is also sensitive to vibrations. The mechanosensitivity of these cells is not totally unexpected as they have the longest non-motile cilia in the larva, and each has a collar of microvilli. If its mechanosensitivity is confirmed, it would suggest that the startle response is mediated by a different set of cells depending on the site of stimulation: by CRs in the anterior and by CRs *and* pygPB^{unp} in the posterior. This could explain the difference in threshold and overall response profiles seen between anterior and posterior stimulation. Preliminary experiments showed detailed larvae, but not decapitated larvae are still able to get startled (data not shown). This suggests that the essential receptor must be at least present in the episphere region.

Behavioral responses are in many cases encoded by multiple populations of mechanosensory cells, encoding each different parameters of the stimulus (Maksimovic et al. 2014; Patella and Wilson 2018; Ohyama et al. 2015; Pirschel and Kretzberg 2016). Thus, CRs and pygPB^{unp} may act in synergy to properly shape the response to posterior stimulation. As serotonin is thought to be an antagonist of ciliary arrests and increase ciliary beating (Verasztó et al. 2017; Doran et al. 2004), the expression of serotonin in pygPB^{unp} suggests that this cell may promote ciliary beating upon posterior mechanical stimulation by releasing serotonin at the synapses with MC and with the prototroch itself (Verasztó et al. 2017). This effect may help explaining the observed increase stimulus threshold needed to elicit ciliary arrests when the larva is stimulated closer to the pygidium

.

Chapter 3 Mechanotransduction channels expressed in collar receptors and in other sensory neurons

"Molecular evolution of mechanosensation cannot be understood the same way rhodopsin evolution contributes to understanding vision"
-Fritzsch et al[6]

[6] (Fritzsch et al. 2007)

Statement of contributions and publication status

Sara Mendes performed most of the immunostainings for *PKD1-1* and *PKD2-1* shown in this chapter, as well as some of the cloning steps leading to the promoter constructs. The SEM micrographs were acquired by Jürgen Berger. The phylogenetic trees, the gene promoters, WMISH data and remaining immunostainings were acquired by the author of this study. The figures were likewise assembled by the author.

The expression patterns of *PKD1-1* and *PKD2-1* as well as the PKD phylogenetic trees were published elsewhere in abbreviated form (Bezares-Calderón et al. 2018). The plasmid construct was used in a previous publication (Verasztó et al. 2017).

Introduction

The molecular basis of neurosensory mechanotransduction

Neurosensory mechanotransduction is the process by which a sensory cell transforms mechanical forces from the environment into electrical signals that can be used by the nervous system (Delmas and Coste 2013; Chalfie 2009). The speed of such transformation in mechanosensory cells is so fast, that the transformation is thought to be mediated by a molecular complex directly sensing the mechanical force and gating the flux of ions through the cell as function of such force (Corey and Hudspeth 1983; Walker et al. 2000; O'Hagan et al. 2005). Such mechanoelectrical transducer (MeT) channel has evolved a molecular mechanism to detect a specific feature of the mechanical force, either by a deformation of the membrane in which it is embedded, or by a pull of a molecule tethered to it (Kung 2005). The exact molecular mechanism is yet to be elucidated.

An unexpectedly diverse set of molecules have shown all the features of *bona fide* MeT channels (Ernstrom and Chalfie 2002). These include members in the Degenerin/ Epithelial sodium /Acid Sensing Channel (DEG/ENaC/ASIC) superfamily, in the TRPN family of TRP channels, in the Piezo family, in the Transmembrane Channel-like (TMC) protein family, and in the stretch-activated channels in the K2P potassium channel family.

The first gene in animals to be identified as a *bona fide* MeT channel was MEC-4, a member of the DEG/ENaC/ASIC superfamily. MEC-4 was originally identified in a screen for gentle touch insensitive mutants (Chalfie and Au 1989). This channel is expressed in *C.elegans* touch receptor neurons (TRNs) and it is required there to preserve the mechanoreceptor currents *in vivo*, and to gate sodium ions (O'Hagan et al. 2005). The related genes DEGT-1 and MEC-10 are required for harsh nose touch in multidendritic neurons (Chatzigeorgiou and Schafer 2011; Chatzigeorgiou et al. 2010), but it also forms heteromeric channels with MEC-4 in TRNs (Chen et al. 2015). DEG-1 is another DEG/ENaC channel expressed in ciliated nociceptor neurons in *C.elegans* and it is responsible for most of the mechanoelectrical current seen in these cells (Geffeney et al. 2011). In *Drosophila,* the DEG/ENaC homologs Pickpocket and Balboa are expressed in nociceptor neurons, where they are specifically required for detecting mechanical noxious stimuli (Zhong et al. 2010; Mauthner et al. 2014). Mice lacking ASIC3, a member of the DEG/ENaC channels, show defects in proprioceptive tasks, and the DRG neurons expressing it show altered muscle spindle sensitivity mammals (Lin et al. 2016). This gene also may have a role in mechanical nociception in the skin free-nerve endings (Price et al. 2001). ASIC2 in mouse is expressed in aortic baroreceptor neurons and contributes to the baroreflex in the circulatory system (Lu et al. 2009). Definitive evidence for the role of DEG/ENaC as MeT channels in either *Drosophila* or mammals is lacking.

Another protein family found only in animals is TRPN (also called NOMPC in invertebrates) (Peng et al. 2015). This channel has been shown to be responsible for mechanically activated currents in both *Drosophila* and *C.elegans* (Yan et al. 2013; Kang et al. 2010). The ortholog in *Drosophila* is expressed in ciliated mechanosensory organs, including proprioceptors, sound-receiving neurons, and labellar sensilla, where it is required for locomotion, hearing, and food texture detection, respectively (Cheng et al. 2010; Liang et al. 2011; Effertz et al. 2011; Walker et al. 2000; Lehnert et al. 2013; Sánchez-Alcañiz et al. 2017). It is also expressed in non-ciliated multidendritic touch-sensitive cells in the larva, where it is required for responses to gentle touch (Yan et al. 2013). *NOMPC* confers mechanosensitivity to other neurons in the fly that do not normally express it and forms a mechanically gated channel when expressed in heterologous systems (Yan et al. 2013). The ortholog in *C.elegans*, TRP-4, is expressed in food-contacting ciliated mechanosensory neurons (Sawin et al. 2000), and it is partially required for nose touch avoidance behaviors (Chatzigeorgiou and Schafer 2011). Since TRPN is absent in amniotes (Schüler et al. 2015), its function in vertebrates is less studied, but it has been found expressed in hair cells where it is important for auditory function (Sidi et al. 2003).

Piezo is a recently described channel with 24 to 36 transmembrane domains in animals, plants and some protists (Coste et al. 2010). In vertebrates and *Drosophila*, this gene is expressed in neural and non-neural tissues, and when expressed in culture cells both the invertebrate and vertebrate homologs form channels and induce mechanically activated currents (Coste et al. 2010; Coste et al. 2012). In *Drosophila* larva, the single Piezo homolog is required for driving mechanically-activated currents in isolated nociceptive neurons (Kim et al. 2012).In vertebrates, mostly mammals, Piezo channels have been involved in numerous functions related to mechanosensation. The Piezo2 channel is responsible for the rapidly adapting currents recorded in *in vitro* skin preparations from DRG neurons, which together with its role in Merkel cell mechanotransduction, shows this channel is a main contributor to non-noxious touch sensation (Ranade et al. 2014; Woo et al. 2014). Piezo2 is also required for stretch-induced firing of proprioceptors innervating muscle spindles, which is needed for proper coordination of body movements and posture (Woo et al. 2015). And both Piezo1 and 2 are required for the baroreflex in the circulatory system (Zeng et al. 2018). In fish, Piezo2b is expressed in Rohon-Beard neurons, where it is needed for a response to light touch (Faucherre et al. 2013).

Another group of molecules where evidence of a function as a MeT channel is mounting is the TMC-A subfamily in the TMC superfamily (Keresztes et al. 2003). TMC1 and TMC2 belong to this subfamily and are expressed in hair cells in the mammalian ear (Kurima et al. 2002). Mutations in both genes lead to the loss of mechanotransduction currents in hair cells (Pan et al. 2013; Kawashima et al. 2011). TMCs are also needed for stimuli-evoked microphonic potentials (i.e.

electrical potentials recorded from the hair cell population) in the lateral line of zebrafish (Chou et al. 2017). Altering residues in the predicted pore forming region of TMC1 alters calcium selectivity in mechanically activated currents in hair cells (Pan et al. 2018), consistent with a function as a MeT channel. The *C.elegans* TMC1 homolog forms a Na^+-selective channel (Chatzigeorgiou et al. 2013), but no mechanotransduction related function has so far been described in this animal. The single TMC-A homolog in *Drosophila* is expressed in larval proprioceptors, where it is required proper crawling(Guo et al. 2016; He et al. 2018). It is also required in multidendritic neurons in the labellum for food texture detection in adult flies (Zhang et al. 2016).

Finally, the TREK-1,TREK-2 and TRAAK channels, members of the K_{2P} channel family, have been shown to be activated by mechanical forces directly sensed at the lipid bilayer(Bang et al. 2000; Brohawn et al. 2014; Maingret et al. 1999; Patel et al. 1998). Although these channels are susceptible to stretch, they are not considered specially adapted to be gated by sensory mechanical stimuli, as they seem to rather serve a homeostatic or regulatory function (Delmas and Coste 2013; Sukharev and Sachs 2012), such as regulation of pain in nociceptor neurons (Noël et al. 2009). Their role in invertebrates is much less studied.

Additional channels involved in sensory mechanotransduction

In addition to MeT channels, the mechanotransduction cascade in sensory cells requires other tightly associated components to function and drive behavior. These additional players include other channels, molecular anchors, and signaling molecules. A complete picture of the process is still missing, and it probably varies according to the specific mechanosensory cell. Some molecules have been identified in both invertebrates and vertebrates and may thus play a conserved role in mechanotransduction.

TRP channels

TRPV channels

Besides TRPN channels, various other members in the TRP channel superfamily are expressed in mechanosensory neurons and have long been considered strong candidate MeT channels (Christensen and Corey 2007). For instance, the TRPV homologs in *Drosophila* are expressed in auditory receptor neurons in the antenna and are required for direct auditory transduction, thus questioning NOMPC as the primary mechanical transducer in those cells (Lehnert et al. 2013). In *C.elegans,* loss of TRPV channels abolish responses to mechanical harsh touch stimuli, but mechanosensory currents in nociceptors expressing them are not altered (Geffeney et al. 2011). . TRPV channels are expressed in vertebrate hair cells, but when mutated it does not abolish auditory function (Morgan et al. 2018). In the unicellular alga *Chlamydomonas,* a TRPV homolog is localized to the flagellum where it is required for driving a startle response to mechanical stimulation (Fujiu et al. 2011).

TRPA channels

A TRPA1 channel in *C. elegans* contributes to mechanosensation during repeated stimulation, and they can form mechanosensitive channels in cell culture. However, mutants only showed modest loss of mechanosensitivity to nose touch (Kindt, Viswanath, et al. 2007), and thus play only secondary role. In mice, TRPA1 is expressed in skin nociceptors and modulates firing rates triggered by noxious mechanical but not by noxious thermal stimuli (Kwan et al. 2009). It also enhances mechanosensitivity when overexpressed (Brierley et al. 2011). It is not clear if it plays a modulatory role or a more direct role in mechanosensation. In *Drosophila,* the TRPA channel *Painless* is required for responses to mechanical but also to thermal nociception (Tracey et al. 2003).

PKD/TRPP2 channels

The role in mechanotransduction of the TRP channel family TRPP2 and of the related PKD1 family has been mainly studied in mammalian non-neuronal tissues. TRPP2 (also called PKD2) and PKD1 family of receptors (also called PKD1) have been suggested to play a role in flow sensation in the primary cilia of tubular cells in the kidney (Nauli et al. 2003), and in the embryonic node during the establishment of the left-right body axis (Yoshiba et al. 2012). Polycystin 1(a PKD1 family member) and Polycystin 2 (a PKD2 family member) localize to primary cilia in tubular cells (Pazour et al. 2002; Yoder et al. 2002), and mutations in these genes cause cysts in kidneys, leading to autosomal dominant polycystic kidney disease (ADPKD) (Delmas 2004). PKD2 forms an outwardly rectifying cation-selective channel (Grieben et al. 2017; Shen et al. 2016; Wilkes et al. 2017), that may be activated by increased calcium levels in the cell (Liu et al. 2018). PKD1 forms a heteromeric channel structure with PKD2 (Su et al. 2018), that can mimic the ciliary current when expressed in culture cells (DeCaen et al. 2013), but not in primary cultures (Liu et al. 2018).

The role of PKD2 and PKD1 in mechanotransduction is still unclear. It has been shown that cilia in kidney cells are not entirely necessary for mechanotransduction (Rodat-Despoix et al. 2013). Moreover, neither the PKD2 homomer nor the PKD1-PKD2 heteromer show a mechanically activated current (Delling et al. 2013; Peyronnet et al. 2012), although they do regulate other stretch-activated currents (Peyronnet et al. 2012; Sharif-Naeini et al. 2009). Thus, the consensus picture is that these channels likely have a modulatory effect on mechanically activated channels such as TREK-1, or Piezo (Ranade et al. 2015), at least in mammalian systems.

In other organisms the function of the channels is less clear. In zebrafish, PKD2L1 is a PKD2 homolog expressed in cerebrospinal fluid-contacting neurons (CSF-cNs)(Böhm et al. 2016). These cells are required to detect bending of the spinal cord. CSF-cNs are mechanosensitive and require PKD2L1 for opening a putative MeT channel (Sternberg et al. 2018). In *C.elegans*, the homolog of PKD1, LOV-1, and the homolog of PKD2, PKD-2, have been shown to be required for male copulation behavior, which is driven by both mechanical and chemical cues (Barr and Sternberg

1999). Both PKD-2 and LOV-1 are localized to the sensory cilia of ray neurons in the male's tail (Barr and Sternberg 1999). More recently, it was shown that PKD-2 is not required for mechanically-induced increase in calcium in ray neurons (H. Zhang et al. 2018). In *Drosophila*, a divergent PKD homolog called Brv1 was recently proposed to form a stretch-sensitive channel that is expressed in touch-sensitive neurons, where it contributes to responses to gentle touch (M. Zhang et al. 2018).

Sponges and some unicellular eukaryotes also encode PKD2 homologs. In the unicellular alga *Chlamydomonas*, the PKD2 homolog localizes to the flagellum (the cilium equivalent in this organism), and PKD2 mutant algal cells have defects in mating (Huang et al. 2007), which is a process that requires mechanical interactions between flagella of opposite sexes. And in the non-ciliated unicellular amoeba *Dictyostelium*, PKD2 localizes to the plasma membrane, where it is thought to respond to mechanical signals based on the rheotaxis defects seen in knockout mutants (Lima et al. 2014).

A note on other molecules involved in mechanotransduction

Numerous additional molecules tightly linked to mechanosensory transduction have been identified in either mammals or in invertebrates, and new molecules are being continuously added to the list (e.g.(Murthy et al. 2018; Xu et al. 2018)). However, for most of them it is not known how conserved their function is across animal mechanosensory systems. Thus, they will not be considered for the purpose of this study.

As it can be judged from the previous introduction, most of the molecular insights on mechanotransduction come from the established genetic model systems. Although recent studies start to explore the genetic and molecular basis of mechanosensation in non-conventional organisms(e.g.(Wang et al. 2018)), it is still almost unexplored what is the molecular composition of mechanotransduction complexes, or even in mechanosensory cells in other animal groups and what molecular adaptations are present there. In this chapter, I explore how conserved the main molecules involved in mechanotransduction are in *Platynereis*. I then analyzed the expression of a subset of these genes, focusing on those markers labeling the CRs cells, which I showed in the previous chapter to be sensitive to hydrodynamic disturbances.

Material and Methods

Whole Mount *in situ* Hybridization (WMISH)

Screen for mechanotransduction channel homologs

Aminoacid sequences of functionally characterized channels were collected from GenBank and blasted against the translated *Platynereis* transcriptome assembly version 2[7]. Putative homologs were selected based on E-value, and length of alignment. Their aminoacid sequences were blasted back against the GenBank database to rule out chain homology.

Cloning of gene fragments and probe synthesis

Riboprobes were designed to target only the open-reading frames (ORFs) of a given transcript and to be no longer than 1Kb in length. 27bp- long primers were used to amplify a given fragment and designed using Primer3 (Koressaar and Remm 2007)[8] (see **Table 0-1** in Appendix). Fragments were PCR-amplified with Taq DNA polymerase (EP0402, ThermoFisher Scientific) using as template cDNA from different larval stages (cDNA was synthetized by E. Williams and C. Verasztó). The following program was generally used:

Temperature	Time	Repeat
95°C	5 min	1 time
95°C	30 sec	35 times
Tm°C	30 sec	
70°C	x*30sec	
72°C	10 min	1 time
10°C	∞	

x=predicted #Kb [9]

PCR fragments were directly purified using the QIAquick PCR purification kit (Qiagen), and TA-cloned into TOPO® pCRII (K460001, ThermoFisher Scientific) or TOPO® pCR2.1 vector (450641, ThermoFisher Scientific) according to the manufacturer's instructions and using the chemo competent TOP10 *Escherichia coli* cells included in the kit. Positive clones (assessed by white-blue colony PCR) were purified with the QIAprep plasmid miniprep kit (Qiagen). Inserts were sequenced from both ends using M13 universal primers. Each plasmid produced in this way was archived in bacteria cryopreserved in glycerol at -80°C (see insert sequences in Appendix).

[7] Accessed at http://jekely-lab.tuebingen.mpg.de/blast/, E.Williams and G.Jékely
[8] Accessed at http://bioinfo.ut.ee/primer3-0.4.0/primer3/,
[9] Tm for each primer was calculated using the NEB online calculator: https://tmcalculator.neb.com/#!/main

Obtaining clones from EST library

Genes of interest found in *Platynereis* transcriptome were BLAST-searched against the in-house EST library[10]. Plasmids were purified by miniprep and sequenced from both ends using T7 and SP6 primers. These plasmids provided an independent source of riboprobes that could be used in parallel to those produced by gene fragment cloning. See Appendix for the list of EST clones used in this study.

Riboprobe synthesis

Around 3 µg of each plasmid was linearized with a 5' overhang, or a blunt restriction enzyme (the FastDigest brand was used in most cases, ThermoFisher Scientific) from the side of the insert closest to the 5' on the coding strand. Alkaline phosphatase (FastAP, ThermoFisher Scientific) was included in digestion reactions to reduce recircularization events. Linearized plasmids were purified with the QIAquick kit according to manufacturer's instructions, and in some cases, precipitated with ethanol and sodium acetate to remove additional impurities.

Riboprobe synthesis was performed according to the following recipe:

	Component	Volume (µl)	Source
1	Transcription Buffer 10X	2	11175025910, Roche
2	Plasmid (<1µg)	up to 11 µl	
3	DTT 0.1M	2	
4	DIG mix*	2	11175025910, Roche
5	RNase inhibitor	1	11175025910, Roche
6	T7 or SP6 RNA polymerase	2	11175025910, Roche
7	H_2O	0 to 20 µl	

Components 1-3 and 7 were incubated for 5min at 65°C, followed by a cold shock in ice-water. Components 5 and 6 were added soon after. Transcription reactions were run for ~4hrs at 37°C. RNA was purified using the RNAeasy clean-up kit (Qiagen) according to manufacturer's instructions. RNA was eluted in 25 µl nuclease-free H_2O and 125 µl deionized formamide (4650, Omnipur® Millipore) added for storage at -80°C.

Larvae fixation

Different variations of the standard WMISH protocol were followed in this study. The general procedure consisted in collecting healthy batches from different crosses at 72 hours post-fertilization in a nylon cell strainer (mesh size: 70 µm, Falcon™) placed in a container. The NSW was exchanged three times to remove any contamination. Larvae were rapidly transferred to a 50ml Falcon tube and dislodged from the mesh using ice-cold 4% formaldehyde/PTW as fixative. The

[10] accessed at http://jekely-lab.tuebingen.mpg.de/blast/

formaldehyde was freshly diluted each time from a 16% formaldehyde/water stock (prilled paraformaldehyde, Fluka) made in house.

The larvae were left to settle on ice, after which the supernatant was removed with a sterile pipette. More fixative was added (up to 40ml), and larvae were fixed for 15 to 30min at RT on a nutator. The tube with larvae was put on ice until they sank to the bottom. The fixative was removed, and larvae washed 3 times with PTW for 5 min per wash on a nutator. Larvae were gradually dehydrated in methanol and stored at least overnight at -20°C.

Hybridization

Larvae were gradually rehydrated in PTW and washed 3 times to dilute out the methanol. To increase probe penetration, larvae were incubated in 0.1mg/ml proteinase K (P4850, Sigma) for 1 min without mixing. To quench the proteinase K activity 2mg/ml glycine (#50046, Sigma) was added and exchanged soon after (<1min) with more glycine solution and left incubating for max. 1 min. Afterwards, permeabilized larvae were fixed in freshly prepared 4% formaldehyde/PTW for 20min at RT with gentle mixing. Excess fixative was diluted out with 5 consecutive PTW washes. Larvae were transferred to 2ml microcentrifuge tubes, and then stepped into hybridization buffer. The hybridization buffer used was prepared according to the following recipe and stored at -20°C until use:

Reagent	Final conc.	Source
Deionized formamide	50%	4650, Omnipur® Millipore
Saline-sodium citrate (SSC) pH4.5	5X	
Heparin	50 µg/ml	H3149, Sigma
Torula-RNA	5 mg/ml	R6625, Sigma
Tween 20	0.10%	P2287, Sigma
Nuclease-free H_2O		

Samples were prehybridized for 1 to 4hrs at 65°C in a thermomixer (Eppendorf) shaking at 300rpm. Excess hybridization buffer was then removed and riboprobe solutions pre-warmed at 70°C for 10 min and then added to the larvae. Riboprobe solutions were prepared at varying dilutions (1:124 to 1:11.5 in hybridization buffer) in 250 µl. In general, a 1:24 dilution was used for newly tested riboprobes and the concentration adjusted in successive experiments depending on the signal-to-noise (SNR) ratio. For example, if a first test did not give any signal but also not noise, the probe concentration was increased for the following test. However, if the SNR was too low, the probe concentration was reduced. In general, a good SNR was obtained for low-expression genes by using a mix of different non-overlapping probes at low concentrations (but this method can only be applied for long ORFs). The samples were transferred to an oven prewarmed at 65°C to preclude condensation. Hybridization was run for ~16hrs at 65°C.

Post hybridization washes and primary antibody incubation

Riboprobes solutions were exchanged for post hybridization solution 1 (50% Formamide (Roth GmbH), 2X SSC ,0.2% Tween 20) and washed twice at 65°C for 30min. 2 more washes with post hybridization solution 2 (2X SSC ,0.2% Tween 20) for 30 min at 65°C were done followed by 2 washes 30 min with post hybridization solution 3 (0.2X SSC ,0.2% Tween 20) at 65°C. Post hybridization solutions 1-3 were freshly prepared and prewarmed in a water bath at 67°C for 10min but no longer than 15 min to minimize evaporation. Samples were not shaken during post hybridization washes.

Larvae were blocked with 5% sheep serum (SS; cat# S3772, Sigma)/PTW for 1hr at RT while shaking on a nutator. Monoclonal mouse anti acetylated tubulin at a 1:250 dilution (T7451, Sigma) and sheep anti Digoxigenin-AP at 1:2000 dilution (11093274910, Roche) were pre-blocked in 5%SS/PTW for max. 30min prior to addition to the samples. Antibody incubation was run overnight at 4°C with gentle shaking.

Color reaction

Unbound antibodies were removed by washing the samples at least 8 times with PTW at RT with short washes at first and then increasing the length of each wash. Larvae were rinsed for 10min in freshly prepared and filtered STOP buffer (100mM TrisCl pH **7.5**, 100mM NaCl, 50mM $MgCl_2$, 0.1% Tween 20) to prevent precipitation of the staining buffer. Samples were pre-equilibrated by 2 washes, 5 min each, in alkaline phosphatase (AP) buffer (100mM TrisCl pH **9.5**, 100mM NaCl, 50mM $MgCl_2$, 0.1% Tween 20). The staining solution was prepared according to the following recipe:

Reagent	Final conc.	Source
AP buffer	1X	
Polyvinyl alcohol	6.00%	341584, Sigma
NBT	9 µl/ml	11383213001, Roche
BCiP	7 µl/ml	11383221001, Roche

Larvae were transferred to the staining solution and protected from light. Color reaction was performed for 1min to 5 days at 4°C. A small number of larvae per sample was taken out of the color reaction at different incubation times and placed in STOP buffer to assess the development of a detectable signal. Larvae were then washed with PTW three times at RT. Samples were stored at 4°C until mounting.

Larvae were stepped into 2,2-Thiodiethanol (TDE, 166782, Sigma) using 20% increases in TDE concentration to avoid an osmotic shock that would alter the morphology . Every incubation step was run for at least 10 min. A final concentration of 97% TDE/H_2O (w/w) was used for mounting. Samples were stored at 4°C in the dark.

Generation of promoter constructs

A plasmid construct was generated with restriction and oligo cloning to label the complete morphology of neurons in live and fixed animals. The construct encoded the bright fluorescent reporter tandem-dimer Tomato (tdTomato) (Shaner et al. 2004), 3 copies of the hemagglutinin (3xHA) tag at the N-terminus, and a palmitoylation signal also fused to the N-terminus of the reporter for localizing it to the membrane. The palmitoylation tag allows to label the complete morphology of neurons, including small neurites, while the 3xHA tag allows to stain the labeled cells in fixed larvae using a monoclonal anti-HA antibody. The final construct is called *Palmi-3xHA-tdTomato*_PUC-57 *(*The plasmid map of this construct and its sequence is shown in the Appendix). To make this construct, the ORF encoding tdTomato was amplified by PCR from another plasmid (kindly provided by Martin Bayer, MPI-DB in Tübingen) and cloned into the PUC-57 vector, which already included a 3'UTR from a ribosomal protein with a polyadenylation signal (also used in the plasmids used for *in vitro* transcription). Next, the 3xHA tag was inserted in frame at the N-terminal of the tdTomato ORF using annealed oligos according the protocol described in the Appendix. Correct insertion was confirmed by sequencing and restriction assay. The palmitoylation tag was inserted in the same manner (see oligo sequences in **Table 0-2** in the Appendix).

Cloning promoters

The sequences around the start codon of *PKD1-1*, *PKD2-1* and *NOMPC* were mapped to the most up-to-date assembly of the *Platynereis* genome (access kindly provided by Detlev Arendt). 27bp primers plus appropriate enzyme adaptors were used to amplify the upstream regions of these genes (see **Table 0-3** in the Appendix for primer sequences) using a two-step PCR program:

Temperature	Time	Repeat
98°C	3 min	1 time
98°C 72°C	10 sec x*30sec	35 times
72°C 10°C	10 min ∞	1 time

x=predicted #Kb

Phusion polymerase (NEB) was used for this PCR to minimize introducing changes in the promoter sequencing by amplification errors. The promoter sequences were confirmed by sanger sequencing (see sequences in the Appendix).

Purified PCR fragments were digested for 1 hr at 37°C (FastDigest REs line, ThermoFisher Scientific). The *Palmi-3xHA-tdTomato*_PUC-57 plasmid was digested with the same set of enzymes, but alkaline phosphatase was added to reduce recircularization events (FastAP, EF0654, ThermoFisher Scientific). Ligation and transformation were performed as indicated in the general

methods (see Material and Methods in the Appendix). Plasmids were injected at 250ng/µl as described in the general methods (see Appendix).

Whole mount immunohistochemistry HA

The protocol developed for this study was published elsewhere (Verasztó et al. 2017). In brief, larvae were fixed in freshly prepared 4% formaldehyde/PTW for 15min at RT. 3 consecutive PTW washes were done to remove excess fixative. 3 PTW washes were done afterwards, each for 5 min. On the same day, the samples were blocked for 1 hr in PTWST (PTW, 5% sheep serum, 0.1% Triton-X 100) at RT. Larvae were incubated overnight at 4°C in primary antibody solution (PTWST, 1:250 rabbit anti-HA antibody (Cat# 3724, Cell Signaling Technology), and in 1:250 monoclonal mouse anti acetylated tubulin). Unbound antibody was removed by shortly washing the samples 3 times with PTW, and for 5 more times for at least 20 min per wash. Samples were blocked by incubation in PTWST for 1hr at RT, and incubated in secondary antibody solution (PTWST, Alexa Fluor-488 1:250 (Cat# A-11001, ThermoFisher Scientific) and anti-rabbit Alexa Fluor-633 1:250 (Cat# A-21070, ThermoFisher Scientific) for 2hr at RT. Larvae were washed with PTW for 5 times more preceded by 3 short PTW washes. Larvae were mounted in PTW and imaged using an LSM 780 NLO or a LSM880 confocal microscope (Carl Zeiss GmbH).

Phylogenetic analysis

Phylogenetic analysis in this study was performed for the TRPP and PKD-1 families. The amino acid sequences of the three human genes in the TRPP family (TRPP2, TRPP3 and TRPP5) and the five homologs of the PKD1 family (PKD1L1-3, PKDREJ and PC1) were used as queries in a BLAST search for homologs against the NCBI non-redundant database and in the *Platynereis* transcriptome[11]. The transcripts thus recovered were translated using an online tool[12] (see sequences in Appendix). *Platynereis* PKD1 and PKD2 were also used as queries to find additional sequences. Sequences from *Clytia hemisphaerica* and *Bugula neritina* were recovered by TBLASTN against the Compagen database (Hemmrich and Bosch 2008). Sequences from *Planaria torva* were collected from PlanMine (Brandl et al. 2016). Sequences from *Petromyzon marinus* were collected from Ensembl (Zerbino et al. 2018). To collect homologs of the Polycystin-1 like family only the 11 transmembrane domain was used as query. The final set was curated to equally represent the breadth of animal phyla as much as the available data allowed (see **Table 0-7** in Appendix for sequence IDs).

For the PKD1-PKD2 joint phylogeny the full sequences of the PKD1 and PKD2 homologs were conjunctly aligned using Clustal Omega (Sievers et al. 2011). The sequence set was further reduced

[11]Accessed at http://jekely-lab.tuebingen.mpg.de/blast/ (E.Williams and G.Jékely).
[12]Accessed at http://insilico.ehu.es/translate/

to include only sequences < 90% identical. The alignment was trimmed to include only the 6 transmembrane I domains homologous to the TRP-channel homology region common to both families. For the separate PKD2 and PKD1 family phylogenies any clearly alignable region was included. Only full sequences less than 90% identical were used for the alignment.

For all phylogenies Gblocks alignments (Talavera and Castresana 2007) were used for the phylogenetic reconstruction (see Appendix for alignments). Maximum likelihood trees were recovered with PhyML (Guindon et al. 2010) using the SMS model selection tool (Lefort et al. 2017) and aLRT statistics (Anisimova and Gascuel 2006). Trees were visualized in FigTree[13] and nodes with statistical support < 0.97 were edited with Adobe Illustrator CS5 (Adobe) to show up as polytomies in the final tree. A custom Perl script was used to format the names of each taxon in the tree. The script is available in GitHub[14]. Phylogenetic trees are available in the Appendix.

Image processing and figure assembly

Confocal stacks were visualized in Imaris (Bitplane) and snapshots acquired using clipping planes to highlight planes of interest. Figures were assembled in Adobe Illustrator CS5 and CC (Adobe Inc.). SEM micrographs were pseudo colored in Adobe Photoshop CS5 (Adobe). Snapshots of the electron microscopy volume were acquired with CATMAID (Saalfeld et al. 2009). Scale bars on these snapshots were defined assuming the complete larva is 260 μm long. Scale bars were generated using Fiji/ImageJ (Schindelin et al. 2012).

[13]Downloaded from http://tree.bio.ed.ac.uk/software/figtree/

[14]https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/FigS4_Formatnaming_tree.pl

Results

Molecular characterization of penetrating uniciliated neurons

In parallel to the anatomical and physiological characterization of putative hydrodynamic receptors, molecular data can provide additional evidence for the sensory modality of a cell type. In particular, the molecules most tightly associated with the sensory transduction cascade can be of valuable information to map and identify candidate cell types with a particular sensory modality, namely the primary receptors transducing the stimuli to a (electro)physiological change.

A list of mechanosensory cell-associated factors was assembled from the literature, focusing primarily on candidate or confirmed mechanotransduction (MeT) channels. Platynereis' genome encodes multiple homologs in all the families that have been implicated in mechanotransduction (**Table 3-1**). The resulting number of candidate genes was too high to undertake a systematic analysis of gene expression at different stages. An arbitrary subset of these genes was cloned and WMISH performed at the nectochaete stage. Of particular interest was finding which of such genes was expressed in the CR neurons, as these cells were found by calcium imaging to be sensitive to hydrodynamic disturbances (see Chapter 2).

Table 3-1. Number of homologs in each family of mechanotransduction channels found in the Platynereis transcriptome

Superfamily	Subfamily	# Homologs in *Platynereis*
DEG/ENaC channels	DEG/ENAC	9
	ASIC	9
	DEG/ENAC/ASIC	6
	FaNAC	9
	Unkown affinity	42
TRP channels	TRPP/PKD	12
	TRPV	5
	TRPA	3
	TRPN	1
Piezo	Piezo	1
TMC1/2/3	TMC1/2/3	1

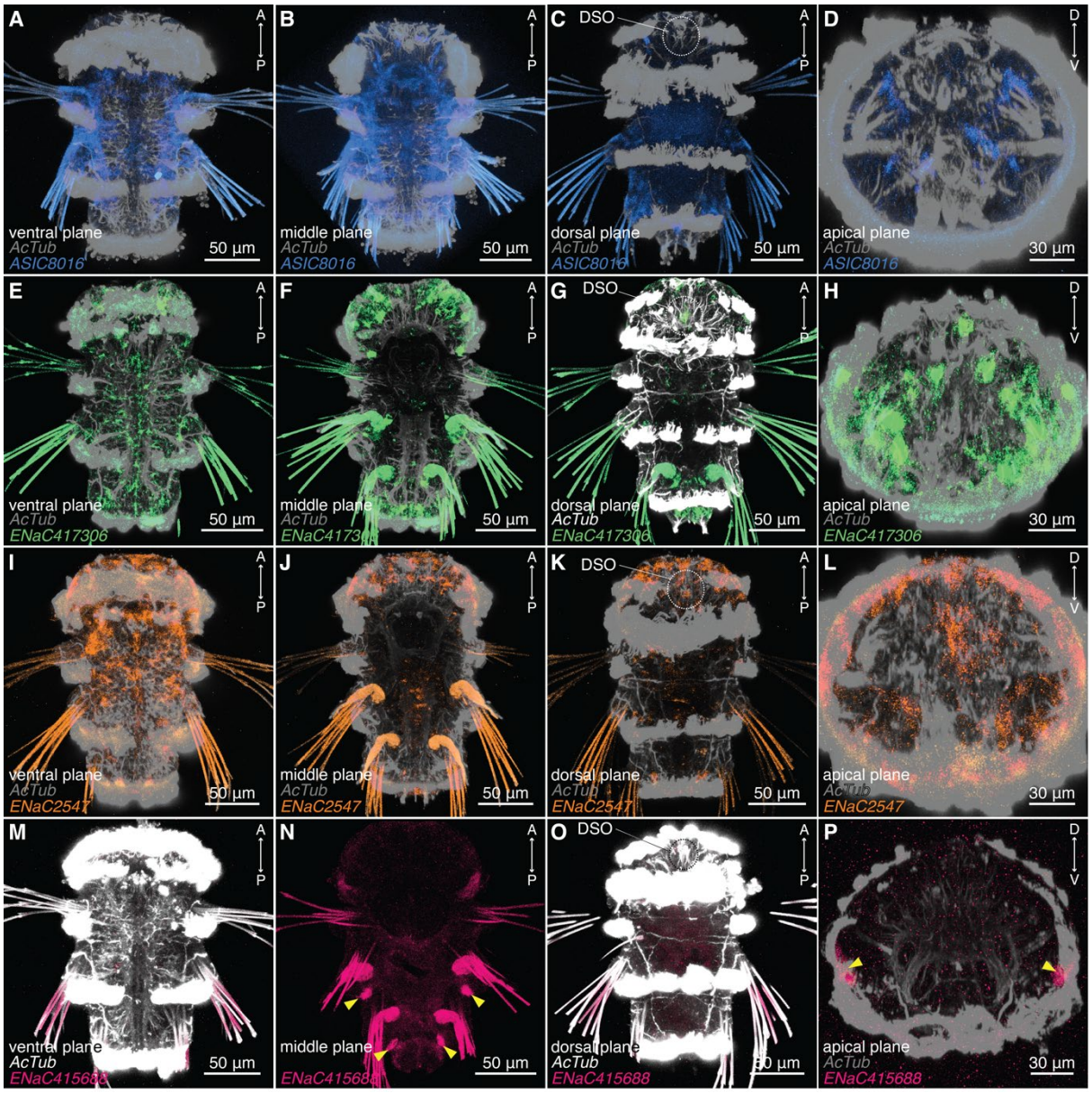


Figure 3-1 *ENaC/ASIC* channels have diverse expression patterns at the nectochaete larval stage. (**A**-**P**) Representative samples highlighting the expression pattern of four ASIC/ENaC genes in different planes. (**A**-**D**) ASIC8016 expression pattern. (**E**-**H**) ENaC417306 expression pattern. (**I**-**L**) ENaC2547 expression pattern. (**M**-**N**) ENaC415688 expression pattern. Ventral view in A, B, E, F, I, J, M, and N. Dorsal view in C, G, K, and O. Apical view in D, H, L, and P. Yellow arrowheads in N and P point to expression signal.

DEG/ENaC Channels

Numerous channels homologous to the DEG/ENaC/ASIC superfamily are encoded in Platynereis genome. Based on BLAST searches they were assigned to the main subfamilies in this group (**Table 3-1**). Some of them have similarity to DEG/ENaC channels, others to ASIC channels, and yet other group is similar to FaNaC channels. However, the great majority is not classifiable in this way. Although there have been phylogenies of certain subfamilies of the DEG/ENaC superfamily(Lynagh et al. 2018; Schmidt et al. 2018), a phylogeny including members from all the groups is needed to resolve the kinship of the different genes in *Platynereis*.

DEG/ENaC Channels show diverse expression patterns at the nectochaete stage

Only four of all the different DEG/ENaC homologs tested showed a reproducible signal at the nectochaete stage (**Figure 3-1**). The expression pattern and expression level were widely different for each of the genes. For instance, *ASIC8016* is expressed in ventrolateral, parapodial and antennal muscles (**Figure 3-1**A-D), as determined by comparing its expression to phalloidin stainings (Fischer et al. 2010). In sharp contrast, *DEG/ENaC 417306* was expressed in a numerous, yet defined set of cells in the episphere, and pygidium (**Figure 3-1**E-H). Expression on cells on the surface layer (epithelial cells) of the episphere was detected for *ENaC2547*, but the SNR was not optimal (**Figure 3-1**I-L). The fourth gene, PduDEG/*ENaC 415688*, was expressed in a much restricted domain, specifically sensory cells in the second and third parapodia (**Figure 3-1**M-P). However, none of these and other genes in this family showed expression in CR neurons.

WMISH using riboprobes against transcripts encoding other channels in this family was performed, but not clear expression was observed.

Transient Receptor Potential (TRP) channels are present in *Platynereis* genome

Homologs of the TRPV, TRPA, TRPP and TRPN families were found in *Platynereis* (**Table 3-1**). Most of these families were represented by more than one homolog (except TRPN where only one gene was found).

At the time this screen was performed the expression pattern of TRPV, and TRPA homologs was the subject of another doctoral thesis. Thus, the present analysis focused on the TRPP and TRPN homologs.

TRPP2 and Polycystin-1 homologs are expressed in CR and other ciliated neurons

PKD2-1 is expressed in numerous penetrating ciliated cells including CRs

12 TRP channels of the TRPP family were identified in *Platynereis* genome (**Table 3-1**). One of the two homologs in the Polycystin-2/TRPP2 subfamily, *PdumPKD2-1* (hereafter *PKD2-1*) was expressed in a range of different penetrating ciliated cells at the nectochaete stage (**Figure 3-2**). High-resolution scans revealed that *PKD2-1* was likely expressed in head CR1 and hCR2 neurons (**Figure 3-2**E). Expression in the putative hCR1 and hCR2 cells is also seen at the trochophore stage (**Figure 3-3**A). *PKD2-1* is possibly expressed in CR neurons in other parts of the body, as high expression was detectable in organs containing CR neurons, such as the DSO (**Figure 3-2**C), the antennae (**Figure 3-2**D-E), and the head and pygidial cirri (**Figure 3-2**I).

PKD2-1 was additionally expressed in a number of penetrating ciliated neurons already found by anatomical means (Chapter 2). This gene was found to be expressed in cells on the dorsal and ventral side of the trunk in a pattern highly similar to that of biciliated cells adjacent to the paratrochs and in the central trunk region (**Figure 3-2**A,C; compare to **Figure 2-6**). High-

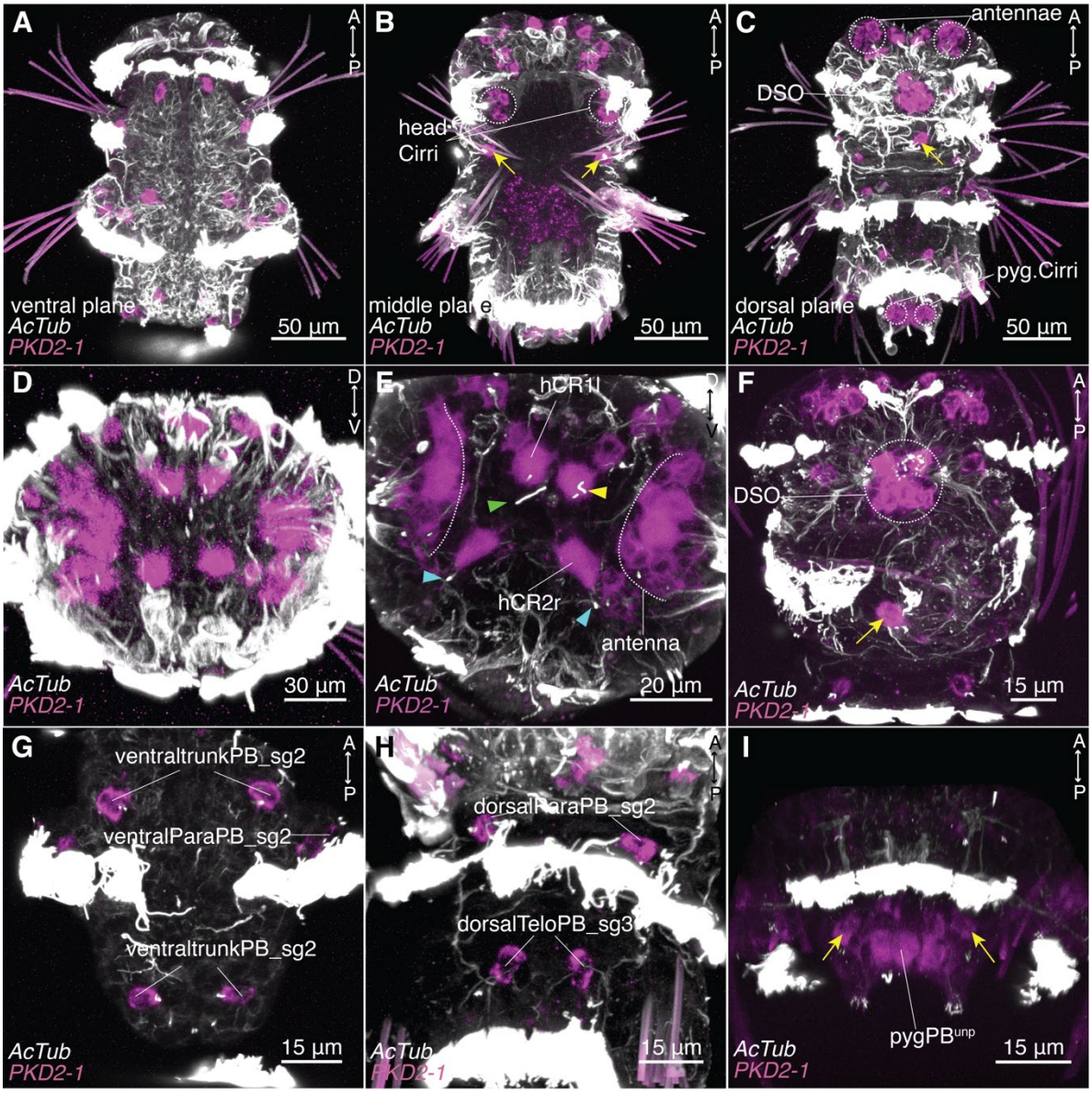


Figure 3-2 *PKD2-1* is expressed in sensory neurons in head, trunk and pygidium. (**A**-**I**) *In situ* hybridization against *PKD2-1*. (**A**-**D**) Different planes of a representative stack showing *PKD2-1* expression in the ventral (A), middle (B), dorsal (C) and apical (D) domains. (**E**) Higher resolution scan of episphere revealing *PKD2-1* expression in hCR1 and hCR2. Yellow arrowheads point to hCR1 cilia, and cyan arrowheads point to hCR2 cilia. Green arrowhead points to the MS1 cilium. (**F**) On the dorsal side of the episphere *PKD2-1* is expressed in multiple cells in the DSO and in other cells in the vicinity. (**G**-**H**) *PKD2-1* is expressed in PB cells in both the ventral (G) and on the dorsal (H) sides. (**I**) *PKD2-1* is expressed in pygPB^{unp}, and in other cells in the pygidium. Yellow arrows in B, C, F and I point to putative CR neurons. Ventral view in A, B, G and I. Dorsal view in C, F and H. Apical view in D and E.

resolution scans revealed that *PKD2-1* was indeed expressed in the biciliated cells of the ventral and dorsal sides of the trunk (**Figure 3-2**G-H).

The pygidium also showed high *PKD2-1* expression levels (**Figure 3-2**C, I). Besides the signal in the pygidial cirri, a patch of expression in the location of the biciliated cell pygPB^{unp} (**Figure 3-2**I). Most of the cells in the pygidial cirri and in the other sensory organs could not be unambiguously identified by *in situ* hybridization.

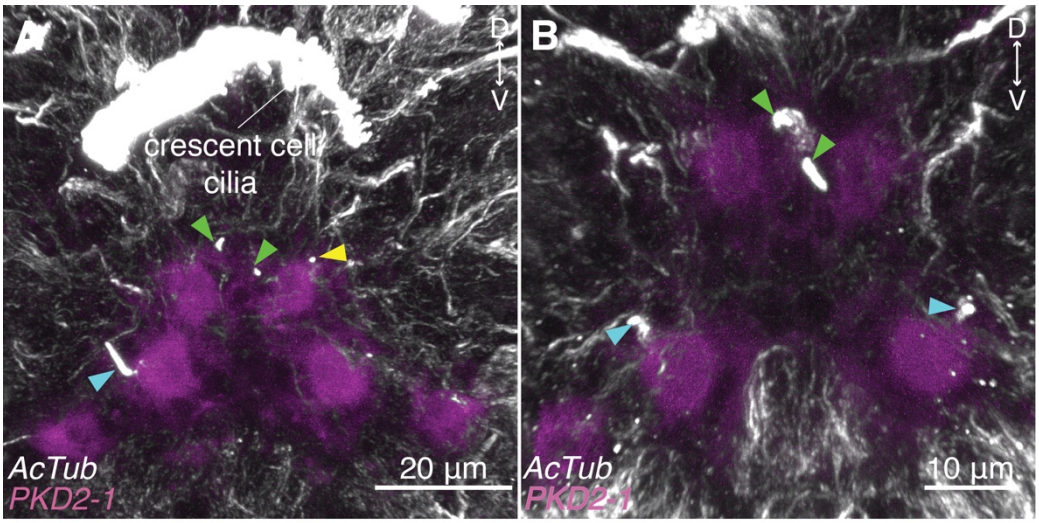


Figure 3-3 *PKD2-1* is expressed in putative hCR cells, MS1 and MS2 at the trochophore stage. (**A**) Apical view of the episphere of a trochophore larva showing *PKD2-1* expression in putative hCR cells (compare to **Figure 3-2**E). (**B**) High resolution scan of the episphere showing additionally weak expression in the MS1 and MS2 neurons. Yellow, cyan, and green arrowheads in A and B point to hCR1, hCR2, and MS cilia, respectively.

No detectable expression was found in MS cells at this stage (**Figure 3-2**D-E). However, faint *PKD2-1* expression in MS1 and MS2 neurons was observed in high-resolution scans of the episphere at the trochophore stage (**Figure 3-3**B). Finally, probes against the second *PKD2* ortholog *PKD2-2* did not yield any detectable expression at the nectochaete stage.

To confirm and further define the set of cells expressing *PKD2-1*, a ~1.5Kb fragment upstream of the start codon was cloned upstream of a membrane-tagged reporter and injected at the one-cell sage. This promoter construct labels in a mosaic pattern the complete morphology of neurons, including dendrites and axonal structures (Verasztó et al. 2017). In the episphere, the *PKD2-1* promoter construct labeled the unciliated neurons hCR1 and hCR2, hPU2l and hPUc1 (**Figure 3-4**A). In the trunk, paratroch and trunk PBs cells were also labeled, as recognized by their position and neurite projection pattern (**Figure 3-4**B-C). The construct labeled the giant biciliated pygPB^{unp} in its entirety, including its cilia and biaxonal morphology (**Figure 3-4**D, F). This confirms that this sensory cell expresses *PKD2-1*. A penetrating biciliated neuron in the pygidium (probably pygCirrusPBc1) was also labeled with the construct (**Figure 3-4**E). The promoter revealed that other uniciliated and biciliated cells express *PKD2-1* such as the uniciliated neurons hPU2 in the ventral side of the episphere (**Figure 3-4**G). The gland biciliated neurons spinPBs also express *PKD2-1* (**Figure 3-4**H-J).

These data, together with the WMISH expression patterns indicate that *PKD2-1* is expressed in many of the penetrating ciliated cell types identified in the EM volume (see Chapter 2), including the CRs and pygPB^{unp}. No MS neurons were labeled with the promoter construct, thus suggesting these cells do not express this gene at the nectochaete stage.

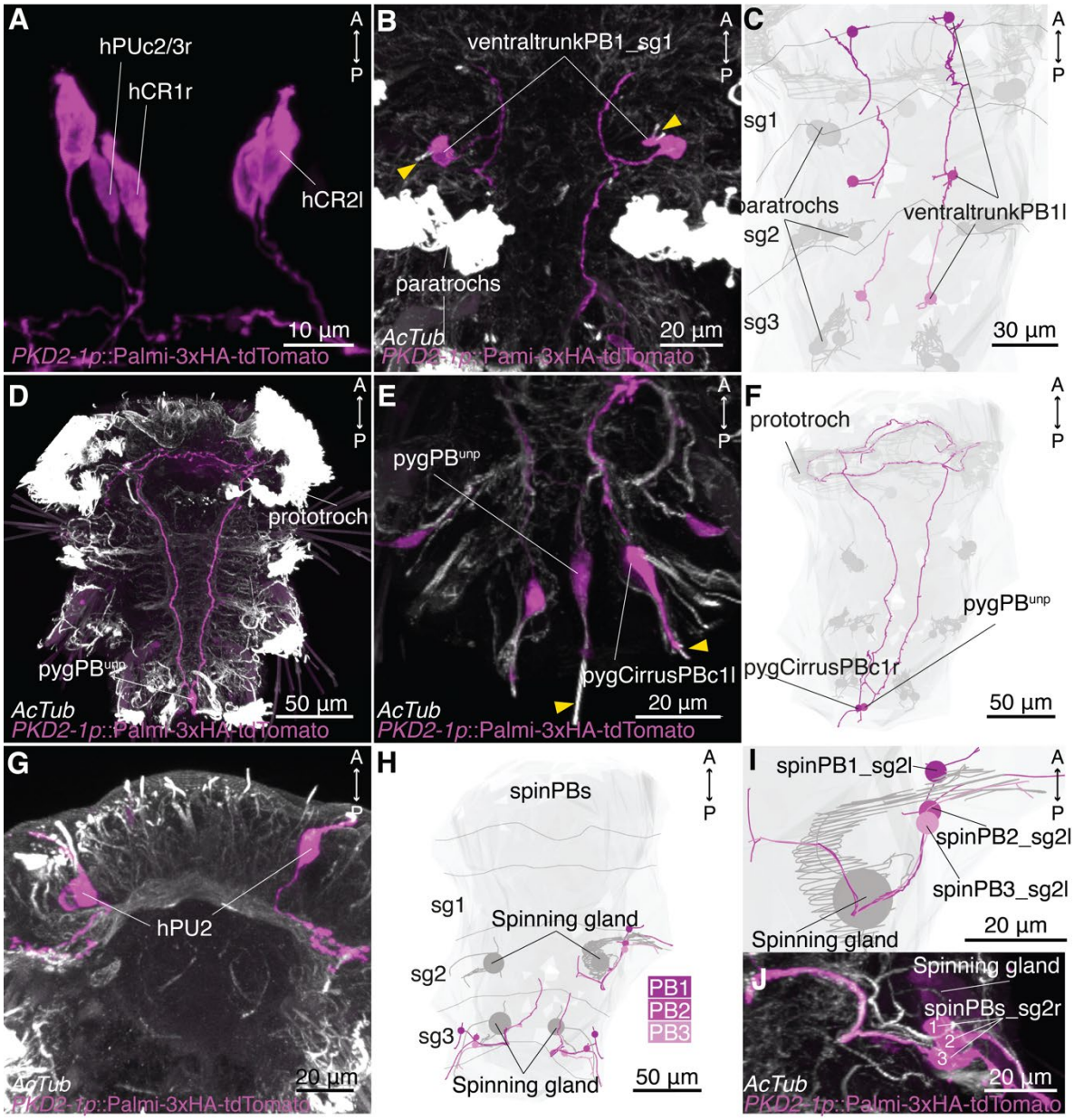


Figure 3-4 *PKD2-1* promoter construct drives expression in CRs, pygPB^{unp} and other sensory cells. (**A**) *PKD2-1* promoter labeling of hCR1, hCR2 and one of the episphere PUc (hPUc2/3). (**B**) The reporter labels biciliated cells in the trunk (ventraltrunkPBs). (**C**) Reconstruction of the ventraltrunkPBs cells in the electron microscopy volume shows a coincident morphology to the cells shown in B. (**D**-**E**) *PKD2-1* promoter drives expression of the reporter in the sensory/ciliomotor neuron pygPB^{unp} and in the single biciliated cell in the pygidial cirrus pygCirrusPBc1. (**D**) The promoter also labels pygPB^{unp} including the bifurcating neurite projection that innervates the prototroch. (**F**) EM reconstruction of pygCirrusPBc1 and pygPB^{unp}. (**G**) Two cells similar to hPU2 were also labeled by the construct. Only one was found in the EM volume (see **Figure 2-2**B). (**H**-**I**) EM reconstruction of biciliated sensory neurons in the spinning gland (spinPBs). (**J**) The three neurons are labeled by the promoter construct. Segmental boundaries in C and H are indicated by dotted lines. Yellow arrowheads in B and E point to pairs of cilia of PB neurons.

PKD1-1 is almost exclusively expressed in CR neurons

A similar expression pattern to that observed for *PKD2-1* was detected with probes against *PdumPKD1-1* (from now on, *PKD1-1*), one of the 10 *Platynereis* genes belonging to the receptor-like Polycystin-1 family (often called TRPP1 family) (**Figure 3-5**). Like *PKD2-1*, *PKD1-1* was expressed in hCR1 and hCR2 neurons, and in cells in the organs where CR neurons are located, but arguably

in fewer cells than *PKD2-1* (**Figure 3-5**). In contrast to *PKD2-1*, no *PKD1-1* expression was observed in trunk, gland or parapodial PUs/PBs (**Figure 3-5**A, B).

The *PKD1-1* promoter (~2.5Kb fragment upstream of the start codon) labeled all the CR neurons that can be unambiguously identified, and where not possible (such as in the antennae, or the cirri) in cells that by location could be CR neurons (**Figure 3-6**). The unpaired dorsal CR, doCR^{unp}, as well as some putative CR neurons in the DSO and ventralpygCRs are some of the CR neurons labeled with this reporter. As shown by WMISH, the promoter failed to label any of the gland, or trunk/parapodial penetrating ciliated cells. In the light of the known role of this gene family in kidney cells, it was also interesting to find reporter expression in one of the head kidney cells (data not shown). Finally, and as in the case for the *PKD2-1* reporter, pygPB^{unp} was also frequently labeled with the *PKD1-1* reporter construct (**Figure 3-6**D-F). Thus, *PKD1-1* labels almost exclusively a subset of *PKD2-1*-expressing cells, apparently being this subset mostly composed of the CR neurons and pygPB^{unp}. The co-expression of these genes suggests that the known interaction between their protein products (Tsiokas et al. 1997; Qian et al. 1997) is also conserved in *Platynereis*, as a well as their close association with ciliated cells (Barr and Sternberg 1999; Pazour et al. 2002; Yoder et al. 2002).

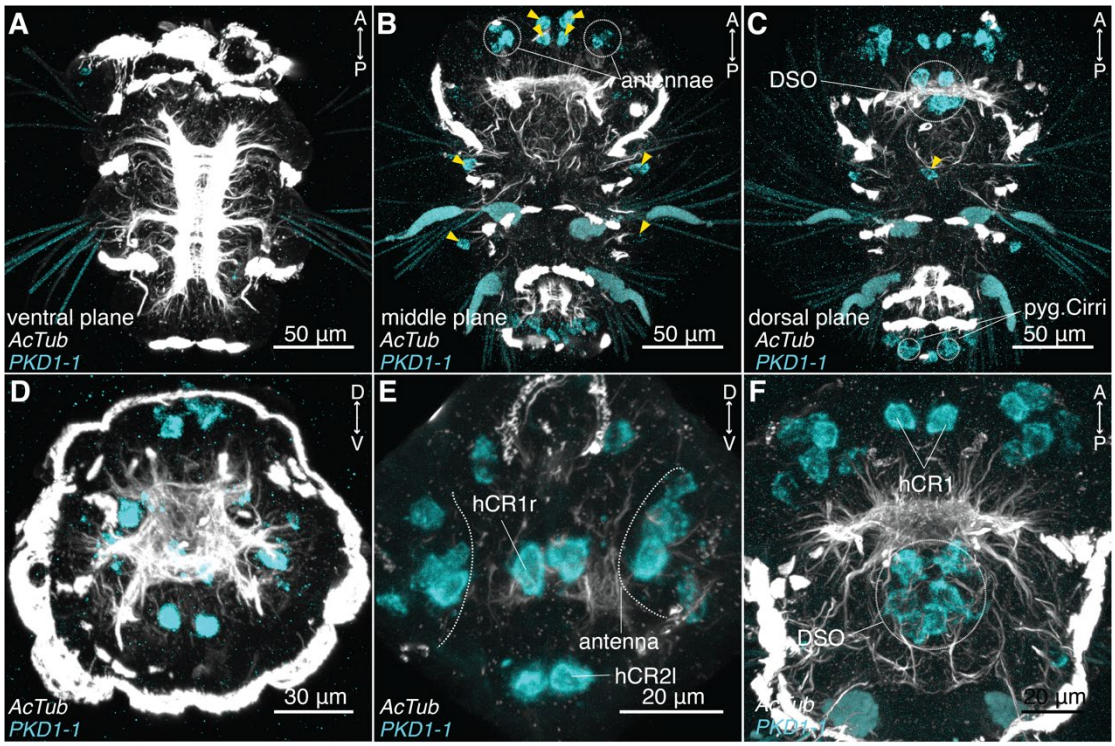


Figure 3-5 *PKD1-1* is expressed in CRs in the episphere, and in other locations where CRs are present. (**A**-**F**) *In situ* hybridization against *PKD1-1*. (**A**-**D**) Different planes of a representative stack showing *PKD1-1* expression in few cells in the head, parapodia and pygidium. (**A**) Ventral plane shows absence of *PKD1-1* expression in ventral trunk. (**B**) Middle plane highlighting *PKD1-1* expression in head region and parapodia. (**C**) Dorsal domain showing *PKD1-1* expression in dorsal sensory organ (DSO), and pygidium. (**D**) Apical plane showing *PKD1-1* expression in the episphere. (**E**) Higher resolution of the episphere showing *PKD1-1* expression in hCR1, hCR2 and in cells in the antenna (compare to *PKD2-1* expression in **Figure 3-2**). (**F**) Dorsal plane highlighting *PKD1-1* expression in the DSO. Yellow arrowheads in B and C point to putative CR neurons. Ventral view in A and B. Dorsal view in C and F. Apical view in D and E.

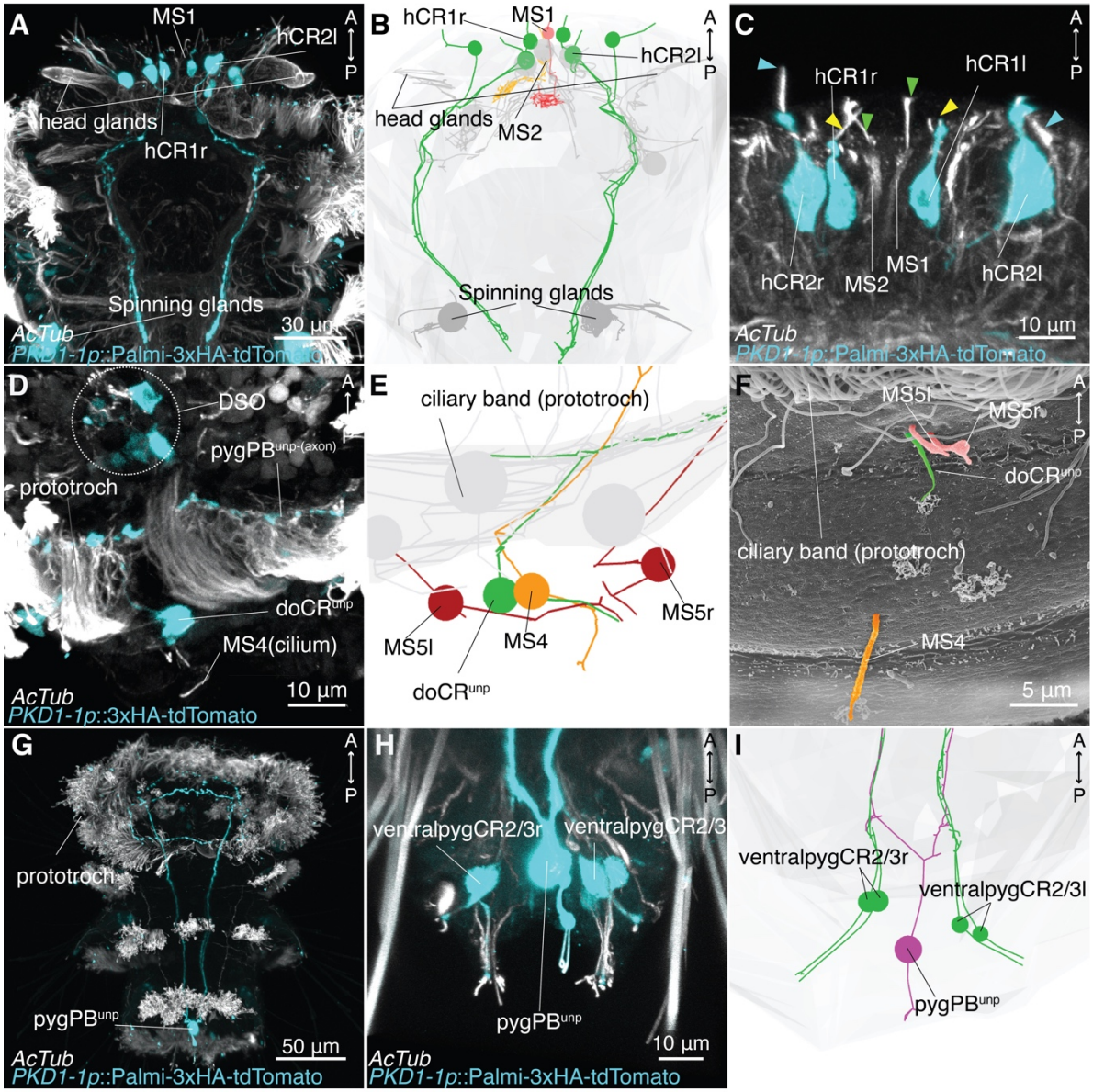


Figure 3-6 *PKD1-1* promoter drives expression in head and pygidial CRs, and in pygPB^{unp}. (**A**) *PKD1-1* promoter labels hCR1, hCR2 and some antennal cells (which could potentially be antCRs). No expression in MS1 and MS2 is detectable. (**B**) EM volume reconstruction of hCR1, hCR2 and MS1 neurons oriented as in A for comparison. Note the coincident morphology of hCR neurite projection. (**C**) Close-up view of hCR1 and hCR2 neurons labeled with the *PKD1-1* promoter construct. Yellow, cyan, and green arrowheads point to hCR1, hCR2, and MS1 cilia, respectively. (**D**) *PKD1-1* promoter drives expression in sensory cells in the DSO, and in doCR^{unp}. In the sample here shown, pygPB^{unp} was also labeled, as evident by its neurite projection along the prototroch nerve. (**E**) EM volume reconstruction showing doCR^{unp} and the other uniciliated neurons in the vicinity (of the MS type). doCR^{unp} projects down the prototroch nerve (see **Figure 2-4**D). (**F**) SEM micrograph showing the cilia of doCR^{unp} and of the nearby MS cells colored as in E. (**G**-**I**) *PKD1-1* promoter construct labels the sensory/ciliomotor neuron pygPB^{unp} and CRs in the ventral side of the pygidium. (**G**) pygPB^{unp} labelled by the construct. (**H**) Close up of the pygidium showing labelling of ventral pygidial CRs as well as pygPB^{unp}. (**I**) EM volume reconstruction showing ventralpygCRs and pygPB^{unp} in a similar orientation as in H.

PKD1-2 is expressed in sensory cells also expressing PKD2-1 but not PKD1-1

PdumPKD1-2 (from now on, *PKD1-2*) another member of the Polycystyin-1 family, show low levels of expression at the nectochaete stage in the main sensory areas of the larva, including the episphere, ventral and dorsal trunk, as well as a pair of cells in the pygidial cirri (**Figure 3-7**A-C). In the episphere, *PKD1-2* is weakly expressed in a subset of antennal sensory neurons (**Figure**

3-7D), but no signal was found in hCR1, hCR2 or MS neurons. It is also expressed in few cells of unknown identify in the DSO (**Figure 3-7**C). In the trunk, this gene is expressed in ventral and dorsal penetrating ciliated cells that look similar to those expressing *PKD2-1*(**Figure 3-7**A,C; compare to **Figure 3-2**A,C). Although future work using a promoter construct should aim to refine the expression domain of *PKD1-2* , it can be concluded for now that this gene is express in a subset of sensory cells, but none of them corresponding to hCR1 and hCR2, which are the CRs shown to get activated by hydrodynamic stimuli.

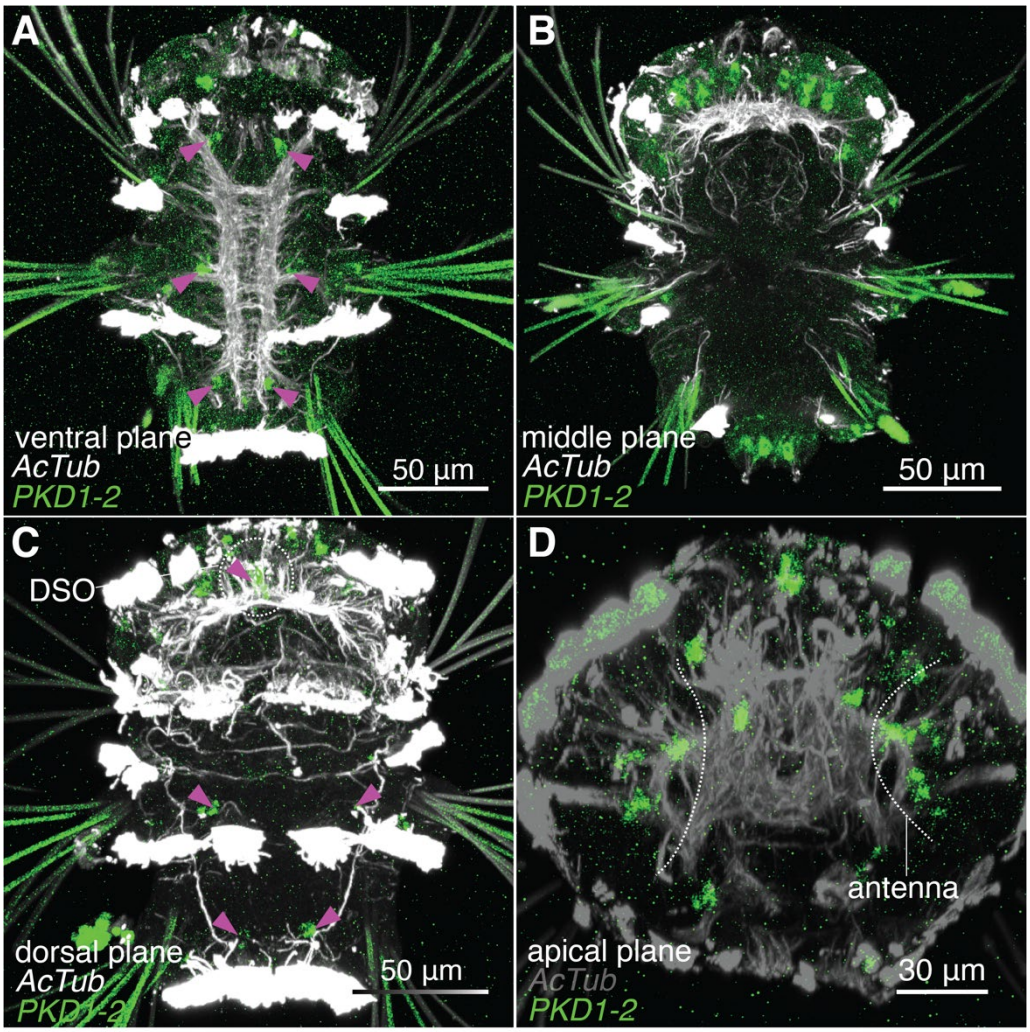


Figure 3-7 PKD1-2 expression in episphere, trunk and pygidium. (**A**-**D**) Representative sample of *in situ* hybridization against *PKD1-2*. (**A**)Ventral plane shows expression in cells in the three trunk segments (magenta arrowheads). (**B**) Middle plane highlights expression in cells located in the episphere and in the pygidium. (**C**) Dorsal plane shows weak expression in dorsal paratroch-associated cells and in the DSO (magenta arrowheads). (**D**) Apical plane showing weak expression in cells in the episphere (mostly within the antennal region). Note absence of detectable signal in hCRs. Ventral view in A and B, dorsal view in C, anterior view in D.

PKD2-1 is an ortholog of TRPP2 proteins and PKD1-1 is a novel but conserved homolog in the PKD1-like family

Joint TRPP2 and PKD-1 phylogeny

The *Platynereis* genome encodes multiple members of the TRPP2 and receptor-like Polycystin-1 families (**Table 3-1**). To clarify the orthology/paralogy relationships of each of those genes with the experimentally characterized families a phylogenetic analysis was carried out.

TRPP and PKD1-like receptors are part of the group 2 of the TRP channel superfamily (Venkatachalam and Montell 2007). The two families share 6 transmembrane domains (The TRP channel homology domain), and an "unusual" TOP domain (or Polycystin domain) (Grieben et al. 2017; Shen et al. 2016) (**Figure 3-8**A-B). This region was used to reconstruct a common maximum likelihood phylogeny (**Figure 3-8**C).

The resulting phylogeny showed a well-supported PKD2/TRPP2 group that includes *Platynereis* PKD2-1(magenta arrowhead in **Figure 3-8**C). The PKD1-like proteins split into different well supported families including a polycystin-1 (PC1) family, a PKD1L1 family and a LOV-1 family, named based on the experimentally characterized members in each (**Figure 3-8**C). Interestingly, PKD1-1 does not belong to the any of these families, but to another well-supported group without known members (cyan arrowhead in **Figure 3-8**C). For its part, PKD1-2 was grouped in the same family that includes *C. elegans* LOV-1 (green arrowhead in **Figure 3-8**C). It will be therefore hereafter called PKD1-2/LOV-1.

PKD2/TRPP2 phylogeny confirms PKD2-1 orthology

To increase the phylogenetic signal and be able to further test the homology relationships of *Platynereis* sequences with the different TRPP families, separate phylogenies of the TRPP/ PKD2-1 and PKD-1-like proteins were reconstructed using the same reconstruction method. The PKD2/TRPP2 phylogeny revealed that the *Platynereis* PKD2-1 and PKD2-2 are closely related proteins grouping with the experimentally characterized homologs in vertebrates and in other invertebrates, thus confirming an orthology relationship (**Figure 3-9**). The vertebrate sequences TRPP2 (Polycystin-2/PC2), TRPP3(PKD2L1) and TRPP5 (PKD2L2) are paralogous sequences to each other, with TRPP2 and TRPP5 more closely related to each other than TRPP3. The duplication may have occurred after the split with lampreys, at the base of the ray-finned and lobe-finned clade. Thus, these three sequences are orthologous to the invertebrate PKD2 sequences. A clade of PKD2 present in unicellular organisms was also recovered.

The alignment revealed that both PKD2-1 and PKD2-2 conserve the characteristic polycystin domain, which is needed for channel assembly and heterodimerization with PKD1 (Shen et al. 2016; Salehi-Najafabadi et al. 2017), as well as the coiled coil domains which are also involved in

heterodimerization with PKD1 (Giamarchi et al. 2010; Celić et al. 2008) (see alignment in Appendix).

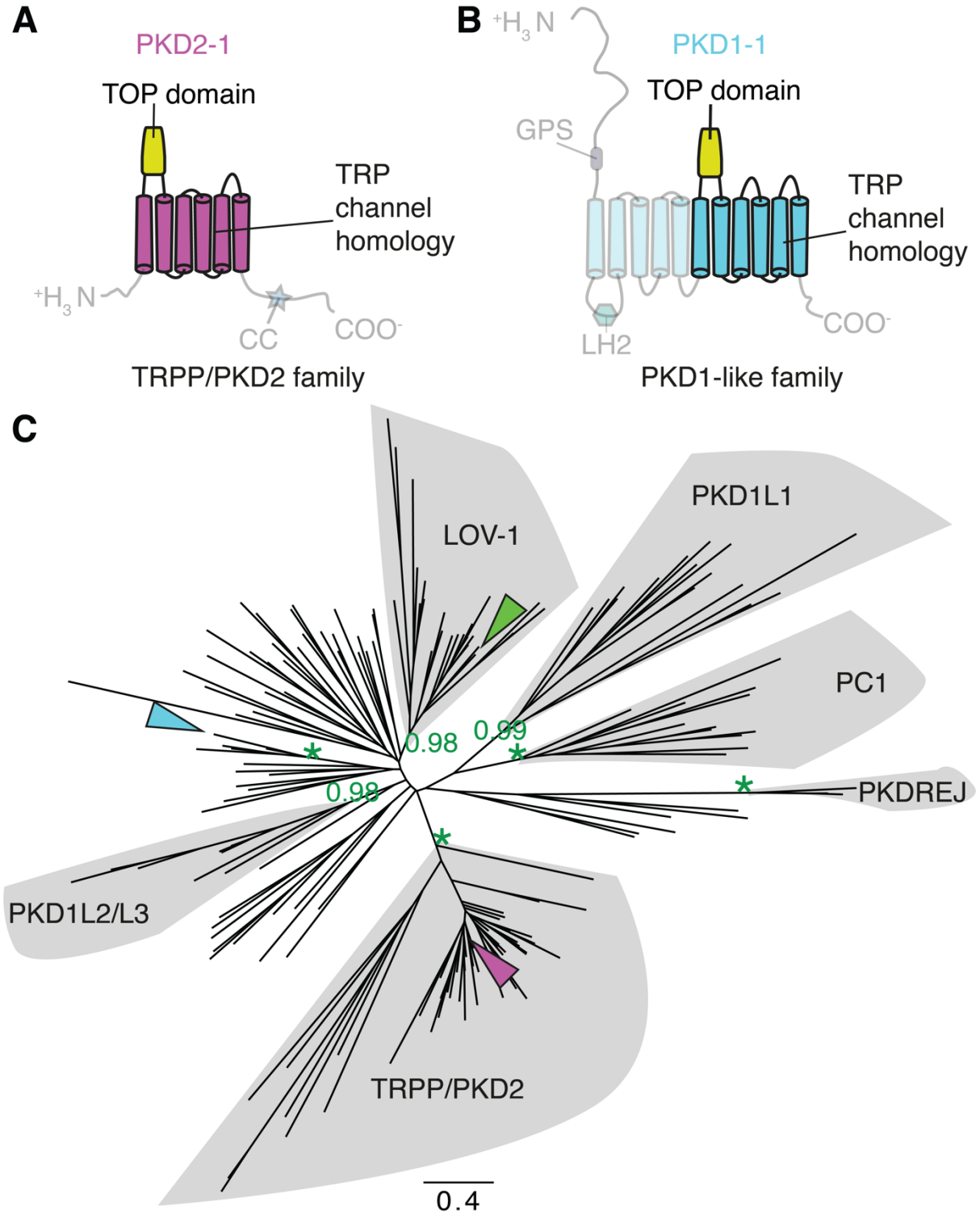


Figure 3-8 Phylogenetic affiliations of *Platynereis* homologs in the TRPP/PKD2 and PKD1-like phylogeny. (**A**-**B**) Secondary structure of PKD2-1 (A) and PKD1-1 (B) highlighting the TRP channel homology region used for the phylogenetic analysis shown in C. The remaining structural components are faded for clarity. (**C**) Maximum likelihood phylogeny of TRPP/PKD2 and PKD1-like protein families. Magenta, cyan, and green arrowheads point to the location in the tree of *Platynereis* PKD2-1, PKD1-1, and PKD1-2, respectively. Support probabilities are indicated only for relevant nodes. Asterisks indicate a support probability of 1. Well supported groups are named after experimentally characterized members. CC: coiled-coil motif; GPS: G-protein coupled receptor proteolytic site; LH2: lipoxygenase homology domain.

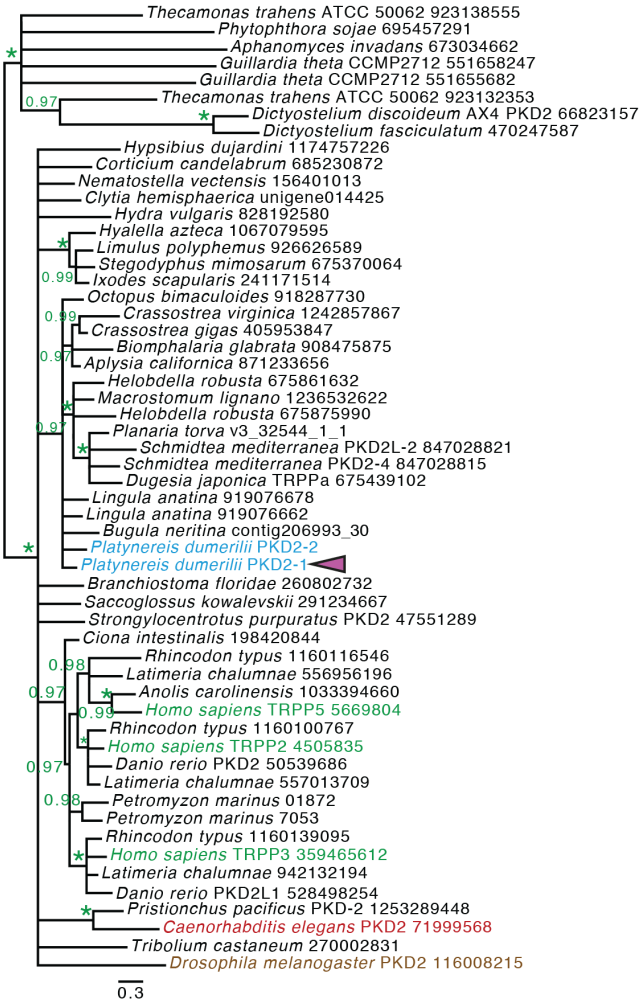


Figure 3-9 Platynereis PKD2-1 is an ortholog of TRPP/PKD2 channels. Maximum likelihood phylogeny of TRPP/PKD proteins from metazoans and unicellular organisms. Only nodes with support probabilities ≥ 0.97 are shown. Asterisks indicate a support probability of 1. Magenta arrowhead points to *Platynereis* PKD2-1. Human, *C. elegans* and *Drosophila* homologs are highlighted in different colors. Species names are shown for each sequence followed by the gene name (if any existing) and the database identifier (list of IDs in **Table 0-7** in Appendix).

Platynereis PKD1 homologs fall in conserved and novel bilaterian subfamilies

The PKD1-only phylogeny increased the statistical support for the groups recovered in the joint TRPP2/PKD1 phylogeny, and further clarified the degree of conservation of the different families in this group (**Figure 3-10**). The PC1 and PKD1L1 families are statistically well-supported groups with proteins across bilaterians, including the experimentally characterized vertebrate Polycystin-1 and PKD1L1 proteins. *Platynereis* as well as other Lophotrochozoan and Deuterostome organisms had an ortholog in each of these groups. No Ecdysozoan or non-bilaterian sequences were part of these clades. The LOV-1 family was also statistically well supported in this phylogeny. Besides PKD1-2/LOV-1, it also includes members across the bilaterians, with the exclusion of Chordates. As observed in the joint phylogeny, PKD1-1 was not part of any of these clades. It instead groups with other Lophotrochozoan sequences and together form a small, but well-supported clade. The alignment revealed that PKD1-1 does have nearly all the protein domains seen in the canonical

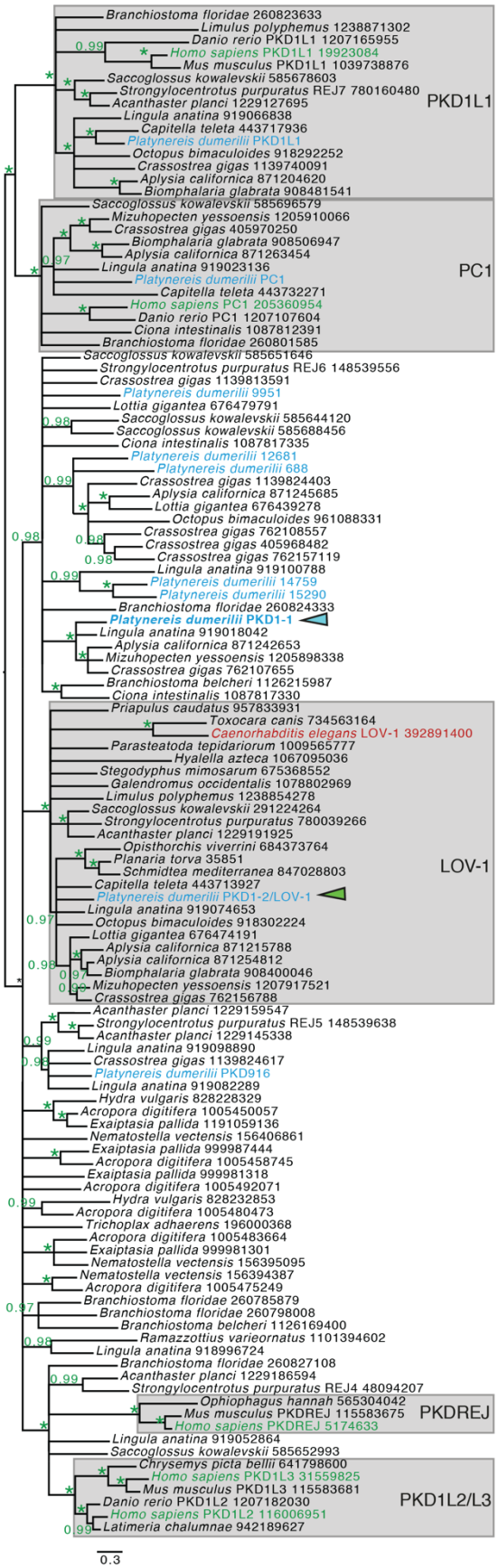


Figure 3-10 Platynereis PKD1-1 is a novel homolog in the PKD1-like family. Maximum likelihood phylogeny of PKD-1-like proteins. Only nodes with support probabilities ≥ 0.97 are shown. Asterisks indicate a support probability of 1. Cyan and green arrowheads point to Platynereis PKD1-1 and PKD1-2 (also called LOV-1). *Platynereis*, human, and *C.elegans* homologs are highlighted in different colors. Species names are shown for each sequence followed by the gene name (if any existing) and the database identifier (list of IDs in **Table 0-7** in Appendix).

PC1 (**Figure 3-8**B; alignment available in the Appendix). The only domain that absent in PKD1-1 is the egg-jelly receptor domain REJ.

The remaining 7 PKD1 homolog genes found in the *Platynereis* transcriptome were in general quite divergent from each other, and thus spread across the whole tree. The PKDREJ and PKD1L2/L3 families, which have been characterized in vertebrates, do not include any of these proteins nor any many member from the Lophotrochozoans. The non-bilaterian homologs (mostly Cnidarians and one Placozoan sequence) did not fall in any of the well-supported groups.

The TRPN channel *NOMPC* is expressed in MS and CR neurons

A single gene homolog to the TRP channel TRPN/NOMPC was found in the *Platynereis* transcriptome (**Table 3-1**). The NOMPC sequence recovered from the transcriptome encodes 29 ANK domains followed by the 6 transmembrane domain and the TRP signature domain (**Figure 3-11**A). The transcript obtained from the transcriptome assembly has a ~500aa long non-structured segment at the N-terminal. Although 5'RACE-PCR did not confirm the presence of such a long stretch, the genome did show this stretch is encoded directly upstream of the more conserved region.

At the trochophore stage, *NOMPC* is expressed in only a reduced number of cells including two penetrating uniciliated cells at the center of the episphere, and a pair of cells on the dorsal side (**Figure 3-11**B). At the nectochaete stage, the gene is expressed in multiple cells in the episphere, trunk and pygidium, albeit at low levels (**Figure 3-11**D-I). *NOMPC* expression in the episphere shows a complex but cell-specific pattern (**Figure 3-11**E, G-I). Notably, it is expressed in the center of the episphere, tightly associated to the MS1, and MS2 cilia (**Figure 3-11**G, H). The expression of *NOMPC* in MS1 and MS2 at the nectochaete stage suggests that the cells expressing this gene at the trochophore stage are the MS cells (**Figure 3-11**B-C).

Two cells adjacent to the MS1 cell also express *NOMPC* (green arrowheads, **Figure 3-11**H). These cells may have a cilium, and due to its proximity to MS1 they could be the hCR1 cells (yellow arrowheads, **Figure 3-11**H), but due to the low expression of the gene this is cannot be fully confirmed. At the trochophore stage faint expression is also seen in the vicinity of the hCR1 cilia (**Figure 3-11**B). At the nectochaete stage, *NOMPC* is also expressed in a pair of bilateral ciliated cells ventral to the MS1/hCR1 group (blue arrowheads, **Figure 3-11**H). On the dorsal side of the episphere, the expression is closely associated to the acrotroch ciliated sensory neurons (**Figure 3-11**I). The only ciliated cells in this region are the MS3 neurons (see **Figure 2-4**). There is also putative expression in the DSO (**Figure 3-11**F, I).

In the ventral side of the trunk, *NOMPC* is weakly expressed in what appear to be trunk penetrating ciliated cells (magenta arrowheads in **Figure 3-11**D, compare to **Figure 2-6**). This expression closely resembles that of *PKD2-1* (compare to **Figure 3-2**A, G). Very weak expression is also seen

in cells on the dorsal side of the trunk, in the location of dorsal paratroch PU/PB cells (magenta arrowheads in **Figure 3-11**F). *PKD2-1* is also expressed in this region (**Figure 3-2**C, H). In the pygidium, *NOMPC* expression was limited to the cell clusters belonging to the pygidial cirri (, **Figure 3-11**E). Thus, contrary to *PKD1-1* and *PKD2*-1, this gene is not expressed in pygPB^{ump}.

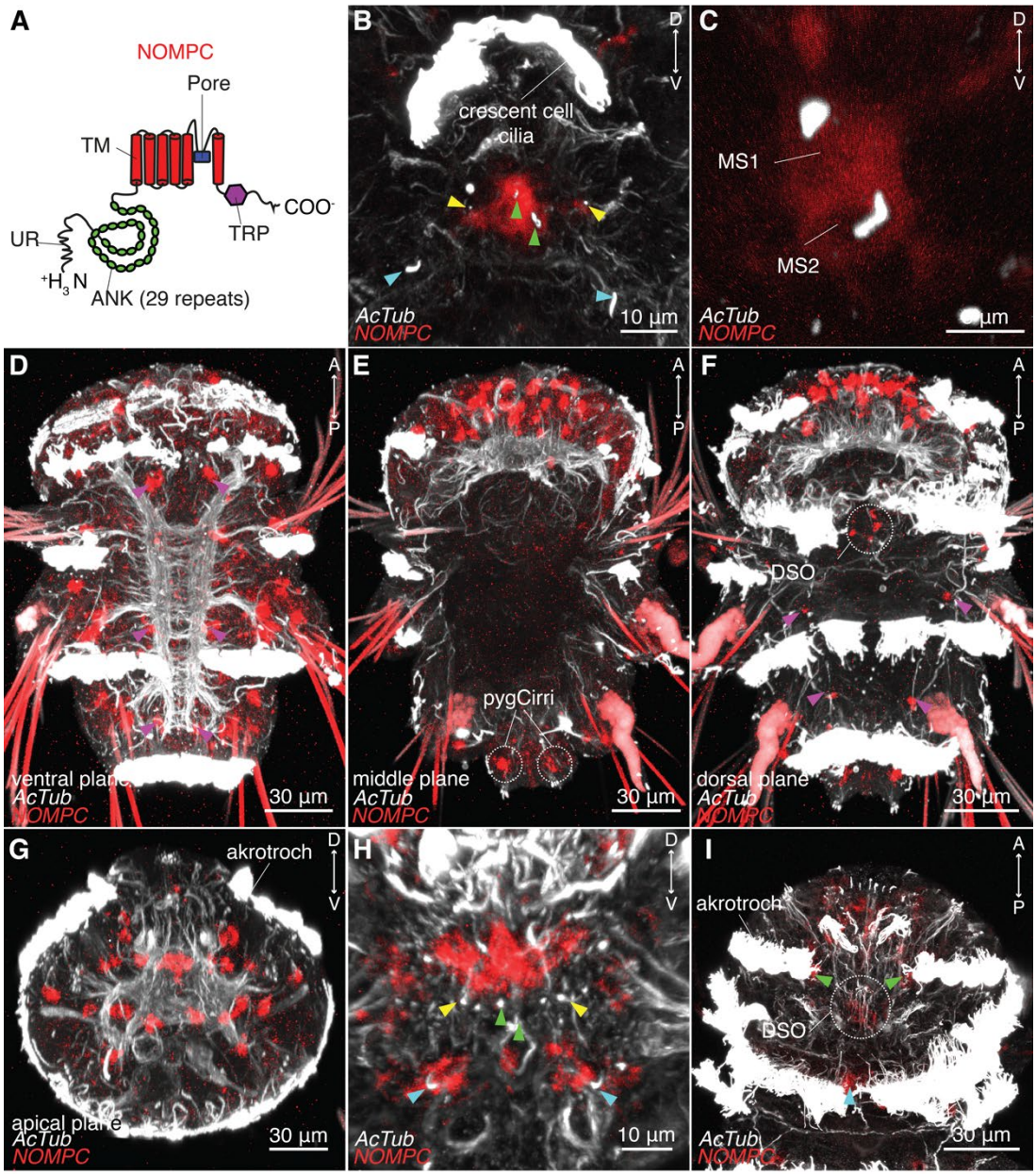


Figure 3-11 *NOMPC* is expressed in multiple sensory neurons including MS and probably CRs. (**A**) Secondary structure of *Platynereis* NOMPC. (**B**-**I**) *In situ* hybridization against *NOMPC* showing expression in cells in the episphere, trunk and pygidium. (**B**-**C**) Expression at the trochophore stage in the episphere. (**B**) MS cilia (green arrowheads), hCR1cilia (yellow arrowhead) and hCR2 cilia (cyan arrowheads) are visible. (**C**) Close up showing *NOMPC* expression in MS1 and MS2. (**D**-**G**) Representative larva used to highlight *NOMPC* expression in different planes. (**D**) Ventral plane highlighting expression in cells in the trunk (putative trunkPU/PBs) (magenta arrowheads). (**E**) Middle plane showing numerous cells expressing NOMPC in the episphere. NOMPC-expressing cells are also visible in the pygidial Cirri (pygCirri). (**F**) Dorsal plane showing NOMPC expression in DSO and weak expression in dorsal paratroch-associated cells (magenta arrowheads). (**G**) Apical plane showing cell-specific NOMPC expression in multiple cells. (**H**) Higher resolution scan of central episphere showing putative NOMPC expression in MS1 and MS2 (green arrowheads point to the corresponding cilia). NOMPC may be also expressed in hCR1 (yellow arrowheads point to cilia) and hCR2 neurons (cyan arrowheads point to cilia). (**I**) NOMPC (weak) expression in the DSO and putatively in MS3 cells (green arrowheads) and MS4 region (cyan arrowhead). ANK: ankyrin domain; TM: Transmembrane domain; TRP: TRP signature domain; UR: unstructured region.

NOMPC promoter drives expression in MS cells and other penetrating uniciliated cells but not in CRs

The complex expression pattern of *NOMPC* required the use of a promoter construct to confirm and complement the set of cells where this gene is expressed at the nectochaete larval stage. A 6.3Kb fragment upstream of the first conserved ANK domain was cloned to avoid missing the real start codon.

The *NOMPC* promoter drove expression in MS1 and MS2, as expected from the WMISH data (**Figure 3-12**A). The cells were completely labeled, including the synaptic region in the brain plexus. The morphology recovered is strikingly similar to that reconstructed from the serial TEM data (**Figure 3-12**Aa). The construct also labeled the MS3 neurons, which have a single cilium projecting out next to the acrotroch ciliary band (**Figure 3-12**B-C).The neurons fully labeled the neurite projection to the brain region where they synapse in the vicinity of the MS1 neuron (**Figure 3-12**Ba).

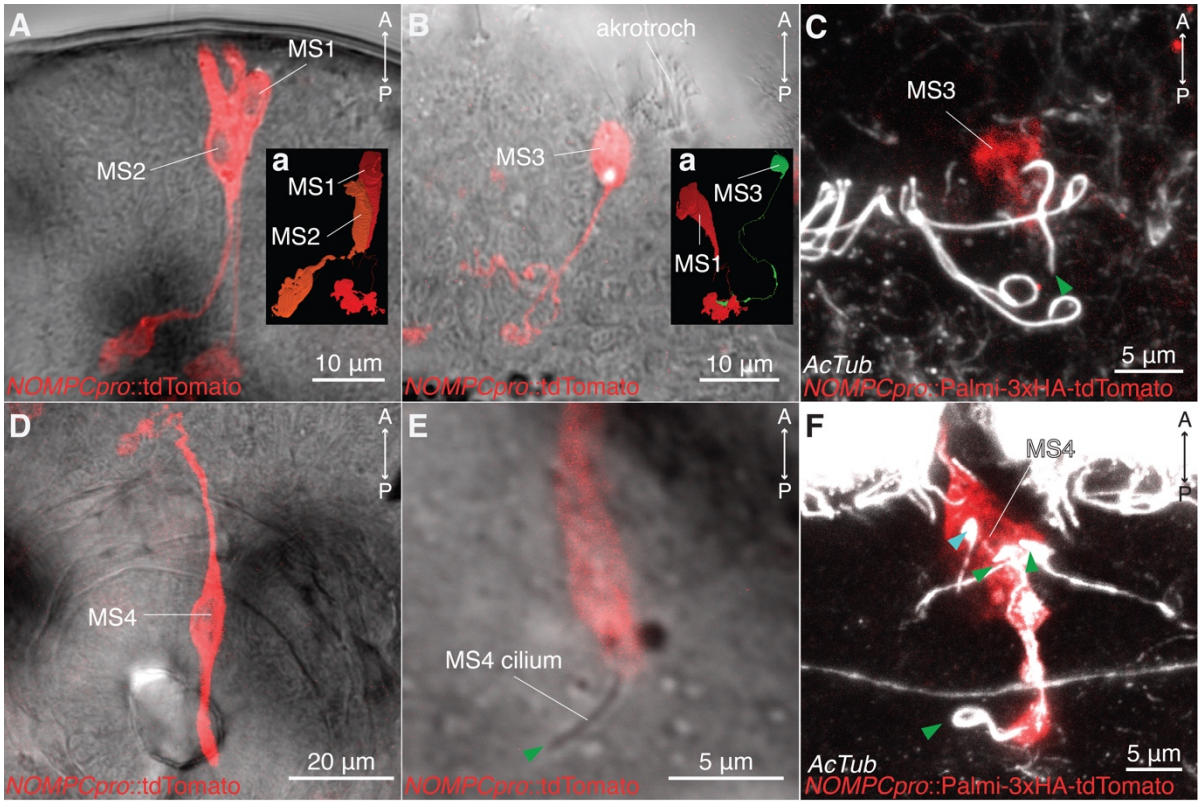


Figure 3-12 The *NOMPC* promoter construct labels MS neurons. (**A**-**B**,**D**-**E**) Bright field composite images showing expression of the tdTomato reporter driven by the *NOMPC* promoter in injected living animals. (**C**,**F**) Immunostainings against the HA tag of the tdTomato reporter expressed under the NOMPC promoter. (**A**) The construct labels the complete morphology of MS1 and MS2. (**Aa**) 3D visualization of MS1 (red cell) and MS2 (orange cell) reconstructed from the ssTEM data. (**B**) The promoter drives reporter expression in MS3, recognized by its proximity to the acrotroch. (**Ba**) 3D visualization of one of the MS3 cells (green cell) and of MS1 (red cell).(**C**) Immunostaining labels MS3 and its long cilium (green arrowhead). (**D**-**E**) Labelling of MS4 . The cilium is also visible (D). (**F**) MS4 also gets labelled in fixed specimens. The cilia of MS4 and MS5 neurons are visible (green arrowheads), as well as that of the doCR^{unp} neuron(cyan arrowhead, see **Figure 2-4**C). Dorsal views in all panels.

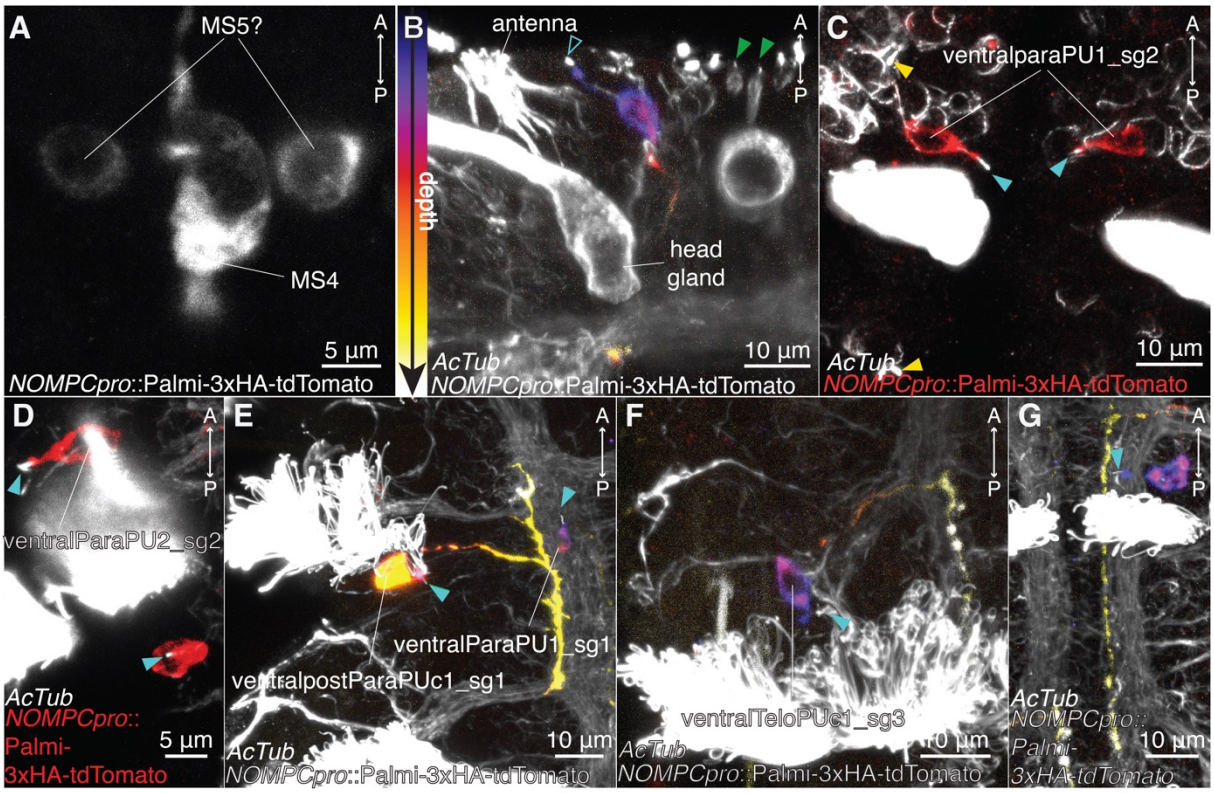


Figure 3-13 NOMPC promoter drives reporter expression in penetrating ciliated sensory cells in episphere and ventral side of the trunk. (**A-G**) Confocal microscopy projections of immunostainings against the HA tag of the tdTomato reporter expressed under the NOMPC promoter. (**A**) MS4 and putative MS5 neurons show tdTomato expression. Dorsal view. (**B**) An unidentified ciliated neuron in the episphere is labelled by the promoter (cilium, cyan arrowhead). Green arrowheads point to MS1 and MS2, which are not labelled by the construct in this specimen. The construct signal is colored by depth using a Fire LUT. (**C-G**) Confocal stacks showing expression of the reporter construct in penetrating uniciliated sensory neurons in the trunk (cyan arrowheads point to individual cilium). E-G are color-coded by depth as in B. Yellow arrowheads in C point to cilia of ventral trunk biciliated neurons. Ventral views in B-G.

One of the dorsal MS neurons, MS4, was also fully labeled with the reporter expressed under the *NOMPC* promoter (**Figure 3-12**D-F). The long neurite that projects anteriorly to the brain plexus can be clearly seen (**Figure 3-12**D), as reconstructed in the EM volume (**Figure 2-4**D-E), as well as the long sensory dendrite projecting posteriorly and the single long cilium (**Figure 3-12**E-F). A single specimen was additionally labelled in what by its relative position to MS4 seem to be the MS5 neurons (**Figure 3-13**A). The signal was weak, and the specimen was not embedded in TDE, which together precluded a confirmation of the cells' identities by analyzing their neurite projections. The only other cell in the episphere that was labelled by the construct was an unidentified sensory neuron on the ventral side (**Figure 3-13**B). As this cell could not be matched to any penetrating ciliated neuron identified in the electron microscopy volume, it may rather belong to the group of non-penetrating ciliated neurons.

The *NOMPC* promoter construct also labelled additional penetrating uniciliated neurons in the trunk (**Figure 3-13**C-G). In particular, the reporter labelled ventral penetrating uniciliated neurons associated to the bodytrochs, both anterior and posterior to the ciliary bands ((**Figure 3-13**C-D). In some cases, the complete neuron morphology could be recovered (**Figure 3-13**E-G).Thus,

while *NOMPC* is expressed in penetrating *uniciliated* neurons in the trunk, *PKD2-1* is rather expressed in penetrating *biciliated* neurons in that same region.

Piezo is not expressed at the nectochaete stage

A single gene encoding a homolog to the transmembrane channel Piezo was found in *Platynereis* (**Table 3-1**). None of the 6 riboprobes tested resulted in a clear expression pattern (see Appendix for probe sequences tested). Thus, this gene may not be expressed at the nectochaete larval stage.

Discussion

Platynereis has homologs in all channel families implicated in mechanotransduction

In this chapter, a survey was carried out for homologs in *Platynereis* to the main channel families implicated in mechanotransduction in other animals. The main purpose of this search was to find mechanosensory cell markers for the different populations of sensory cells identified by anatomical means. This approach may be justified on the grounds that a large but common set of molecules in both vertebrates and molting invertebrates have been implicated in mechanotransduction, and at least in some cases, by totally independent forward genetic screens (Chalfie 2009; Marshall and Lumpkin 2012).

Platynereis encoded homologs to all the families: from a single homolog in the Piezo, TRPN and the TMC-A families to several homologs like those in the DEG/ENaC family (**Table 3-1**). This last group may have been expanded at the base of the annelids, as the leech genome also has several ENaCs (Simakov et al. 2013). A phylogenetic analysis of this superfamily is sorely needed to place and prioritize the numerous homologs and find those that are closer to MEC-4 and MEC-10, the experimentally characterized ENaC MeT channels in *C.elegans*.

Some of the proteins found in the screen showed clear orthology to experimentally characterized groups. A single NOMPC was found in *Platynereis* that has 29 ankyrin repeats like its ortholog in *Drosophila*. This is consistent with the widespread conservation of the domain architecture for proteins in this family (Schüler et al. 2015). This repeats have been shown to be essential for the gating spring mechanism that opens the NOMPC pore upon detection of mechanical force (W. Zhang et al. 2015). The conserved protein architecture thus suggests that the *Platynereis* NOMPC homolog can also form a functional MeT channel, as shown for its *Drosophila* counterpart (Yan et al. 2013).

Molecular evolution of Polycystins

The phylogeny of bilaterian PKD1 and PKD2 homologs revealed ancient subfamilies and clarified the evolutionary kinship of experimentally characterized members in these groups. One important finding was the conservation in bilaterians of the Polycystin-1 and of PKD1L1 proteins. These proteins seem to be absent in Ecdysozoans and non-bilaterians, but homologs in *Platynereis* and in other Lophotrochozoans and Deuterostomes indicate that these proteins—whose members have been only experimentally studied in vertebrates—were already present at the base of the bilaterian split. In vertebrates, these proteins participate in seemingly different biological processes: PC1 acts in kidney cell physiology (Nauli et al. 2003), while PKD1L1 has rather a role in left-right axial patterning (Field et al. 2011; Kamura et al. 2011). Both genes, however, seem to have a role in cilia-

mediated flow detection. Moreover, PKD1L1 can rescue the flow-detection deficiency of PC1 loss-of-function mutants (Grimes et al. 2016). The deep evolutionary split of these subfamilies nonetheless suggests that their functions are not fully interchangeable.

The PKD1 phylogeny also shows that proteins that are often considered analogous in function are actually phylogenetically distant. For instance, LOV-1 in *C.elegans* is not a direct ortholog of neither vertebrate PC1 nor PKD1L1, which warns against direct comparisons to the findings in vertebrates or against assumptions of such homologs to faithfully reproduce human diseases like ADPKD (Barr and Sternberg 1999; Barr et al. 2001). Adoption of an evolutionary framework is likely to lead to more meaningful comparisons of function in order to understand the differences, and to better identify functional principles.

The LOV-1 family is nonetheless also likely to have evolved at the dawn of Bilaterians and later lost in Deuterostomes. Most of the remaining PKD1 families are more phylogenetically restricted, including that defined by *Platynereis* PKD1-1. Most of these sequences belong to Lophotrochozoans. Newly assembled Lophotrochozoan genomes and transcriptomes should help refine these groups (Y.-J. Luo et al. 2018; Marlétaz et al. 2019).

The TRPP2 phylogeny showed a more conservative picture. Only a major group of metazoan PKD2 proteins was found. The tree revealed that the two main PKD2 channels studied in mammals PC2 and PKD2L1 only occur in vertebrates. Nonetheless, it has been shown that the permeability features of these channels is different (Kleene and Kleene 2012; Kleene and Kleene 2017; Liu et al. 2018; DeCaen et al. 2016; DeCaen et al. 2013; Hulse et al. 2018). It is thus important to characterize the channel properties of invertebrate channels to know what mechanism is more general, or if the invertebrate orthologs use a different selectivity mechanism. The more divergent PKD2 homolog Brv1 in *Drosophila* that has been recently implicated in mechanosensation (M. Zhang et al. 2018) should be added in future phylogenetic analyses, as they are reported to be quite divergent to TRPP2 molecules, but still considered part of this family (Gallio et al. 2011).

Although PKD1-1 is not a direct ortholog of the experimentally characterized members in the same family, it conserves many of the functional domains in PC1 and PKD1L1 known to be useful for the association with PKD2 proteins (Sharif-Naeini et al. 2009; Tsiokas et al. 1997; Qian et al. 1997; Newby et al. 2002). Thus, the co-expression of PKD1-1with PKD2-1 in CRs and pygPB^{unp} neurons suggests that these two molecules interact *in vivo*, like other PKD1/PKD2 molecules (Barr and Sternberg 1999; McLaughlin 2017; Nauli et al. 2003; England et al. 2017). Indeed, PKD2-1 was not only co-expressed with PKD1-1, but also with PKD1-2, the LOV-1 ortholog. The expression pattern of LOV-1 did not overlap with that of PKD1-1. These suggests that PKD2-1 may be able to interact with most of the PKD1 homologs, but in different subsets of cells. It would be thus especially interesting to find the expression pattern of the conserved PKD1L1 and PC1

homologs, as in mammals they localize to the cilium, and interact with PKD2 proteins. PKD1 homologs may be in general good markers of sensory cells, exemplified by the differential expression of 10 of these genes in *Nematostella* recovered by single-cell sequencing (Sebe-Pedros et al. 2017).

Identifying mechanosensory cells: molecular, anatomical and functional information

The identification of mechanosensory cells is far from trivial. As seen from previous chapter, many cells share the ciliated morphology and some even the collar that are commonly associated to mechanosensory cells. Molecular information is also not sufficient on its own to know a cell is a mechanoreceptor—even some well-known mechanotransduction channels such as Piezo or TRP channels are expressed in other sensory cells. Therefore, the two sources of information need to be combined to identify potential mechanosensory cells. However, only after physiological characterization it can be said that a neuron has been truly identified. Following these requirements, the first and only example so far of an identified mechanosensory cell in the nectochaete larva is the CR neuron. This cell type was found to co-express *PKD1-1*, *PKD2-1* and putatively also *NOMPC*. It has also the morphology of a mechanosensory cell—namely a collar receptor with a cilium exposed to the environment—and calcium imaging revealed that it responds to hydrodynamic stimuli.

NOMPC expression in MS neurons, a putative mechanosensory cell type

The expression of the MeT channel *NOMPC* in multiple sensory neurons of the nectochaete larva suggests that not only CRs, but also other cells may have a mechanosensory modality at this stage. Particularly relevant is to identify the set of neurons expressing this gene that overlaps with the set already proposed to have mechanosensory modality by anatomical and physiological means in Chapter 2 and in previous reports. The clear expression of *NOMPC* in MS1 and MS2 neurons suggests that these cells have a mechanosensory modality despite their lack of response to the stimulus that triggers startle response (see Chapter 2). In fact, the expression pattern of *NOMPC* in MS1 and MS2 at the trochophore stage allows to conclude that the putative mechanosensory cells reported by Marlow et al to be present at the early trochophore stage (24 hours post-fertilization), are the MS cells (Marlow et al. 2014). These authors assigned mechanosensory modality to these cells primarily based on their sensory morphology and on the expresion of another TRP channel, TRPV. Marlow et al also report the expression in the mechanosensory cells of *mir-183*, a highly conserved microRNA involved in sensory cell development including vertebrate hair cells (Mahmoodian Sani et al. 2016; Christodoulou et al. 2010; Pierce et al. 2008).

Based on anatomy and morphology other putative mechanosensory cells were found. For instance, penetrating biciliated neurons in the ventral and dorsal side of the trunk co-express *PKD1-2*, *PKD2-*

1 and *ENaC417306*. Co-expression of homologs to these genes is common in mechanosensory, or in nociceptive neurons (e.g.(Turner et al. 2016)). Functional characterization of putative mechanosensory cells could follow a similar pipeline to this work. Namely, the characterization of a behavior, followed by the identification of putative candidate cells and the assessment of their activity upon a stimulus that is thought to drive their activation.

"Absence of evidence is not evidence of absence (in gene expression)"

The list of homologs was far greater than what could be analyzed in this study. Only the expression pattern of a small proportion of these genes could be determined; the list nonetheless serves as a rich source for future comparisons to other invertebrate and vertebrate systems. The failure to detect expression of many of the channels only underscores the fact that transduction channels are usually in low abundance in mechanosensory cells (Beurg et al. 2009). More sensitive techniques such as single molecule RNA FISH will be required in those cases where traditional WMISH does not work (Pichon et al. 2018; Gaspar et al. 2017). However, many of the genes were only tried once or twice, especially those in the more numerous families. A more rationalized expression screen is needed—especially focusing on those conserved families revealed by phylogenetics—to sample the diversity of expression patterns of conserved genes important for mechanotransduction. The availability of a map of all penetrating ciliated cells (see Chapter 2)—and in the near future, of all the sensory cells in the nectochaete larva cells (Jékely Lab, unpublished observations)—should make possible to identify the cells expressing these sensory markers (e.g. by high resolution scans or by promoter constructs labeling sensory dendrites).

One of these genes was the single homolog of the Piezo family. None of the seven probes tested, alone or in combination, produced a detectable expression pattern. Like NOMPC, Piezo is among the few confirmed MeT channels in animals (Coste et al. 2012), thus being an especially important marker to study mechanosensation. However, this gene has been only found to be expressed in non-ciliated mechanosensory cells—such as multidendritic nociceptors, Rohon-Beard cells and free-nerve endings in *Drosophila*, zebrafish and mammals, respectively (Kim et al. 2012; Ranade et al. 2014; Faucherre et al. 2013; Woo et al. 2015). As few non-ciliated sensory neurons have been found in *Platynereis*, Piezo may mainly play a role in non-neuronal tissues in this animal.

MS neuron as a motile-ciliated mechanosensory cell type of unknown modality

Among the genes reported to be expressed in the now identified MS1/MS2 cells are *mir-34* and *foxJ*, two molecules with a conserved role in the development of motile cilia (Yu et al. 2008; Song et al. 2014; Vij et al. 2012). This observation independently supports the motility of MS cilia described in Chapter 2.

Although a motile cilium is at first glance incompatible with a mechanosensory modality, mechanosensitivity is actually a fundamental requirement for motility even in non-sensory cilia

(Wiederhold 1976). MS neurons would not be an unusual case of a mechanosensory cell with a motile cilium. For instance, mechanosensory cells in statocysts—gravity and balance organs—in ctenophores, and in mollusks have motile cilia (Grossman et al. 1979; Tamm 1982; Tamm 2014). The motility of the statocyst cilia in Nudibranchs was shown to aid the amplification, modulation and transmission of the mechanical signal (Stommel et al. 1980).

Active motility is also known to be required for signal non-linear amplification and frequency selectivity in sound-receiving organs such as the ear in vertebrates and the antennae in insects (Göpfert and Robert 2003; Göpfert and Robert 2001; Hudspeth 1997; Hudspeth et al. 2000). At least in the antenna of *Drosophila* this active mechanism is based on cilia motility of the auditory sensory neurons (Göpfert and Robert 2003; Karak et al. 2015). Interestingly, *Drosophila NOMPC* mutants show reduced oscillations in the antenna (Göpfert and Robert 2003), and loss of non-linear signal amplification, while *TRPV* mutants—which are also expressed in sound receptors—showed gained amplification (Göpfert et al. 2006). The co-expression of these molecules in MS neurons poses the question as to whether a similar amplification mechanism of a yet-to-be-identified stimulus is at play.

Based on preliminary experiments showing that the deflection of the cilium by laminar flow shows calcium increases in MS1 and MS2 (data not shown), it could be proposed that the signals sensed and amplified by MS neurons are minute changes in flow around the episphere. The direct innervation of MS neurons onto the motoneurons controlling the direction of swimming in the larva (Verasztó et al. 2018; Randel et al. 2014) may reflect how MS neurons control swimming direction based on these changes in flow. These changes may in turn be due to self-generated flow or to changes in the swimming orientation. Detection of changes in flow seems to be one of the main functions of CSF-cNs in zebrafish (Sternberg et al. 2018; Böhm et al. 2016), which are also mechanosensory cells with motile cilia. Flow detection has also been hypothesized to be the function of the PKD2/PKD1L1-expressing ciliated cells in the Kupffer's Vesicle of the fish medaka (Kamura et al. 2011). Thus, MS cells may eventually be added to the group of animal motile ciliated mechanosensory cells specialized to detect flow.

The evolution of mechanosensory cells in *Platynereis*

It is currently unclear what evolutionary relationship, if any, the CRs and the MS neurons have to other mechanosensory neurons. A way into elucidating such relationships is to compare their "molecular fingerprint" with cell types in other animals (Arendt 2005). So far, extensive expression data are lacking for CRs, although their morphological similarity to collar receptors in other annelids suggests that this cell type is conserved at least in this phylum. Based on other reports and on the *PKD2-1* expression pattern at the trochophore stage, it can only for now be concluded the CRs may develop within the conserved apical plate region demarcated by the expression of the

transcription factors *six3,* and *foxQ2* (Marlow et al. 2014). Defining with more precision the molecular fingerprint of these cells using the *PKD* genes as reference markers would be key for comparisons with other mechanosensory cells.

Although the sensory morphology of MS neurons has not been reported in other annelids, molecular data obtained for the trochophore stage suggests that this cell type may be conserved even in other phyla (Marlow et al. 2014). MS neurons were suggested to develop in the *six3+/foxQ2+* apical plate zone, and as previously mentioned, they also were found to express *foxJ*. Co-expression of these three transcription factors have been found in the apical plate of various invertebrate larvae (see references in Marlow et al. 2014), thus suggesting that MS neuron homologs may be an ancestral cell type in the apical organ of other animals. Although Marlow et al made the comparison between the multiciliated crescent cell in the apical organ of *Platynereis* and motile uniciliated cells in the apical organ of other larvae (Nezlin and Yushin 2004; Chia and Koss 1979), the cilium motility and the molecular fingerprint of MS cells rather suggests that they are the analogous or even homologous cell type to those other putative mechanosensory cells. More in-depth comparisons could be made with the planula larva of the Cnidarian *Nematostella*, where a (putative) motile ciliated sensory cells with similar molecular fingerprint to MS neurons have been identified (i.e. six3/*foxQ2*/ *foxJ*/*TRPV*+ cells) (Sinigaglia et al. 2015; Sinigaglia et al. 2013).

The proposed function for MS cells in controlling swimming direction may be also conserved in other invertebrate larvae. Sea urchin embryos have a motile apical tuft that by protein purification were found to be highly enriched for the glutathione transferase theta enzyme (Jin et al. 2013). When this enzyme is chemically inhibited, the apical tuft gets bent and the embryos have difficulty changing direction upon collision with an obstacle. Although more specific manipulations should be carried out to assign such function to the apical tuft, these results are suggestive of a mechanosensory function in larval navigation mediated by apical organs with motile cilia.

Chapter 4 The genetic analysis of the startle response in an ethological context

"…labeling something as behavior is contingent on the outcome of an investigation into what the animal does to ensure its survival in its native habitat".

Krakauer et al, 2017[15]

[15] (Krakauer et al. 2017)

Statement of contributions and publication status

SEM dry point fixation of wildtype and mutant larvae, as well as image acquisition was done by Jürgen Berger (MPI, Tübingen). Copepods and advice on predation experiments were contributed by Ruben Almeda (DTU, Denmark). The remainder of the work presented in this chapter was carried out by the author of this thesis. The figures were also assembled by the author.

The *PKD1-1,* and *PKD2-1* mutant lines, their startle response phenotype and the predation experiments were published elsewhere (Bezares-Calderón et al. 2018).

Introduction

Genetic analysis of startle responses

The use of genetic analysis and genetic tools has increased our understanding of the molecular and neuronal mechanisms driving behavior. Such approach is most powerful when combined with simple behavioral assays that can be reproduced in independent lines and in a large number of animals. Startle responses are amenable to genetic dissection due to their unambiguous behavioral patterns, the (in many cases) known wiring diagrams, and the simple stimulation assays needed to reliably evoke them.

The genetic analysis of the nose and body touch withdrawal responses in *C.elegans* has been useful both as a platform for the discovery of molecules important for mechanotransduction, and at the same time as a model to understand circuit mechanisms and how these are regulated (Chalfie et al. 2014). The genetic screen on touch-insensitive mutants has led to the discovery of the MeT channels MEC-4/MEC-10 (O'Hagan et al. 2005; Chalfie and Au 1989), and regulatory proteins such as the stomatin-like MEC-2 (Zhang et al. 2004; Goodman et al. 2002; Huang et al. 1995)(reviewed in (Schafer 2015)). The availability of uncoordinated and of mechanosensory insensitive mutants has also allowed to test hypothesis on the role that certain molecules and neurons play in the startle circuit. For instance, a combination of uncoordinated mutants and mutants for specific enzymes involved in neurotransmitter synthesis led to the implication of tyramine-expressing cells in the arrest of head oscillations that occur during the touch withdrawal response (Alkema et al. 2005). Similarly, mutations in the glutamate receptor showed its requirement for nose-touch withdrawal responses but not for triggering responses to other noxious stimuli, which are nonetheless both mediated by the same cell (Hart et al. 1995). The role of dopamine in habituation to mechanical stimulation has also been shown in the withdrawal response using genetic manipulation (Kindt, Quast, et al. 2007).

In other systems amenable to genetic analyzes, simple startle response assays have also been instrumental for the discovery of MeT channels and other associated proteins (Kernan et al. 1994; Nicolson et al. 1998). Due to pleiotropic effects, few studies use mutants to understand startle circuit function. Studies to interrogate startle circuits in these systems make use instead of genetic tools such as optogenetics or thermogenetics that can be targeted to defined subsets of neurons to manipulate their function in an acute manner (reviewed in (L. Luo et al. 2018; Luo et al. 2008)).

Genome editing with CRISPR

The genetic analysis of behavior has greatly benefited from new technological advances. Genome sequences are available for an ever-increasing number of organisms, and tools to manipulate gene function are becoming inexpensive and high in throughput. In this way forward and reverse

genetics approaches are becoming a reality in non-conventional organisms with interesting and unexplored behaviors, including those related to mechanical forces.

A site-directed genome editing tool was recently developed leveraging the natural sequence specificity of the prokaryotic antiviral defense system CRISPR (Clustered Regularly Interspaced Sequence Palindromic Repeats) (Jinek et al. 2012). In the most popular version of this system, a synthetic small guide RNA (sgRNA) couples to the endonuclease Cas9 to form a ribonucleoprotein complex that binds to a complementary sequence ~20bp long found in the genome to be edited. Once the RNP is bound to the targeted site, Cas9 makes a double-strand break in the DNA around the binding region (**Figure 4-1**). For the cut to take place the target sequence has to be flanked on one side by a 3 bases-long protospacer adjacent motif (PAM) sequence. The simplicity of this system has opened the possibility to do genetic analyses in any animal with a sequenced genome and a means of transformation in place.

The CRISPR/Cas9 editing system holds a great promise for research on non-standard aquatic/marine animals, as shown by the growing number of taxa where this technique has been successfully used, which includes cnidarians (Cleves et al. 2018; Ikmi et al. 2014; Momose et al. 2018; Sanders et al. 2018), crustaceans (Martin et al. 2016; Nakanishi et al. 2014; Kumagai et al. 2017), echinoderms (Lin and Su 2016), urochordates (Sasaki et al. 2014) and mollusks (Perry and Henry 2015)(reviewed in (Momose and Concordet 2016)).

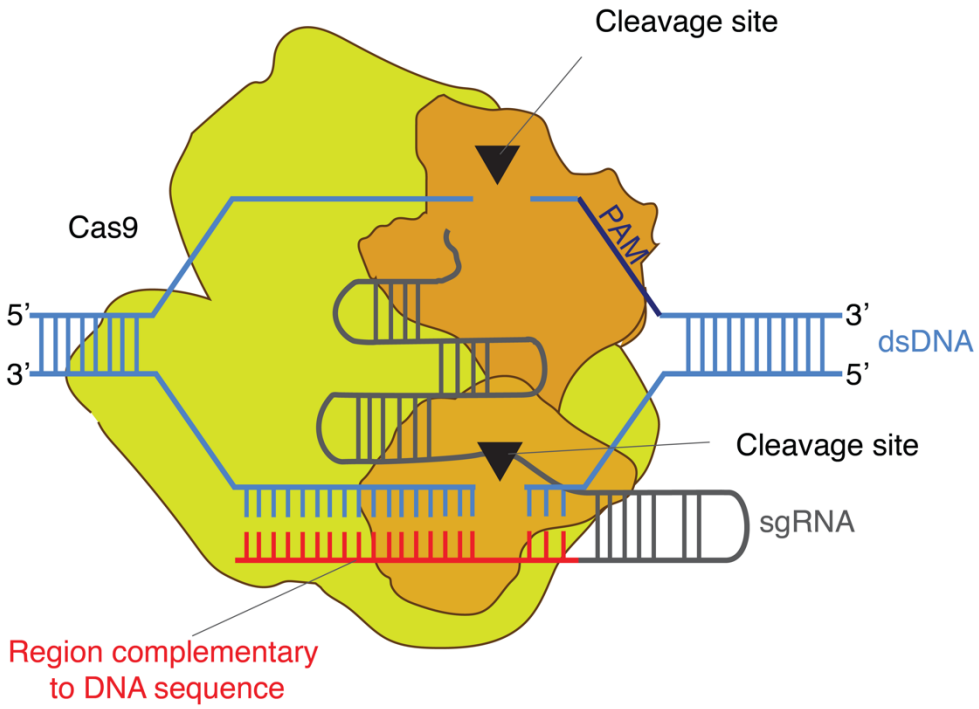


Figure 4-1 The CRISPR/Cas9 genome editing system. Schematic showing the mechanism of action of the Cas9-sgRNA ribonucleoprotein (RNP) complex. The Cas9-sgRNA complex unwinds the dsDNA and a complementary sequence in the sgRNA (red) anneals to one of the DNA strands. The annealing process and the recognition of other sequence motifs such as the protospacer adjacent motif (PAM) lead to the double strand cleavage in the unwound DNA (black arrowheads). Each cut is made by two distinct RNA cleaving domains in the Cas9 (colored in different shades of orange). Redrawn with permission from (Hsu et al. 2014).

So far, few studies have used CRISPR to address the genetic basis of behavior in these organisms, but recent work in cnidarians showcase how this tool can be used to understand marine ecology and behavior. For instance, in the larva of the anthozoan sea anemone *Nematostella,* the conserved neuropeptide GLWamide was mutated with CRISPR/Cas9 to generate frameshift mutations that caused a delay in the transition from larva to polyp, thus strengthening the conserved role of this peptide family in life-cycle transitions (Nakanishi and Martindale 2018). In the hydrozoan *Clytia hemisphaerica,* CRISPR/Cas9-mediated mutations in a gene coding for a cnidarian opsin expressed in gonad neurosecretory cells caused the loss of light-dependent oocyte maturation and spawning (Quiroga Artigas et al. 2018). This finding gives insight into the molecular mechanisms underlying the coupling between environmental conditions and timing of reproduction.

Startle responses in a natural context

Few studies have used genetics to dissect the startle responses in planktonic organisms. However, a sizeable number of studies have been performed to analyze the ecological relevance of startle responses in zooplankton. In most cases, these behaviors have been studied in the context of prey-predator interactions.

Zooplankton

A great variety of ciliates and flagellates have a jumping response against flow (mentioned in Chapter 1) that reduces its mortality rate relative to non-jumping ciliates when incubated with a predatory copepods generating feeding currents (Jakobsen 2001; Jakobsen 2002). The escape response in the rotifer *Polyarthra* (also alluded in Chapter 1) is displayed when a predatory copepod approaches it or when it is caught in the feeding current of the water flea *Daphnia.* This may explain the decrease in predation rate relative to other rotifer species lacking such fast escape response (Gilbert and Williamson 1978; Gilbert 1987). Similar predation experiments using a filter-feeding copepod with a selection of different rotifer species showed that those with escape or avoidance behaviors were less likely to be predated (Williamson 1987).

The type of startle response can determine the susceptibility to predation. In direct competition experiments, copepods with a fast escape response were less susceptible to predation than *Daphnia,* which has a slower escape response (Browman et al. 1989). Similar experiments show copepods escape more easily from barnacles than non-evasive brine shrimp (Trager et al. 1994). Other species of cladocerans use a sinking strategy to escape from predators. Upon an attack by a predatory copepod, *Bosminia* and *Chydorus* display a "dead man response" involving the cessation of all movement and the encroaching of the body, which results in passive sinking. This cladocerans may use the dead man response to reduce their hydrodynamic signal and thus make it more difficult for a rheotactic predator to relocate them (Kerfoot 1978).

The role of startle responses in predator avoidance has not been widely studied in other planktonic groups different to crustaceans. One of the few studies in ciliated larvae showed that sabellid larvae displaying a chaetae extension response are predated less than younger larvae that do not have chaetae (Pennington and Chia 1984). The authors conclude the chaetae play a defensive role, but this could be due either to the presence of the chaetae, or to the presence of the behavior, or to both factors.

Using genetics to analyze behavior in an ethological context

In most predation experiments in zooplankton, comparisons between species or between larval stages of the same species have been performed. Although direct observation of the prey-predator interactions supports a role of the startle responses in reduced predation, the behavioral versus the morphological component behind reduced predation cannot be separated. Both morphology and behavioral adaptations have been shown to be important for survival (Ohman 1988).

By reducing the number of possible interpretations of the predation results, genetic analysis of startle responses in a prey-predator context can help elucidate the features that are important for survival. In *C.elegans*, mutants that fail to suppress the head movements that accompany the touch withdrawal response fall prey more often to the ring traps of nematophagous fungi than wildtype worms (Maguire et al. 2011). This shows that even this subtle feature of the startle response is crucial for survival.

The use of genetics combined with an ecologically relevant stimulus can additionally reveal new features of the behavior. In *Drosophila* larvae, the use of a natural predator triggered additional range of escape behaviors besides the initial response seen with artificial stimulation (Robertson et al. 2013). In the same study, genetic inactivation of the nociceptors involved in one of these behaviors revealed these cells are required to trigger the escape response upon a predator attack, but it is not the main response affecting the survival of the larvae. In another study, naturalistic looming stimuli were obtained from real predator attacks to *Drosophila* flies and then used to elicit escape responses in intact flies or in flies impaired for fast escapes. These experiments showed that the few milliseconds between the fast and slow escape responses actually are relevant for increasing the survival rate (von Reyn et al. 2014).

Even when no genetic manipulations are involved, the study of startle responses considering the natural framework reveals the importance of the different responses and can help understand the underling circuitry. Exposure of crayfish to predatory dragonfly nymphs revealed that the fast escapes are used upon predator approach, while the more variable tail flips are preferentially used in the case the crayfish is captured (Herberholz et al. 2004). Finally, a recent report analyzes the startle behaviors in wild population of fish and shows that the escape responses seen in the

laboratory do occur in the wild, but the presence of other fish and the elements in the natural scene are also factors that determine the final escape response profile (Hein et al. 2018).

In this chapter, the CRISPR technique is used to generate frame-shift mutations in the three genes found to be expressed in in *Platynereis* CRs: *PKD2-1*, *PKD1-1*, and *NOMPC*. Previous studies have made use of other targeted genome editing tools in *Platynereis*, such as TALE or zinc finger nucleases (Gühmann et al. 2015; Bannister et al. 2014). It is shown here that CRISPR effectively induces mutations that can be transmitted to the next generation to end in stable mutant lines. The results analyzing the phenotype of the mutants in the context of the startle response are also presented. At the end of the chapter, the ecological relevance of the startle response is evaluated using the mutants generated.

Materials and Methods

Generation of mutant lines using CRISPR/Cas9

Cloning and mRNA synthesis of Cas9

The full ORF of the GFP-labeled *Streptococcus pyogenes* Cas9 (*Sp*Cas9-GFP) (Jinek et al. 2013) was PCR-amplified (see **Table 0-5** in Appendix) from plasmid pMJ920, (kindly provided by Jennifer Doudna, Addgene plasmid #42234), and cloned into the *Pdum*IVT plasmid using restriction cloning. From this plasmid *Sp*Cas9-GFP mRNA was synthesized as described in the Materials and Methods in the Appendix.

Identification and cloning of sgRNAs

Single guide RNAs (sgRNAs) targeting 18 or 20 nucleotides of the second exon of *PKD1-1*, *PKD2-1* and *NOMPC* were designed with the online tool ZiFiT (Sander et al. 2010). The sequence of each prospective target was blasted against *Platynereis* genome database to select oligos with ≥1 mismatch to any other sequence in the genome. One or two targets were selected for each gene (**Table 4-1**). Oligos were designed for each target and cloned into either plasmid DR274 (kindly provided by Keith Joung, Addgene plasmid #42250 (Hwang et al. 2013), or plasmid pX335-U6-Chimeric_BB-CBh-hSpCas9n(D10A) (kindly provided by Feng Zhang, Addgene plasmid #42335 (Cong et al. 2013)) using the general oligo annealing method (see Materials and Methods in the Appendix and **Table 0-5** for list of oligos).

Table 4-1 List of small guide RNAs (sgRNAs) used in this study. All sgRNAs were injected at 20ng/µl except for NOMPC_T3, which was injected at 50ng/µl.

Target gene	sgRNA ID	sequence (5'->3')	PAM	Length (bp)	Parent Vector (#Add gene)	sequence matching to other loci (5'->3') *	Matching off-target contig
PKD2-1	truT2	GGGCCACCCG TCAGACTG	GGG	18	#42250	xxxCCACCCGTCA GACTG	g_s292699
PKD1-1	truT1	TGTGTCAACTT TACCCCA	GGG	18	#42335	xTGTGTCAACTTT ACCCCA	g_sc_V2_ 141633
	T5	AGGCGGATAT CAGTGAAGCG	TGG	20	#42335	xxxCGGATATCAG TGAAGCx	g_sc_V2_ 28914
NOMPC	T3	TCTCCCTCACA GCCAGCATC	AGG	20	#42335	xCTCCCTCACAGC CAGCxxx	scaffold62 8500

*Sequenced blasted against *Platynereis* genome assembly (courtesy of Detlev Arendt): AllPduGenomic 2,953,208 sequences; 8,390,881,062 total letters.

sgRNA synthesis

To produce sgRNAs from these plasmids, PCR templates were obtained either with the T7 primer or with an oligo-specific primer and with the reverse primer used for general *in vitro* transcription (see **Table 0-5** for primer list). The following PCR program was used:

Temperature	Time	Repeat
98°C	1.5 min	1 time
98°C 50°C 72°C	10 sec 30 sec 7 sec	35 times
72°C 10°C	10 min ∞	1 time

The fragments were purified with the QiaQuick PCR purification kit (Qiagen) and sequenced using the T7 primer and the reverse primer used for amplification. Ca. 1µg template was used for *in vitro* transcription with the MEGAshortscript T7 kit (AM1354, Ambion/ThermoFisher Scientific) according to the manufacturer's instructions. Transcription reactions were run overnight at 37°C. sgRNAs were purified with the MEGAclear Transcription Clean-up kit (AM1908, Ambion) according to the manufacturer's instructions, eluted with nuclease-free water, aliquoted and stored at -80°C. RNA integrity was verified by gel electrophoresis and the concentration measured from a 1:10 dilution using a spectrophotometer.

Microinjection of Cas9-GFP/sgRNA mix

*Sp*Cas9-GFP mRNA and one or more sgRNAs were combined in a single solution prior to injection from freshly thawed -80°C stocks and injected at 300, and 20 or 50ng/µl, respectively (**Table 4-1**) into *Platynereis* zygotes (Tübingen WT strain). Healthy green fluorescent larvae were selected at 24 hours post-fertilization (hpf) using an Axiozoom stereoscope (Carl Zeiss GmbH, Jena), and a fraction of them genotyped, while the rest was cultured at 7dpf if evidence of Cas9-mediated gene editing was obtained by genotyping (see below).

Extraction of genomic DNA

genomic DNA (gDNA) of single larva, or of batches of 2 or 40 larvae was extracted by transferring them to 5, 10 or 20 µl of QuickExtract DNA Extraction solution 1.0 (Epicentre), respectively. The sample was then vortexed for 15-20 seconds, incubated at 65°C for 6 minutes in a thermocycler, vortexed for 15 seconds more, and finally incubated at 98°C for 2 minutes. Samples were stored at -20°C until further use.

gDNA of adult worms (older than 2 months) was obtained as follows: animals were taken out from their tubes, rinsed in mixed sea water (MSW) and anesthetized in a 1:1 mixture of 7.5% (w/v) MgCl2 (dissolved in NSW): MSW (Bannister et al. 2014). The tail or pygidium of the anesthetized animal was cut with a scalpel and transferred to 20 µl QuickExtract solution and its gDNA extracted as described above. The sampled worms were then transferred to non-treated plastic 6-well plates (NuncTM, cat# 150239, ThermoFisher Scientific) and filled with MSW for recovery.

After determining their genotype, the worms were cultured in boxes sorted by genotype, or cultured individually in glass beakers.

Genotyping by PCR

Loci to be genotyped were amplified from gDNA in 20 µl (see Materials and Methods in the Appendix for PCR recipe and components) using a defined set of primers (see **Table 0-5** for primer sequences) and the following PCR program:

Temperature	Time	Repeat
98°C	3 min	1 time
98°C	10 sec	35 times
66°C	20 sec	
72°C	x	
72°C	10 min	1 time
10°C	∞	

x=45 sec for *PKD2-1* and *NOMPC*, 2 min for *PKD1-1*

The PCR amplicons were analyzed by agarose gel electrophoresis, the reactions diluted with one volume of water and directly sequenced with a specific nested primer (see **Table 0-5** for sequencing primers list) using Sanger sequencing. Evidence of successful genome editing was manually assessed in the chromatogram by a sudden drop in sequence quality in the targeted region.

Touch startle assay

Nectochaete larvae from heterozygote crosses were collected by phototaxis and transferred to a small petri dish. Larvae selected blindly (i.e. without watching through the stereomicroscope) were transferred in 40 µl NSW to a glass slide placed under a stereomicroscope. Immediately after being transferred, the larva was touched with a manually-held fine tungsten needle tool (RS-6063, Roboz) on the anterior side. Only swimming larvae were assayed. The larva was touched 2-3 times more if failed to get startled to verify lack of response. The results were recorded in paper at the moment of the experiment. For easier interpretation, only two categories were defined: responder or non-responder larvae.

After being assayed, larvae were transferred to 4 µl QuickExtract solution for gDNA extraction. PCR genotyping was performed as indicated in the preceding section. Genotype was assessed manually by reconstructing the genotype of each sample from the chromatogram. The *PKD2-1*$^{\Delta 137}$ allele was assessed in most cases by gel electrophoresis as the band could be separated from bands of other alleles.

The results were plotted in R using a custom-written script deposited in GitHub[16].

[16] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig3I-J_Barplots_mutants.R

Vibrating filament startle assay

Phototactic larvae with the predicted mutant genotypes *PKD1-1*$^{i1/i1}$ or *PKD2-1*$^{mut/mut}$ were tethered with glue as described in Materials and Methods in Chapter 1 for the kinematics experiments. The same stimulation equipment and protocols were used. Probe speed was assessed in the same way. Each tested larva was genotyped after the experiment to verify the homozygote (or trans-heterozygote) genotype. All the measurements were deposited in GitHub[17].

The results were analyzed and plotted in R using a similar pipeline as that described in the script used to analyze the kinematics of the startle response in WT larvae (`FigStartleKynetics.R`). The script was deposited in GitHub[18].

Wild type and mutant nectochaete larvae for CR morphology assessment

Crosses of homozygote or trans-heterozygote *PKD1-1* or *PKD2-1* worms were set in parallel to wildtype crosses. Batches resulting from both wildtype and mutant crosses were incubated under equal conditions for 72 hours at 18°C (i.e. until they reach the nectochaete larval stage).

SEM

Samples were processed for SEM in parallel using the protocol described in Chapter 2. *PKD2-1*$^{mut/mut}$ and its wild type control were fixed in 3% Glutaraldehyde for 3 days, and *PKD1-1*$^{i1/i1}$ and its wild type control for 1 month. Sibling larvae of those processed for SEM were genotyped as described above. SEM micrographs were taken at the same magnification for mutant larvae and wildtype controls. ≥3 larvae of each genotype were analysed.

Cilia length

Data acquisition

To mount the larvae, an object slide with spacers was prepared as follows: two bands made each of 2 layers of Tesa® tape (Tesa GmbH) were stick to the object slide leaving a gap of 15 mm between each band. An individual larva was added to the slide in 45 µl. 5 µl of 1M $MgCl_2$ were gradually added to the drop containing the larva while mixing. A coverslip was carefully placed on top. A stack of the episphere of the larva was acquired with an AxioImager microscope (Carl Zeiss GmbH) using bright field illumination and DIC alignment. The stack acquisition was automatized using the AxioVision software (Carl Zeiss GmbH) to keep the magnification, resolution, Z-step speed and lighting conditions constant. Both mutant and age-matched wildtype larvae were imaged on the same day. Mutant genotype of each larva was verified by sanger sequencing.

---

[17] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip (Folder: Kynematics_measurements/)

[18] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig3K_S7U_Analysis_startledataPKDmutants.R

Data analysis

Cilia length using the measure tool of ImageJ implemented with the macro `CiliaLengthMeasurementwithrandomization.ijm` (see code in Appendix). To make the measurements blind for the genotype of each larva, the script randomly opens the list of files that includes wildtype and mutant files (each file is previously labeled only with a number code to avoid revealing the genotype of the larva). Each list of measurements was used as input for the script `CiliaStats.R` (see code in Appendix).

Predation experiments

Centropages typicus culture

Cultures of the copepod *Centropages typicus* were obtained from a continuous culture at the National Institute for Aquatic Resources (Technical University of Denmark, DTU) by Dr Rodrigo Almeda. His description of the culture method is quoted below:

"Specimens of *C. typicus* were originally isolated from zooplankton samples collected in the Gullmar fjord (Sweden) by vertical tows with plankton nets (500 μm mesh). Cultures of *C. typicus* were kept in 30 L plastic tanks with sterile-filtered seawater (FSW, salinity 32 ppt), gently aerated, at 16 ± 1 °C in dark. Copepods were fed *ad libitum* with a mix of phytoplankton (the cryptophyte *Rhodomonas sp.*, the diatom *Thalassiosira weissflogii* and the autotrophic dinoflagellates *Heterocapsa triquetra*, *Prorocentrum minimum* and *Gymnodinium sanguineum*) and with the heterotrophic dinoflagellate *Oxyrrhis marina*. Phytoplankton cultures were kept in exponential growth in B1 culture medium and maintained at 18°C and on a 12:12-h light/dark cycle in glass flasks. *O. marina* was fed the cryptophyte *Rhodomonas salina* and maintained at 18°C in 2-L glass bottles."

Experimental setup

C. typicus shipped from Denmark were kept in the same room as the worm cultures (18°C) for maximum 1 week before the experiment and daily fed with *O. marina*. Adult stages of active male and female copepods were individually selected with a pipette and transferred to a beaker filled with NSW without food for 2 hr prior to the experiment. 25 phototactic larvae of each genotype (25 age-matched wild type and 25 *PKD2-1*$^{mut/mut}$) were transferred to an experimental container that consisted of a Ø7 cm glass beaker of 250ml capacity filled with 200ml NSW and wrapped in aluminium foil to create a dark environment. Care was taken to accurately count the larvae and avoid losing them during transfer. The copepods were transferred afterwards (5-10 per container) and the container covered with more foil. An additional container with 25 wildtype and 25 mutant larvae but without copepods was prepared to control for mortality not related to predation. Virtually all experiments (except 1 batch) were run in duplicate or triplicate at 18°C for 12h or 24h without shaking (Almeda et al. 2017).

At the end of the incubation period, both dead and alive copepods were recovered, and their number registered. Survivor larvae were counted as follows: the incubation water was poured onto a large petri dish placed under a stereomicroscope, where individual larvae were located and directly transferred to 4 µl QuickExtract solution for gDNA extraction. The water was poured back and forth to the experimental container until no further larvae were found. Only healthy-looking larvae were counted. Larvae were genotyped as described above.

Data analysis

The number of survivor larvae of each genotype for each experiment as well as the incubation times and the number of copepods used were recorded in a tabular form. From this information, predation rates (*I*) for each genotype were calculated according to the following formula:

$$I = \frac{C_i - C_f}{NxT}$$

taken from (Almeda et al. 2017), where C_i and C_f are initial and final larva concentrations (larvae/L), respectively, *N* is number of alive copepods at the end of experiment and *T* is incubation time in days. The concentration of prey was determined based on the initial incubation volume (0.2 L).

Statistical analysis

The null hypothesis of equal predation rates was tested against the alternative of a higher predation rate of mutant larvae with a non-parametric one-sided exact Wilcoxon-Pratt signed rank test. The test was implemented using a script written in R (`Fig3M_BoxplotsPredRates.R`, see code in Appendix). The results were also plotted using this script. The data used as input for this script is available at GitHub[19].

Figure assembly

All plots were generated in R and edited in Adobe Illustrator CC (Adobe Inc.).

[19] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip (Folder Predator_assay).

Results

The role of polycystin genes PKD1-1 and PKD2-1 in the startle response

To get a deeper insight into the molecular and cellular mechanisms by which CR neurons detect water-borne vibrations, the role of the genes found in the preceding chapter to be expressed in these cells was investigated. These genes include the polycystin-related molecules *PKD1-1* and *PKD2-1* and the MeT channel *NOMPC,* albeit the expression of this gene in CRs was only determined by WMISH.

Generation of frame-shift mutations in PKD1-1, PKD2-1 and NOMPC with the CRISPR/Cas9 system

To address the genetic contribution of *PKD1-1, PKD2-1* and *NOMPC* to the startle response, loss-of-function mutations in each of these genes were generated using the CRISPR/Cas9 genome editing system (Jinek et al. 2012). sgRNAs ranging in size between 18-20 nt were designed to target the first exons of the genes of interest, co-injected with mRNA encoding SpCas9-GFP (Jinek et al. 2013) and mutations identified by Sanger sequencing (**Table 4-1**). sgRNA/Cas9-induced mutations were recovered for all the three genes targeted at different efficiencies.

In *PKD2-1* a single sgRNA targeted to the second exon generated multiple deletion alleles (**Figure 4-2**A). Three of these deletions (4, 5 and 14 bp) caused frame-shifts in the reading frame that in turn created premature stop codons soon after the mutation site and before the 6 transmembrane (TM) domains (see Version 1 of the corresponding mutant sequences in the Appendix) (**Figure 4-2**B). However, alternative start codons were found in the resulting sequences that could in theory still produce proteins with the 6 TM domain and C terminus almost intact (see Version 2 sequences in the Appendix). The same sgRNA induced a deletion spanning the downstream part of the targeted exon as well as the exon/intron boundary, and part of the second intron (**Figure 4-2**A). Given that the splicing signal (the GT sequence) in this allele was destroyed, and that correct splicing may no longer occur, it is uncertain if the protein could nevertheless be produced from this locus (see Version 1 of predicted sequence in Appendix). Assuming that splicing of the second exon can still occur, then the resulting protein will still be in frame, albeit with the first TM domain deleted (see Version 2 of predicted sequence in Appendix). Thus, the alleles generated in *PKD2-1* are potentially loss-of-function alleles, provided that the alternative start codons are not used *in vivo*. For convenience, the alleles generated with this sgRNA are called as a whole *PKD2-1*^{mut}.

Two sgRNAs targeted to different regions of the second exon of *PKD1-1* induced each a single base modification, a deletion and an insertion (**Figure 4-2**C). Therefore, like the mutations generated in the *PKD2-1* locus, these mutations changed the reading frame of the gene with a predicted stop codon occurring few residues downstream of the modification (see Version 1 of

predicted sequence in Appendix). Such truncated proteins lack most of the conserved domains in the protein and thus would not be functional (**Figure 4-2**D). But as in the mutant *PKD2-1* alleles, putative alternative start codons found in the mutated genes could still produce a functional protein (see Version 2 of predicted sequence in Appendix).

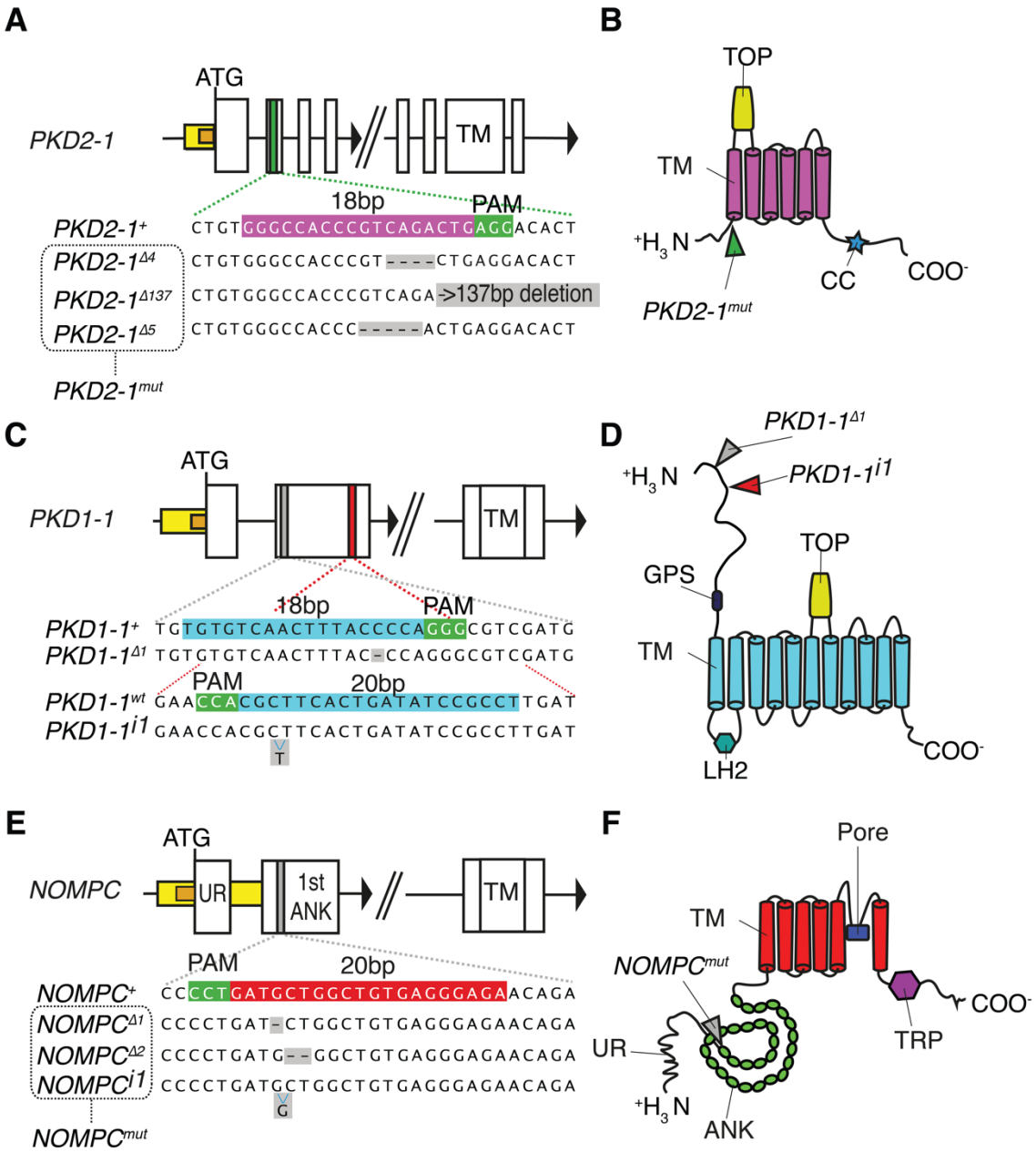


Figure 4-2 Generation of frameshift mutations in *PKD2-1*, *PKD1-1* and *NOMPC* with CRISPR/Cas9. (**A**) Top: Schematic view of the *PKD2-1* locus. Bottom: Wildtype and mutant sequences around the 18 bp target region (magenta, PAM sequence in green) located in the second exon. *PKD2-1*^{mut} is used to refer to any of the three alleles generated. (**B**) PKD2-1 secondary structure. The green arrowhead points to the approximate site of the mutations. (**C**) Top: Schematic view of the *PKD1-1* locus. Bottom: Wildtype and mutant sequences around the 18bp and 20bp target regions (cyan, PAM sequence in green) located in the second exon. (**D**) PKD1-1 secondary structure. The grey and red arrowheads point to the approximate site of the mutations. (**E**) Top: Schematic view of the *NOMPC* locus. Bottom: Wildtype and mutant sequences around the 20bp target region (red, PAM sequence in green) located in the second exon (encoding the 1st ankyrin repeat. *NOMPC*^{mut} is used to refer to any of the three alleles generated. (**F**) NOMPC secondary structure. The grey arrowhead points to the approximate site of the mutations. In A, C, and E: white boxes represent exons and lines introns. Gap in genomic sequence is indicated by a double slash. ATG: translation initiation site, Yellow box: promoter fragment, Orange box: 5'UTR, TM: transmembrane domain. Size of boxes is only approximately to scale. See **Figure 3-8** and **Figure 3-11** for meaning of protein domain abbreviations.

Finally, a single sgRNA targeted against *NOMPC* in the region encoding the first Ankyrin repeat domain generated three frameshift allelic variants (**Figure 4-2**D). All three alleles generated premature stop codons, and thus the resulting proteins would only retain the non-structured N-terminal element (if this fragment is indeed part of the final polypeptide) (**Figure 4-2**E; see Version 1 of predicted sequence in Appendix). Alternative start codons were also found in all the mutant sequences, the three of them conserving the TM domains, but *NOMPC*$^{\Delta 2}$ and *NOMPC*^{i1} would still lack the first three conserved Ankyrin repeats (see Version 2 of predicted mutant sequences in Appendix).

Injected worms carrying the mutations described above were backcrossed to reduce off-target effects, after which heterozygote and homozygote lines were established. All homozygote lines were fertile with no gross morphological defects at the larval stages.

PKD1-1 and *PKD2-1* mutant larvae, but not *NOMPC* show severe defects in the startle response elicited by touch stimulation

The effect of these mutations on the ability of nectochaete larvae to display the startle response was then investigated. To phenotype mutant larvae, touch stimuli were used as surrogates for vibrations, since touch also triggered the startle response in freely swimming and tethered larvae with no visible differences to the response induced by vibrations (data not shown). Specifically, a thin filament was used to touch a single larva in the anterior end and counted as responsive if this stimulus triggered the elevation of parapodia. In some cases, larvae reacted even before being touched (as expected from a response tuned to hydrodynamical stimuli), but the response was only scored after direct contact. To avoid any bias from the experimenter, each larva was genotyped only after being phenotyped (see Materials and Methods).

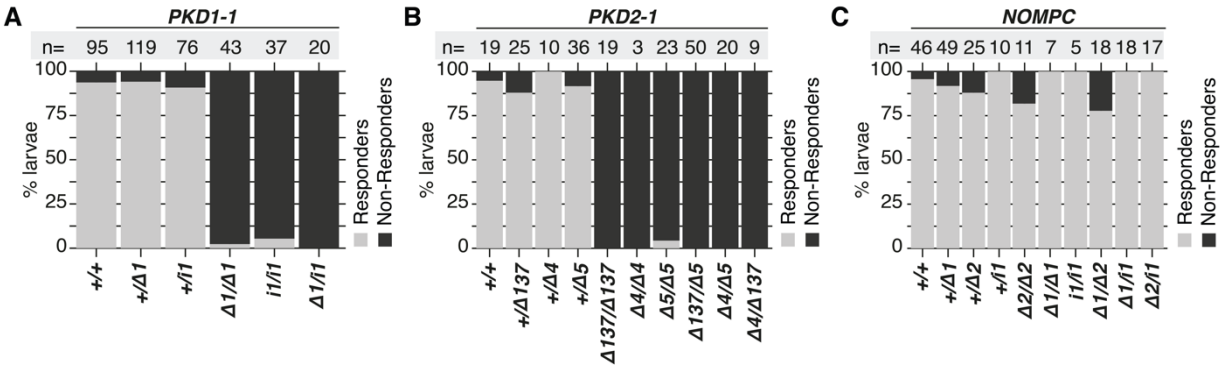


Figure 4-3 *PKD1-1,* and *PKD2-1* homozygote mutant larvae do not display the startle response upon touch stimuli. (**A**-**C**) Stacked bar plots of the percentage of larvae responding (gray bars) or not responding (black bars) to touch. (**A**) Percentage of responding and non-responding *PKD1-1* larvae heterozygote, homozygote and trans-heterozygote for *⊿1* and *i1* alleles. (**B**) Percentage of responding and non-responding *PKD2-1* larvae heterozygote and homozygote for *⊿4*, *⊿5*, and *⊿137* alleles, as well as combinations thereof. (**C**) Percentage of responding and non-responding *NOMPC* larvae heterozygote and homozygote for *⊿1*, *⊿2*, and *i1* alleles, as well as combinations thereof. Number of phenotyped larvae in A to C is indicated above each corresponding bar. Larvae counts were pooled from 18 (*PKD1-1*), 14 (*PKD2-1*), or 13 (*NOMPC*) batches. WT and heterozygote larvae analyzed are siblings of homozygotes or trans-heterozygotes here analyzed.

The startle response upon touch stimulation was completely abolished in virtually all *PKD1-1*$^{\lrcorner 1/\lrcorner 1}$ and *PKD1-1*$^{i1/i1}$ larvae (97.7% and 94.6%), while it was unaffected in most of their heterozygote and wild-type siblings (>90% in all cases) (**Figure 4-3**A). The two alleles did not complement each other, as evidenced by the failure of all *PKD1-1*$^{\lrcorner 1/\ i1}$ trans-heterozygote larvae to respond to the stimulus.

This drastic effect on the startle response was also observed in *PKD2-1* mutant larvae (**Figure 4-3**B). Most or all *PKD2-1* mutant homozygotes and trans-heterozygote larvae failed to get startled upon touch stimulation (**Figure 4-3**B). Non-responding larvae in wild-type and heterozygote controls was much lower (12% in *PKD2-1*$^{wt/\lrcorner 137}$ larvae was the highest value observed). This result strongly suggests that the phenotype in the startle response is specifically due to in-target mutations. Mutations in *NOMPC* did not cause any effect on the touch-induced startle response (**Figure 4-3**C). *NOMPC*$^{\lrcorner 1/\lrcorner 1}$, *NOMPC*$^{i1/\ i1}$ and *NOMPC*$^{\lrcorner 2/\lrcorner 2}$ mutant larvae were for the most part unaffected, with the last genotype showing the highest percentage of non-responders (18%). Combinations of these alleles did only show a slight increase in the percentage of non-responders (22% of *NOMPC*$^{\lrcorner 2/\lrcorner 1}$ were non-responders). Thus, these results do not support a crucial role of this gene in the startle response at this stage.

PKD1-1 and *PKD2-1* mutant larvae do not respond to hydrodynamic stimuli

The touch assay uncovered a role of *PKD1-1* and *PKD2-1* in the startle response upon touch stimulation. However, with this crude assay it was not possible to determine if the response was completely abolished, or if only specific features of it were absent. To address this possibility as well as to more directly compare to results obtained in wild-type larvae using vibrational stimulations, individual *PKD1-1* or *PKD2-1* larvae were tethered to a glass-bottom dish and stimulated with a vibrating filament using a range of stimulation values (see Chapter 1).

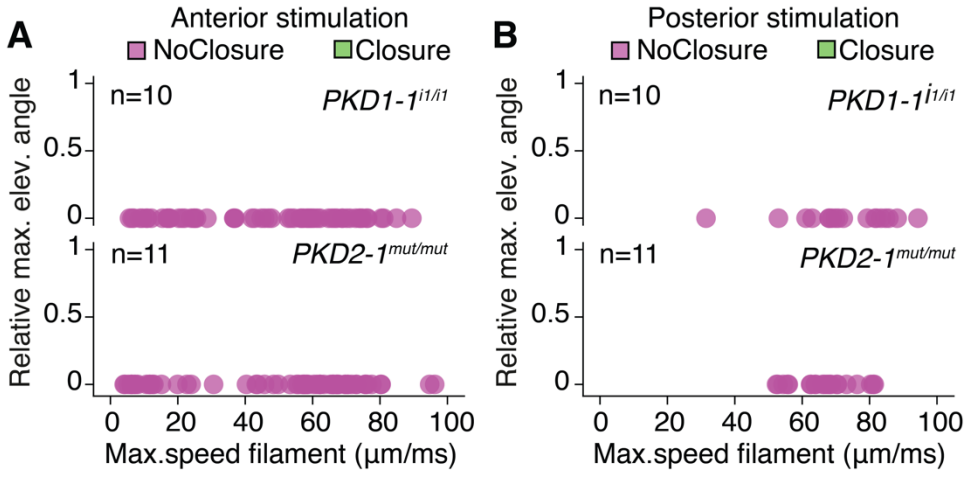


Figure 4-4 *PKD1-1* and *PKD2-1* homozygote/trans-heterozygote mutants do not display closures or parapodial elevation upon hydrodynamic stimuli. (**A**-**B**) Scatterplot of fractional maximal parapodial angles observed in *PKD1-1*$^{i1/i1}$ and *PKD2-1*$^{mut/mut}$ larvae stimulated with a filament placed 100µm from the anterior (A) or from the posterior (B) end. Observations are colored according to the effect the stimulus had on the prototroch (closure vs no closure). Closures and Parapodial elevation responses were virtually absent in these observations.

None of the *PKD1-1* or *PKD2-1* mutant larvae tested showed any trace of parapodial elevation response (either LowE or WideE types) or of induced ciliary band closures upon anterior stimulation (**Figure 4-4**A). Even stimulating the anterior end with filament speeds up to 89 μm/ms or 100 μm/ms, respectively, failed to trigger any response; these values robustly elicited ciliary closures and a parapodial elevation response in wild-type larvae (**Figure 1-3**B, E). Similarly, when the same larvae were stimulated from the posterior, neither elevation events nor ciliary arrests were observed at strong levels of stimulation (**Figure 4-4**B; compare to **Figure 1-5**B-C).

These results, together with those obtained from the touch stimulation assay, strongly suggests the mutations generated in both *PKD2-1* and *PKD1-1* abolished the expression of functional products, or drastically compromised their function. The total lack of a startle response in otherwise normal larvae suggests these genes are playing an essential and specific role in the behavior. The expression analysis performed in Chapter 3 indicated that the group of cells in which both genes are expressed is limited to CRs and pygPB^{unp}. Thus, malfunctioning or absent PKD1-1 or PKD2-1 in these cells may be causing the overt lack of response to mechanical stimuli.

CR neurons have a normal sensory morphology in PKD mutants

Morphology of hCR1/hCR2 microvilli collars is normal in PKD1-1 and PKD2-1 mutant larvae

The inability of *PKD1-1* and *PKD2-1* mutant nectochaete larvae to get mechanically startled could be due to gross developmental abnormalities in the cells where these genes are expressed, or to more subtle defects affecting the mechanotransduction cascade. To address this issue, the sensory morphology of CR neurons (specifically of hCR1 and hCR2 as they were shown to respond to hydrodynamic stimuli) in *PKD1-1* and *PKD2-1* homozygote mutants and in age-matched wild type controls was compared by SEM.

hCR1 and hCR2 neuron cilia were present and in a normal arrangement in both *PKD1-1* and *PKD2-1* mutant larvae (**Figure 4-5**F,G,P,Q). At the SEM level, wildtype larvae used for the comparison showed two main collar morphologies: an "open" collar, where the microvillar structure was visible (**Figure 4-5**D, N), and a "closed" collar where the microvilli appeared covered by a thin (cuticular) layer (**Figure 4-5**E, O). The latter may be the normal structure and the former only a case where the cuticular layer is broken. Both conformations were also observed in the mutants (**Figure 4-5**I,J,S,T). In some mutant larvae, the collars did not look as round and well-formed as in the WT. It is important to note these observations were neither made under a blind experimental design, nor any morphological parameter was quantified. In conclusion, no qualitative differences in hCR1 and hCR2 sensory morphology were evident by SEM in *PKD1-1* and *PKD2-1* mutants when compared to wildtype samples.

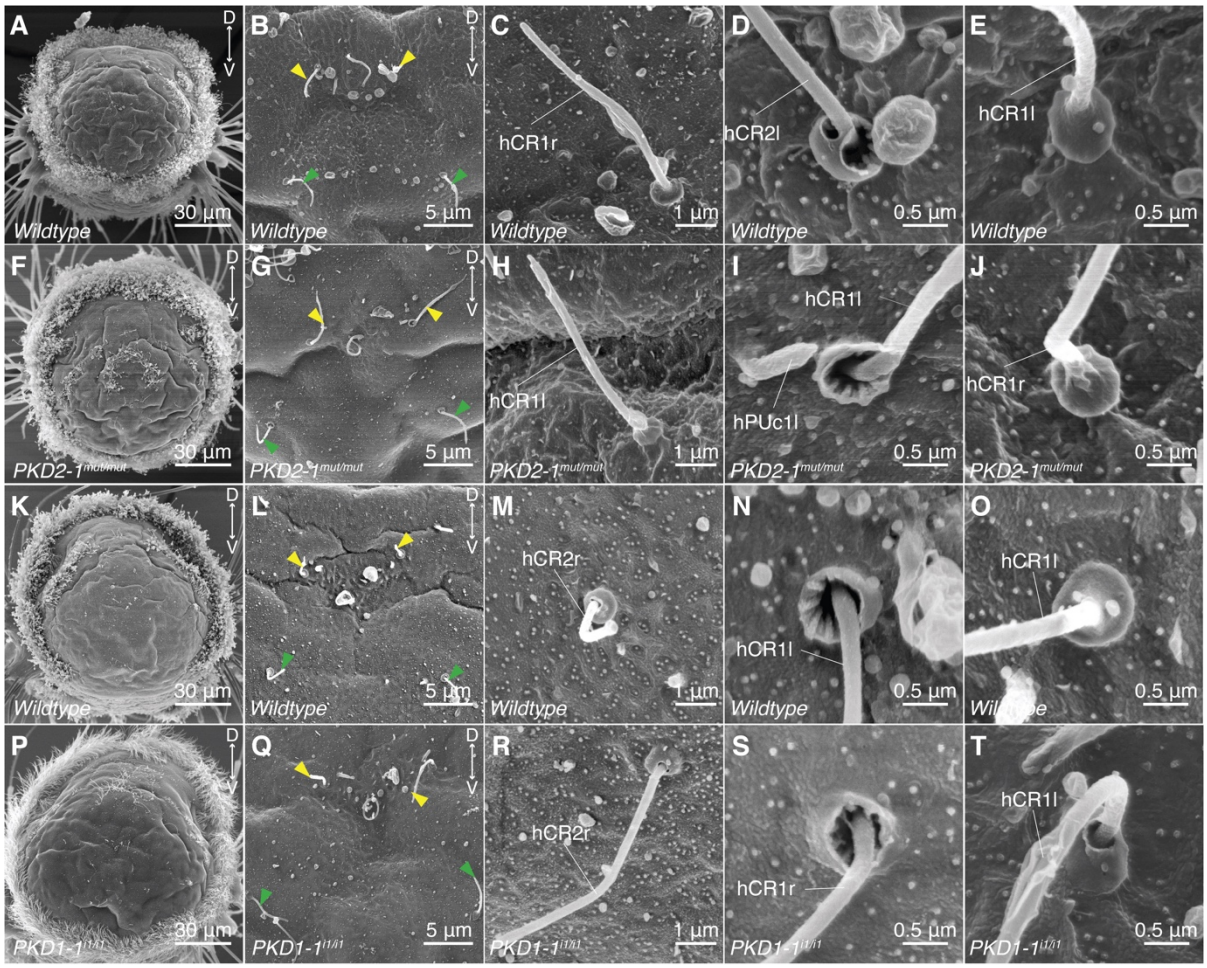


Figure 4-5 CR sensory cilia in *PKD1-1* and *PKD2-1* homozygote mutants do not show obvious developmental defects. (**A**-**J**) SEM micrographs of age-matched wildtype (A-E) and *PKD2-1* $^{mut/mut}$ (F-J). (**K**-**T**) SEM micrographs of age-matched wildtype (K-O) and *PKD2-1* $^{il/il}$ (P-T). Apical overview of nectochaete larvae in A, F, K, and P. Close up view of episphere in B, G, L, and Q, to show configuration of hCR1 cilia (yellow arrowheads), and hCR2 cilia (green arrowheads). Close-up views showing a complete cilium (C, H, M, R), an "open" collar with visible microvilli (D, I, N, S), or a "closed" collar with a cuticular layer covering it (E, J, O, T).

hCR1 neurons have a longer cilium in PKD2-1 mutants

Length is one of the morphological features of a cilium that is most susceptible to be quantified, and that is tightly regulated by the cilium assembly machinery (Ishikawa and Marshall 2011). Length is also regulated by mechanosensory processes, in which Polycystins have been found to play a role (Besschetnova et al. 2010).

The length of the cilia found in the larval episphere of *Platynereis*—belonging to MS1, MS2, hCR1 and hCR2—was measured by bright field microscopy in *PKD2-1* mutants and age-matched wildtype controls. A first measurement was done from larvae whose genotype was known (Materials and Methods). A significant difference was observed only in the length of hCR1 cilium (**Figure 4-6**A).

.

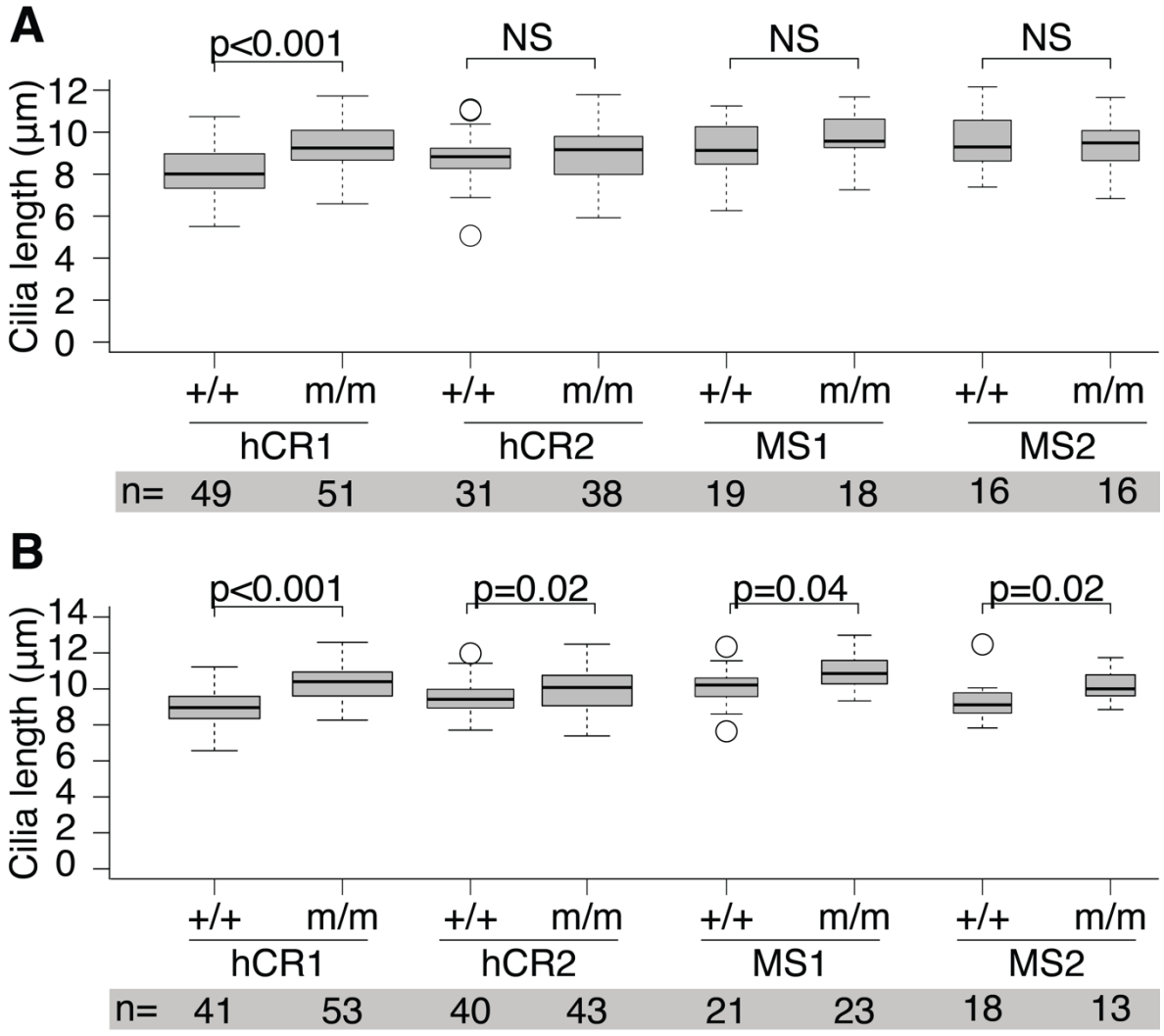


Figure 4-6 Comparison of cilia length of hCR1/2 and MS cells between wildtype and *PKD2-1*$^{mut/mut}$ reveals significant differences in cilium length in randomized measurements. (**A**-**B**)Tukey boxplots of cilia length measured from bright field images in (A) pre-sorted measurements or (B) randomized measurements. A two-sided Kolmogorov-Smirnov statistic was used to compare distributions. +/+: wildtype genotype; m/m: *PKD2-1*$^{mut/mut}$ mutant genotype. Box width as defined in **Figure 1-4**.

To rule out that this difference was due to a bias introduced by the prior knowledge the experimenter had on the genotype of the larva being measured, the videos were processed in randomized order and the identity of the genotype was hidden. The significant difference in length of the hCR1 cilium between wildtype and *PKD2-1* mutant larvae was still obtained (**Figure 4-6**B). Intriguingly, now the p-value of the difference in length for the remaining cilia was significant at a 0.05 level. Although these results suggest that hCR1 cilia in *PKD2-1* mutants are slightly but significantly longer than normal, the trend to observe longer cilia in the mutants in cells not expressing this gene underscores the need to rule out other indirect differences affecting this result.

PKD2-1 mutant larvae are more susceptible to predation by copepods

PKD1-1 and *PKD2-1* mutant nectochaete larvae have an overall normal morphology, yet they are unable to display the startle response. It thus seems that their main defect is an insensitivity to hydromechanical stimuli, and thus their failure to get startled by such stimulation. Startle responses

have evolved in many instances as a way to escape and deter predators (Sillar et al. 2016). To assess how important is the startle response in *Platynereis* nectochaete larvae for survival in the presence of a predator and taking advantage of the seemingly specific phenotype of PKD mutant larvae, a direct survival competition was performed. Namely, *PKD2-1*$^{mut/mut}$ larvae and age-matched wildtype were incubated together with *Centropages typicus* (**Figure 4-7**A), a predatory copepod that uses both suspensivorous and ambush tactics to capture its prey (Calbet et al. 2007). *PKD2-1* mutant larvae were more readily available than *PKD1-1* mutants at the time this experiment was designed.

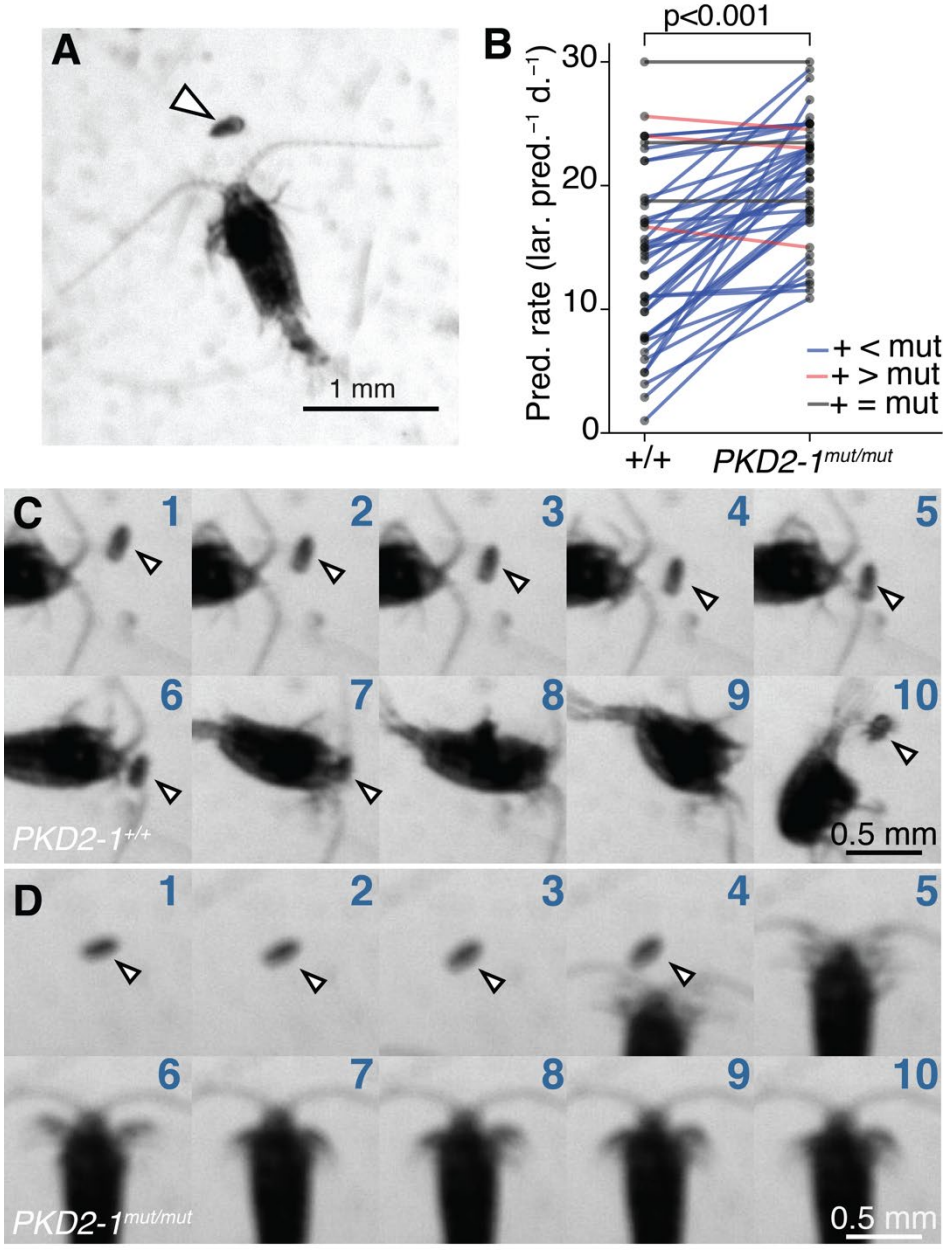


Figure 4-7 PKD2-1 mutant nectochaete larvae are more susceptible to predation by the copepod C. typicus than age-matched wildtype. (**A**) Snapshot showing an adult Centropages typicus female (in the center) and a Platynereis nectochaete (arrowhead). (**B**) Predation rates of wildtype (+/+) and *PKD2-1*$^{mut/mut}$ larvae in direct competition experiments. Paired values are joined by blue, grey or red lines if predation rates were higher, equal or lower in mutant than in wildtype larvae, respectively; data from 42 trials with 12 batches. One-sided exact Wilcoxon-Pratt signed rank test, P = 5.2e-10. (**C**-**D**) Time series (each numbered from 1 to 10) showing wildtype (C) or PKD2-1$^{mut/mut}$ (D) nectochaete larvae being attacked by *C. typicus*. Arrowheads indicate the position of the larva. Note the larva has the chaetae extended in C10. Note the mutant larva did not escape the attack of the predator.

These direct competition experiments revealed a statistically significant increase in the predation rate of *PKD2-1*$^{mut/mut}$ larvae over the co-incubated wildtype counterparts (**Figure 4-7**B). The results were obtained using multiple batches and different incubation times. Only in three out of 42 trials the predation rate was higher in wildtype larvae. In the negative controls (i.e. larvae without predators), 86% or more larvae survived, while less than 70% of the larvae survived when incubated with copepods[20].

Encounters between wildtype or mutant larvae and *C. typicus* were also captured in time-lapse recordings (**Figure 4-7**C-D). The recording presented in **Figure 4-7**C shows that although the copepod is able to locate and attack the wildtype larva, the larva displays the startle response before it is being captured (**Figure 4-7**C6), which may have contributed to cause the aversive reaction of the copepod (**Figure 4-7**C9) that leads in the end to the larva being released (**Figure 4-7**C10). In contrast, the *PKD2-1*$^{mut/mut}$ larva does not react to the approach of the copepod (**Figure 4-7**D). Thus, if it is assumed the major difference between *PKD2-1*$^{mut/mut}$ and wildtype larvae is the lack of a startle response in the former, it can be concluded that such behavior is important to escape and/or deter predators like *C. typicus*.

[20]Dataset available at https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip

Discussion

The CRISPR/Cas9 system was used to generate frame-shift mutations in *PKD1-1, PKD2-1,* and *NOMPC,* the three genes that were found to be expressed in CR neurons. The mutant lines generated were used to evaluate the role of these genes in the startle response. The evidence presented in this chapter supports a role of *PKD1-1* and *PKD2-1* in the behavior. By incubating *PKD2-1* mutants lacking the startle response and wildtype larvae with a planktonic predator, support the hypothesis that the startle response is part of a behavioral mechanism for predator avoidance and defense.

Genome editing with CRISPR/Cas9 in *Platynereis*

The use of the CRISPR/Cas9 system in *Platynereis* was straightforward and with similar effects as in other organisms. Even though the efficiency assessment of the technique was carried out, no obvious differences to reports in other species was noted. One of the main concerns when using the CRISPR/Cas9 system—and that it also applies to work in *Platynereis*—is the high rate of off-target effects seen in some studies (reviewed in (X.-H. Zhang et al. 2015)). Thus, additional controls needed to be included to minimize confounding effects in the analysis of the phenotype.

To decrease off-target effects at the sgRNA level, one of the two sgRNAs targeting *PKD1-1* and the sgRNA targeting *PKD2-1* were designed to be 18bp long instead of the usual length of 20bp (**Figure 4-2**). Such truncated sgRNAs have been shown to reduce off-target mutagenesis by as much as 5000 fold (Fu et al. 2014). The *NOMPC* locus was targeted only by a 20bp-long sgRNA, but the target sequence chosen had at least three mismatches to the closest sequence at the PAM-proximal region, which have been shown to be less tolerated than those at the PAM-distal region (Fu et al. 2013). By considering these two factors in the sgRNA design, the targeting specificity may have been improved. Off-target prediction algorithms could be used in future work to further improve the specificity of the designed sgRNA (e.g.CRISPOR (Haeussler et al. 2016)). A *Platynereis* genome assembly with good coverage will be required to use this prediction tool.

Due to the fact that the experiments leading to the mutant lines reported in this study started in 2013, (not long after the initial work proving the gene editing potential of the CRISPR technology), no major considerations in the election of the Cas9 used were made. Since the publication of the engineered Cas9 used in this study great improvements in the specificity of the enzyme have been performed (see (Hu et al. 2018) for a recent example). Future work attempting to use this technique should make use of these new variants to further decrease the likelihood of off-target effects.

Ultimately, the only way to determine the degree of off-target effects is to use whole-genome sequencing or other genome-wide approaches such as exome sequencing, GUIDE-seq or BLESS (reviewed in (Zischewski et al. 2017)). Beside their elevated cost, using these techniques was not

possible as they require a reference genome with good coverage. Instead, the lines were outcrossed twice to unlink the off-target from the on-target mutations. Outcrossing is a widespread technique in genetics used to remove background mutations in mutant screens. It is important to note, however, that this protocol does not ensure the total "cleaning" of a line, as mutant variants may be retained due to balancing selection, or to the introduction of new mutant variants (Sarin et al. 2010). Although these issues are likely to be more relevant in non-targeted mutagenesis, an assessment of the level of off-target induced by the sgRNAs used would be optimal.

A second technique widely used in genetics to verify a locus-specific effect is gene-complementation analysis. Heteroallelic combinations were performed with the mutant alleles of each gene in order to support a role of *PKD1-1* and *PKD2-1* in the startle response. Only in the case of *PKD1-1*, the *i1* and Δ1 alleles were generated by different sgRNAs, while the alleles in *PKD2-1* and *NOMPC* were generated by the same sgRNA. However, since the genealogy of each mutant line carrying a different allele is unique, the combination of alleles is expected to help alleviate the concern of an off-target having an effect on the phenotype. The blind assignation of genotype to each phenotype recorded also strengthens the validity of the conclusions.

In conclusion, different strategies to minimize the unspecificity inherent to the CRISPR/Cas9 system were followed, but more rigorous analysis of the off-target effects and the use of the most specific variants and accurate design tools will have to become a norm, as genetics analysis takes off in this organism.

The role of Polycystins in the startle response

Phenotype and expression of PKD1-1 or PKD2-1

Calcium imaging experiments provided correlational evidence that CRs and pygPB^{unp} have a sensory role in the startle response (see Chapter 2). As the common expression domain of *PKD1-1* and *PKD2-1* is restricted to these two cell types, the complete abolishment of the startle response when either of these genes is mutated strongly suggests that CRs and pygPB^{unp} are the main drivers of the response. The *PKD1-1* or the *PDK2-1* promoter constructs could be used to genetically activate or inhibit CRs and pygPB^{unp} in order to directly show this functional connection. The mosaicism seen in animals injected with such constructs could be leveraged to show the differential contributions of both cell types to the response.

Function of PKD1-1 or PKD2-1 in CR/pygPB^{unp} mechanotransduction cascade

The abolishment of the startle response in *PKD1-1* and *PKD2-1* mutants (*PKD* mutants from here on) further implies that these two genes are integral to the function of CRs and pygPB^{unp} in this behavioral context. To confirm that the lack of a response in PKD mutants is due to a defect in these cells, the mutant genes could be rescued by expressing in CRs and in pygPB^{unp} their intact versions using promoter constructs. These experiments could help address the role of polycystin-

related molecules in mechanotransduction including the use of human PKD homologs to analyze their ability to rescue the response.

As mentioned before (see Introduction to Chapter 3), the role of polycystins in the mechanotransduction cascade is far from clear, and thus there are still many open possibilities regarding the exact function of *PKD1-1* and *PKD2-1* in this process. Possible models based on what is currently known and experiments that address this issue are discussed below.

The *PKD1-1* or *PKD2-1* mutant larvae neither show obvious defects in their overall morphology nor in the structure of their CR cilia. This in contrast to analysis in hair cells showing structural defects in microvilli (Steigelman et al. 2011). Similar to *Platynereis, C.elegans* mutants for PKD1 and PKD2 do not show obvious morphological defects in the sensory cilia (Barr et al. 2001). Further assessment of the neurite projections and of the dendrite fine ultrastructure in *PKD* mutants is needed to completely rule out structural defects as the main cause of the insensitivity to vibration and touch stimuli.

That said, *PKD2-1* mutants did show an elongated cilium when compared to wildtype larvae. This is in line with a study showing that knocking out either PC1 or PC2 leads to elongated primary cilia in kidney tubular cells and in (Liu et al. 2018). Ciliary length is dynamically controlled by levels of Ca^{2+} and cyclic AMP (cAMP), with reduction in the former and increase in the latter leading to longer cilia (Besschetnova et al. 2010). Besschetnova et al also showed that upon shear stress, ciliary Ca^{2+} levels increase, and cAMP levels decrease, leading to cilium shortening. As PC1 and PC2 morphants did not show this shortening in response to flow, the authors suggest these molecules play a role in modulating cilia length upon sensing mechanical cues. The elongated cilia of hCR1 in *PKD2-1* mutants suggest a similar regulatory mechanism might be at play in CR neurons.

The role of PKD2-1 in such flow-driven regulatory mechanism together with the absence of gross structural defects in *PKD* mutants suggests that PKD1-1 and PKD2-1 are tightly linked to the transduction of the mechanical signal. One model could involve the formation of a PKD1-1/PKD2-1 complex that localizes to the CR/pygPB^{unp} cilia and regulates there the gating of an unknown MeT channel complex upon cilium deflection. Assessing the localization of both proteins could be an important first step towards exploring this model, as polycystins can be localized to cilia, microvilli, ER, and to the plasma membrane (Yoder et al. 2002; Steigelman et al. 2011; Köttgen and Walz 2005; Cai et al. 1999; Foggensteiner et al. 2000). Such experiment—which could be achieved by the use of translational fusions or antibodies—would also reveal whether the two proteins colocalize to the same ciliary compartment. At the same time, it can give hints as to the role of cilia in CR mechanosensation (see Discussion in Chapter 2).

Assessing the effect of the PKD1-1/PKD2-1 complex on the MeT channel will be more challenging, as electrophysiology techniques needed to measure mechanically activated currents are

not yet routine in *Platynereis*. A doable experiment that would indicate if PKD1/PKD2 are part of the mechanotransduction cascade involves calcium imaging in the CR neurons of mutant larvae. It is important to note, however, that even if mutants lack a calcium response it is not a direct indication of PKDs directly affecting the MeT channel, as they may be still totally dispensable for mechanically-induced electric currents, as reported for the TRP channels OSM-9 and OCR-2 in *C.elegans* (Geffeney et al. 2011). Whole-voltage clamp would be required to isolate the MeT current and test if it is absent in mutants (Katta et al. 2015).

Another explanation for the lack of response in the mutants, despite the normal sensory morphology is that PKD1-1 and PKD2-1 play a general role in sensory ciliary function, independently of mechanotransduction. In support of this is the fact that *PKD2-1* was expressed in more penetrating ciliated neurons in the larva, and not only in CRs and pygPB^{unp} (see Results in Chapter 3). The exact role of polycystin-related molecules in the cilia is not known, but the recent discovery that PKD2 is an essential subunit of the channel conducting the ciliary current of primary cilia support a general role in this organelle's function (Liu et al. 2018).

Were PKD1-1 and PKD2-1 really knocked out?

Although the mutations induced with CRISPR were predicted to cause a shift in the reading frame of both *PKD1-1* and *PKD2-1*, the total absence of the gene products was not verified. In fact, alternative start codons were found downstream of the gene lesion that in principle could still lead to the translation of shortened protein versions with the transmembrane domain intact (see Appendix). Although the penetrance of the phenotype seen in *Platynereis* mutants suggests that the genes were indeed knocked out, truncated PKD2 proteins can still induce loss-of-function phenotypes (Gallagher et al. 2006). In fact, most of the mutations in human *PKD1* and *PKD2* that cause ADPKD and other diseases are missense or truncating mutations, while total loss of function is lethal in mice (Boulter et al. 2001; Lu et al. 1997; Harris and Torres 2009). Thus, in order to better interpret future experiments using the *PKD* mutants reported here, it would be advisable to first characterize the nature of the molecular deficiency. Antibodies raised against the proteins would be a direct way to confirm the absence of the proteins.

Does *NOMPC* have a function in CRs?

In stark contrast to the *PKD* mutants, the homozygous and trans-heterozygous *NOMPC* mutants generated with CRISPR did not show any quantifiable defect in the startle response in the touch assay. This, in spite of the fact that *NOMPC* is also expressed in at least a subset of CR neurons (see **Figure 3-11**H), albeit at low levels and only as indicated by WMISH, and not by promoter labeling. A definite view of the expression pattern of this gene is still needed. Assuming for now that *NOMPC* is indeed expressed in at least in hCRs, the lack of a startle phenotype has still possible interpretations.

One possibility is that *NOMPC* function in CRs may be independent of these cells' role in the startle response. There are numerous examples in multimodal sensory cells (mainly nociceptors) where mutating one channel inactivates only a specific modality, while leaving the other modalities unaffected (e.g. (Kim et al. 2012; Zhong et al. 2010; Chatzigeorgiou et al. 2010)). The *PKD* mutants failed to respond to touch and hydrodynamic stimuli, thus suggesting that CRs are responsible for detecting both cues. Although, these two types of signals may be part of a single sensory modality, it has to be noted that *NOMPC* mutants were only tested using touch stimuli. Thus, it is still possible that *NOMPC* mutants have a defective startle response when triggered by hydrodynamic stimuli. In other words, NOMPC may be required to detect only the subtler water-borne vibrations, and not the more intense touch stimuli.

A second explanation for the lack of a startle response phenotype is that *NOMPC* is required for CR function in the startle response, but the alleles generated are at worst hypomorphic versions of the gene. As for the *PKD* mutant alleles, the *NOMPC* predicted mutations generated a frame shift in the coding region, but alternative start codons were found downstream. In *Drosophila*, null mutations of this single copy gene have markedly reduced viability and obvious locomotion defects (Kernan et al. 1994; Walker et al. 2000). Such drastic defect is in part due to *NOMPC* expression in multiple sensory neurons (Cheng et al. 2010). *NOMPC* is also present in a single copy in *Platynereis* genome and it is also expressed in multiple sensory cells (see Chapter 3). Despite this fact, the homozygote worms do not show any obvious morphological or behavioral defects at the larval or adult stages (unpublished observations). Confirmation of the absence of the NOMPC protein in the mutants is thus needed before completely ruling out a role of this molecule in the startle response. Future genetic manipulations on this gene should rather target the transmembrane domain, as even a missense mutation in this region results in behavioral defects in other animals (Walker et al. 2000; Cheng et al. 2010).

Future work in the genetics of the startle response

The ease of the phenotyping touch assay used study could be extended in the future for analyzing the role of other molecules in the startle response. A particular emphasis could be put in identifying the MeT channel complex for comparisons to other organisms. First in the list of candidate genes to test will be the homologs to known MeT channels. Although *NOMPC* mutants did not show a phenotype, other mutant variants or other channels like the single homolog TMC could be involved in the response. Non-biased methods such as forward genetic screens are difficult to implement in *Platynereis* due to the long generation time. An alternative unbiased approach to find gene candidates is single-cell sequencing. Unfortunately, this method is not sensitive enough to recover the lowly expressed MeT channel candidates. Proteomics may be likewise challenging as it requires enough sample to perform proteome analysis. Despite its challenges, the identification of the MeT channel

in CR cells is a tempting goal that could help us understand the basis of mechanotransduction across animals.

The startle response in an ecological context: A predator defensive behavior?

Startle responses have evolved in many cases as part of a strategy to avoid and deter predators. The survival value of a particular startle response is however often taken for granted, even when as for any other biological process, it is subject to experimental inquiry (Tinbergen 1963). The startle response was described in Chapter 1 and some of its parameters measured. The CR neurons are the best candidates to drive the response based on calcium imaging and genetic analysis presented in this chapter and in Chapter 2. But what is this response actually useful for? Similar responses have been observed in other planktonic animals and implied to deter predators (Pennington and Chia 1984; Thiel et al. 2017). Initial prey-predator experiments were performed here using *Platynereis* mutant larvae and a copepod predator to show that the response is indeed an effective way to reduce mortality by predation. These results more firmly place the response in an ecological context that could be discussed in this light.

The survival value of the startle response

Based on its behavioral features, the survival value of the startle response could lie in its use as an avoidance response (using ciliary closures), or as a defense response (parapodial elevation), or it may rely on a combination of both. The predation experiments do not allow to distinguish which of the features has a bigger role, as the *PKD2-1* mutants used in the experiments do not display either of them. Both avoidance and defense strategies have been recognized as important contributors for determining the dynamics of plankton populations (Ohman 1988). It is for now worth discussing the mechanisms by which the larva may use the response to avoid being eaten.

The startle response as predator avoidance mechanism

The startle response may be an avoidance response because it may allow the larva to stop swimming by shutting down the cilia upon detecting an approaching predator. The use of hydrodynamic signals for successful predator detection and escape has been documented in other animals, even those with well-developed vision (Stewart et al. 2014; Viitasalo et al. 1998; Browman et al. 2011). Moreover, the ambush feeding technique of the copepod used in this study makes it more hydrodynamically conspicuous to the larva (Titelman and Kiørboe 2003), which may help deploy a timely startle response.

Although ciliary arrests may seem insufficient to avoid a predator, at this scale the exact motility patterns and the size of the prey are important factors for their detectability (Titelman and Kiørboe 2003; Kiørboe et al. 2014; Gallager 1993). The lack of a feeding current together with its small size makes swimming *Platynereis* already inconspicuous when compared to other planktonic animals (Kiørboe et al. 2014). But the presence of a sinking response in *Platynereis* may further help avoid

detection by reducing its hydrodynamic signal, as suggested for the dead man response seen in some cladocerans (Kerfoot 1978). Thus, the sinking strategy may be an effective defense for animals that do not have fast jumping escape mechanisms. Moreover, a sinking strategy may be a particularly effective avoidance strategy against predation by ambush-feeding copepods like the one used here , as they are relatively inefficient in detecting their prey by chemical signals, and rely more on hydromechanical cues for prey detection (Svensen and Kiørboe 2000). That said, a fine line between detecting the predator and being detected probably exists, as copepods have well developed hydrodynamic receptors in their antennae (Strickler and Bal 1973), and can accurately locate a prey by its hydrodynamic profile alone (Strickler and Balázsi 2007).

The startle response as defense mechanism against predators

Due to the fast speed of copepod attacks (100mm s^{-1})(Kiørboe et al. 2009) and to the highly developed mechanosensory systems of these animals, *Platynereis* larvae may be spotted before it can avoid the predator, and get attacked (as seen in **Figure 4-7**C). Such attack likely triggers the fast elevation of the parapodia with the concomitant extension of all the chaetae that is the characteristic feature of the startle response. The mutants lacking this reaction may thus be unarmed against copepod attacks and be more easily eaten. The chaetae would serve as little spines that trigger a nocifensive reaction of the predator (as seen in **Figure 4-7**C). Previous experiments have supported a defensive role for chaetae (Pennington and Chia 1984), but it was not clear if the sole presence of these structures was enough to deter predators. As both mutant and wildtype larvae had chaetae, their sole presence is not enough to explain the different predation rates. The results presented here suggest that the extension of chaetae is a major determinant of survival. Although it is perhaps hard to conceive that such small spines would deter a big copepod predator, it has been shown that if copepods have a choice, they will rather choose those prey that are easier to catch and to eat (Kerfoot 1978; Williamson 1987; Gilbert and Williamson 1978; Broglio et al. 2001). Thus, even seemingly sub-optimal defensive behaviors can have noticeable effects in the composition of planktonic populations.

Alternative interpretations of the predation experiments

The predation experiments point to the startle response as the main explanatory variable behind the different survival rates between wildtype and *PKD2-1* mutants. Nonetheless, as *PKD2-1* is not only expressed in CRs and pygPB^{unp}, but also in other sensory cells (see **Figure 3-3**), it is still possible that the mutant larvae may have additional behavioral defects unrelated to the startle response.

For instance, the *PKD2-1* mutants may have a reduced ability in detecting a chemical signal emanating from the predators. Although in principle possible, the diffusion of any chemical emanated from the copepod would be slower than a mechanical signal triggering the response. The

diffusion of chemicals by diffusion in non-turbulent water is slow (Berg 1993; Martens et al. 2015). On the other hand, the aversive response could be due to a post-ingestion effect. Although *PKD2-1* was only expressed in sensory cells, they could in principle trigger the expression of a defensive chemical upon the larva is being ingested. Prey-predator experiments in other planktonic organisms have shown that some predators regurgitate their prey (Gilbert and Williamson 1978; Williamson 1980). Preliminary experiments have in fact shown that stickleback larval fish regurgitate *Platynereis* larvae (unpublished observations). Although the startle response may also induce such response (due to the extension of the chaetae), the effect of chemical defenses cannot be completely ruled out. To clarify the contribution of the startle response to reduced predation, the experiments could be repeated using the *PKD1-1* mutants instead. As this gene is expressed mostly in CRs the number of alternative explanations would be reduced.

Outlook research in the ethology of the startle response

The present experiments set the stage for further investigating the behavioral mechanisms behind this prey-predator interaction. A primary goal would be to actually image the prey-predator encounters to confirm and to clarify the importance of the startle response and of each of its features. It will also help define the type of hydrodynamic interactions that trigger it in a more natural context, which in turn may be useful to use more naturalistic stimuli in mechanistic studies of the behavior. Freely behaving prey-predator experiments in copepods have been performed using a 3D recording system arguing for the feasibility of this idea (Fields and Yen 1992; Bradley et al. 2013). Recording natural interactions could also uncover additional behavioral mechanisms that the larva might use to escape predation.

The direct competition experiments could be repeated using predators with other attack mechanisms, to determine the type of predators to which the startle response is most effective. The startle response may not be equally effective escape response against all predators. The ciliary arrest and sinking response would probably not be effective against visually guided predators. Predators that do not generate a detectable hydrodynamic signal would not trigger the response until contact was made. In fact, preliminary experiments using *Clytia hemisphaerica* medusae (which may be immobile for long periods) did not show an increase in predation rate in wildtype vs mutant larvae, but only in larval stages with chaetae and without chaetae (unpublished observations). Differences in predation success depending on predator and on larval stage have been documented in other prey-predator experiments (Gilbert and Williamson 1978; Pennington and Chia 1984). The extension of chaetae may have evolved as defense specially tuned against larger rheotactic predators, rather than to filter-feeders, which are better avoided by jumping mechanisms (Williamson 1987).

Both direct observation of the prey-predator interactions and direct competition experiments could be further potentiated if combined with additional mutants with more subtle defects in the startle response. In this way the survival value of coordination, of synchrony in ciliary beating, and other features of the startle response could be explored.

Chapter 5 The CR neuronal circuit of the startle response

"Connectomes are absolutely necessary but completely insufficient for understanding nervous system function"

-Eve Marder

Statement of contributions and publication status

Reconstruction and reviewing of neurons in the EM volume was done by the author of this thesis together with Martin Gühmann, Sanja Jasek, Gáspár Jékely, Nadine Randel, Reza Shahidi and Csaba Verasztó. sTEM micrographs were acquired by Reza Shahidi. Snapshots and wiring diagrams were obtained and assembled into figures by the author of this study.

The EM volume, and the reconstruction of pygPB^{unp}, Arc interneurons, ciliomotor neurons, and ciliary bands was previously published (Verasztó et al. 2017; Randel et al. 2015; Randel et al. 2014; Shahidi et al. 2015). The CR circuit was published elsewhere (Bezares-Calderón et al. 2018)

Introduction

Startle neuronal circuits

General importance

Startle responses have been chosen as a behavior that can be used to understand neuronal and circuit mechanisms. This stems from the starting assumption that the fast, and often stereotyped nature of startle responses is implemented by a relatively simple circuit that can be entirely understood. Although this is not often the case, the study of startle circuits has led to a greater understanding about general neuronal and circuit mechanisms. For instance, the tail flip startle circuit in crayfish has led to understand circuit mechanisms such as integration, coincidence detection, decision-making, among others (Edwards et al. 1999). It also provided us with the concept of command neuron (Wiersma and Ikeda 1964). In the sea slug *Aplysia,* a basic withdrawal reflex was used as a model to understand the molecular principles of associative and non-associative learning (Hawkins and Byrne 2015; Byrne and Hawkins 2015). Recordings of negative field potentials in the Mauthner cell axon hillock led to the discovery of non-synaptic electrical inhibition (Furshpan and Furukawa 1962). Startle circuits have been investigated in a multitude of systems, a rare case given the emphasis in modern neuroscience to focus on only a handful of systems.

Common principles emerging from startle circuits

This research has led to the finding of common principles in startle circuits. Speed and coordination are some of the hallmarks of startle responses in many animals. These feats are in many cases achieved by a system of neurons with axons of relatively large caliber that relay the afferent signal without much delay and intermediary synapses to the locomotor system. In crayfish, a set of medial and lateral giant fibers, MG and LG, run along the ventral nerve chord and innervate motoneurons in the abdominal segments (reviewed in (Edwards et al. 1999)). When stimulated, these neurons trigger a fast simultaneous contraction of left and right flexor muscles along the abdomen, causing bending of that region. The sufficiency of these neurons to trigger a whole body response led to the concept of command neurons (Wiersma 1947; Wiersma 1938). The conduction velocity of these cells is extremely fast, and rectifying electrical synapses contribute an efficient mechanism of coincident detection (Edwards et al. 1998). Fish and amphibians have a pair of command-like giant neurons called Mauthner or M-cells that are located in the brainstem region. These cells cross the midline and project down the nerve chord to innervate primary motoneurons in each segment of the trunk (Fetcho 1991). In fish, this contralateral innervation causes the unilateral C-type bending of the body by a single M-cell AP (reviewed in (Korn and Faber 2005)) . Many annelids possess a system of giant fibers in the ventral nerve chord that are involved in fast startle responses (reviewed

in (Drewes 1984)). These neurons have command-like properties in nereid polychaetes and earthworms, as their ablation results in loss of the response and their stimulation in the activation of the startle behavior (Nicol 1948; Kupfermann and Weiss 1978; Theodore Holmes Bullock 1945). At least in nereids, there is anatomical evidence these neurons target motoneurons segmentally arranged in the trunk (Smith 1957; Horridge 1959). In *Drosophila*, a pair of giant neurons (GF neurons) drives simultaneous wing depression and leg contraction (Thomas and Wyman 1984). A single GF spike induces muscle contraction by directly synapsing on giant ipsilateral motoneurons in the thorax (Tanouye and Wyman 1980). GFs are indispensable for and induce fast jump escapes (von Reyn et al. 2014).

Although giant fiber systems definitely play a central role in startle responses, there are often converging pathways that are responsible for endowing flexibility to and for fine-tuning of the startle responses. In zebrafish, additional reticulospinal neurons serially homologous to the M-cell are required for short latency responses, forming a M-cell array (Liu and Fetcho 1999; O'Malley et al. 1996). A whole reticulospinal neuronal population is also activated during escape responses and may explain the presence of a delayed, but still normal C-type escape response in fish with M-cells ablated (Liu and Fetcho 1999; Gahtan et al. 2002; Eaton et al. 1982). It may also explain the weaker and less flexible response induced by artificial activation of M-cells alone (Nissanov et al. 1990). In *Drosophila*, a slower but more finely controlled escape response does not fully require GFs (von Reyn et al. 2014). In crayfish, a less stereotyped and slower tail flip escape response is partially mediated by segmentally iterated interneurons that through local circuits trigger small tail flip contractions (Kramer and Krasne 1984).

The nature of the stimulus often determines the exact neuronal pathway activated in startle circuits. In fish, the rate of approach of a looming visual stimulus determines which neuronal pathway is activated. High approach rates activate the M-cells and other reticulospinal neurons, while low approach rates activate only the reticulospinal population, but not the M cells (Bhattacharyya et al. 2017). GFs in *Drosophila* respond to sound stimuli through direct electrical coupling with the auditory organ cells (Lehnert et al. 2013), as well as to fast looming stimuli conveyed by direct synapses from the optic lobe (Bacon and Strausfeld 1986; Mu et al. 2014; von Reyn et al. 2014). Slower looming stimuli activate additional (as yet unidentified) circuits involved in a more controlled takeoff (von Reyn et al. 2014). Crayfish giant neurons receive direct sensory input from hair cells and stretch receptors (Newland et al. 1997). The giant pathway is triggered by abrupt mechanical stimuli (Olson and Krasne 1981), or in juveniles, by moving shadows (Liden and Herberholz 2008). The non-giant tail flips are generated during more gradually developing mechanical stimuli (Wine and Krasne 1972).

The site of stimulation has different effects on the activation of the downstream circuits. In fish, tail stimulation elicits activation of M-cells alone, while anterior stimulation also activates the complete M-cell array in the hindbrain (O'Malley et al. 1996). Moreover, the specific site of anterior stimulation can trigger different pathways. While auditory and visual inputs activate M-cells, head-tactile input induces fast escapes through one of the M-cell segmental homologs in the hindbrain (Kohashi and Oda 2008). In *C.elegans* touch sensory neurons in the head and in the tail synapse on a partially overlapping set of interneurons that dictate the different responses to anterior and posterior stimulation (Chalfie et al. 1985). Inhibitory interactions between the different populations ensure that only one of the responses is triggered. In crayfish, abrupt stimulation to the tail or to the posterior abdomen activates the LG fibers, while anterior stimuli (above the abdomen) preferentially activates the MG fiber. These two giant fiber systems target partially overlapping sets of local interneurons and motoneurons, which explains the resulting patterns of abdominal flexion (Heitler and Fraser 1993).

Different circuit motifs ensure startle responses are robust and produce a vigorous startle response. In fish, ipsilateral descending (ID) excitatory interneurons are activated by the M-cell on the opposite side and directly activate primary motoneurons (Fetcho and Faber 1988; Kimura et al. 2006). This interaction amplifies and broadens the muscle excitation effect of M cells. Similarly in crayfish, a giant segmental interneuron is activated by LG, and together activate fast flexor muscle motoneurons, thus evoking a powerful abdominal flexion (Kramer et al. 1981).

Other circuit motifs ensure response are synchronous and well-coordinated across segments and body sides. In some cases, this is achieved by the intersegmental giant fiber systems mentioned above, but other circuit elements play a role. The unilateral bending of the body during the fish C-start response requires commissural local (CoLo) inhibitory interneurons directly inhibiting ID interneurons and primary motoneurons on the side opposite to the intended bending (Satou et al. 2009; Liao and Fetcho 2008). In bilaterally symmetric responses such as the escape response in crayfish or the jump escape in *Drosophila*, coordinated muscle contraction is thought to be a consequence of electrical coupling between the giant fibers on each side of the body (Hale et al. 2016). In the *Drosophila* larva, interneuron pathways with partially overlapping inputs and outputs control different behavioral subsequences that are integrated to produce a seamless nocifensive behavior (Burgos et al. 2018).

Startle circuits in planktonic organisms

Characterization of planktonic startle circuits has not been performed extensively. Perhaps the only planktonic animal where physiological and anatomical studies exist is the hydrozoan medusa *Aglantha*. In this animal a ring giant axon around the vellum is activated upon stimuli that activate a swimming escape response (Roberts and Mackie 1980). This giant fiber receives sensory input

from hair cells and activates giant and non-giant motoneurons that synchronously contract the muscles in the vellum (Mackie and Meech 1995; Arkett et al. 1988). This circuit architecture is reminiscent to that seen in giant fiber circuits in bilaterian animals (Roberts and Mackie 1980). Giant fibers in other planktonic animals have been reported, but their association to escape responses has not been definitely made. A decussating giant fiber system exists in phoronids (Temereva and Tsitrin 2014), and it has been physiologically characterized (Wilson and Bullock 1959). The activation of this system has been shown to be necessary for a withdrawal reflex[21]. More recently, a circuit morphologically and topologically similar to the Mauthner-like startle circuit was reported in the planktonic larva of the tunicate *Ciona* (Ryan et al. 2017). This circuit awaits physiological characterization.

Circuit reconstruction using electron microscopy

Although startle circuits are arguably the best characterized circuits in the animal kingdom, new circuit reconstruction approaches have allowed to characterize previously undescribed startle circuits and to identify new elements and motifs in previously studied systems. Perhaps the most powerful approach is based on imaging a piece of tissue or a complete animal using electron microscopy (EM) methods and mapping neural circuits on it by identifying synaptic interactions between reconstructed cells. This is an unbiased method for recovering the topology of neural circuits. It was first used to identify the circuit driving the touch withdrawal response in *C.elegans* (Chalfie et al. 1985). In *Drosophila*, a nocifensive behavior has been analyzed using EM volume reconstruction to identify key neurons and motifs in the circuit (Ohyama et al. 2015; Jovanic et al. 2016; Burgos et al. 2018; Takagi et al. 2017). At least 13 morphologically distinct types of interneurons were found downstream of the nociceptors triggering the response (Ohyama et al. 2015). The near-complete animal neuronal circuitry (the connectome) recently completed in *Ciona* (Ryan et al. 2016), allowed to reconstruct the putative startle circuit mentioned above from the sensory input to the musculature. The transferability of the method allows to reconstruct circuits in non-conventional systems that can be then compared to other startle circuits in systems amenable to experimental interrogation (Hale 2014).

In this chapter, EM volume reconstruction is used in the *Platynereis* nectochaete to characterize the synaptic circuit downstream of the CR neurons. Given the anatomical, physiological and molecular evidence presented in the preceding chapters that point to the role of these cells as the drivers of the startle response, the circuit reconstructed is considered to be the core of the startle circuit. The reconstruction was made on a previously published dataset acquired using serial transmission EM (Shahidi et al. 2015; Randel et al. 2015).

[21]Attributed to (Silen 1954) as referred by (Wilson and Bullock 1959).

Material and Methods

CR circuit reconstruction

Neuron reconstruction and synapse annotation

The chemical synapse circuit downstream of the CRs was reconstructed using CATMAID (Saalfeld et al. 2009). All neurons in the circuit were manually reviewed using the Review widget. Chemical synapses were identified as clusters of vesicles located adjacent to the inner side of the presynaptic membrane (**Figure 5-1**). Every cluster of vesicles was counted as one, without considering the number of sections it spanned.

Defining the CR circuit

The set of first-order interneurons was defined as neurons targeted by at least 2 CRs making together 3 or more synaptic contacts. Additional neurons with only 1 or 2 synapses, or only one upstream CR neuron were included if their bilateral counterparts (as defined by neuron morphology) were already part of the first selected set. As the aim of the reconstruction was to define the most direct pathway from CRs to the locomotor output (both cilia or muscles), only the synaptic partners of the 1st interneuron layer making synapses on cilia and muscles were included in the final circuit (i.e. only muscle motoneurons were considered).

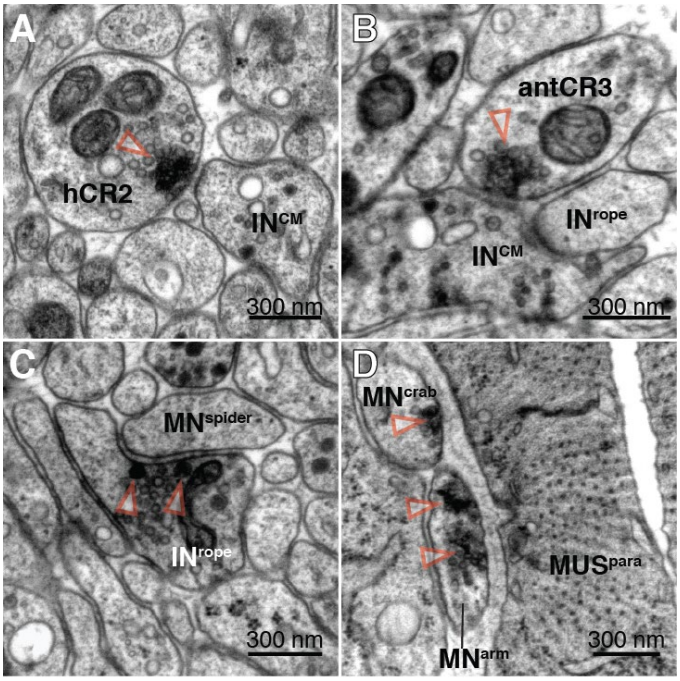


Figure 5-1 Synapses in the electron microscopy volume. (**A**-**D**) High-resolution images taken from different layers showing regions where synapses in the CR circuit were annotated. Synapses were defined as an accumulation of dense-core or clear vesicles clustered on the inner membrane of an axon (arrowheads). (**A**) The head CR neuron hCR2 makes synaptic contact with IN^{CM}, one of the identified interneurons. (**B**) An antenna CR synapses on the same interneuron. (**C**) Another interneuron type in the CR circuit (IN^{Rope}) makes two synapses onto a type of VNC motoneuron (MN^{spider}). (**D**) Two other types of motoneurons (MN^{crab} and MN^{arm}) in the CR circuit make synaptic contact with a muscle cell in the parapodium (MUS^{para}). The resolution of the dataset used for synapse annotation is lower than the one shown here.

Weakly connected cells were double-checked for accuracy of the synapse annotation. Eight fragments with three or more synapses downstream of the CR neurons but without a cell body were not included in the final circuit. A pair of glial cells, 4 sensory cells and two interneurons without a bilateral pair were also not included. The set of sensory neurons, interneurons and motoneurons newly reconstructed for this study were uploaded to the NeuroMorpho database (Ascoli et al. 2007)[22]. The models can be visualized online at the database website[23].

Muscles were named according to the nomenclature proposed by a study run in parallel to this, and that will describe the complete musculature of the nectochaete larva (S. Jasek and G. Jékely, unpublished results).

Visualization of circuits and neuron reconstruction

Circuits shown in this chapter were obtained in CATMAID with the Graph widget. The circuit was filtered for minimum number of synapses and for confidence in synapse annotation (only the most confident annotations were included). Networks were exported as .svg files and edited in Adobe Illustrator CC (Adobe Inc.) for improving the graphic style. Snapshots of neuron morphologies were acquired in CATMAID with the 3D widget. A smoothing function with sigma of 3000-6000 was used. Segmental boundaries and scale bars were applied as described in **Figure 2-1**. The connectivity matrix from musclemotor neurons to muscles was acquired in CATMAID with the Matrix widget and exported as .csv file. The image was imported in FiJi, transformed to 8-bit and the LUT 'glow' applied to the inverted image.

Calcium imaging of muscles activated during the startle response

Experimental setup

These experiments were performed using the same protocol, materials and equipment described for calcium imaging of sensory neurons in Chapter 3. The only distinction in this case is that no mecamylamine (i.e. no anesthetic) was used. Muscles were identified by comparing their shape to that known from previously published muscle stainings (e.g. (Fischer et al. 2010)) and by the reconstructions in the electron microscopy volume. Recordings were made at 4Hz.

[22]The morphology models can be downloaded from this link: doi: 10.13021/degz-cz50
[23]http://neuromorpho.org/index.jsp

Results

Two lines of evidence presented in this study suggest CR neurons play a role in triggering the startle response (described in Chapter 1). First, calcium imaging showed that hydrodynamic disturbances activate hCR neurons (Chapter 2). And second, frameshift mutations in two genes expressed in CRs, *PKD1-1* and *PKD2-1,* abolished the startle response (Chapter 3 and Chapter 4). Thus, reconstruction of the downstream synaptic targets of CRs should in theory consist of those neurons participating in triggering the startle response. This analysis assumes CRs in other parts of the body have also a hydrodynamic sensory modality, which seems reasonable based on the similar dendrite morphology (**Table 0-6**). The CR circuitry was reconstructed in the EM volume previously described (Randel et al. 2015; Shahidi et al. 2015).

CR neurons project ipsilaterally down and up the VNC

As shown in Chapter 2, CRs are located in the episphere, DSO, parapodia and pygidium. CRs were grouped based on their location. The CR^{head} group includes all CRs in the episphere (hCRs, antennal CRs and dsoCRs), and the single unpaired CRs next to the prototroch, $doCR^{unp}$. $CR^{segmental}$ consists of CRs located in the head cirri (located in segment 0) and in-between notopodia and neuropodia in each segment. Finally, $CR^{pygidium}$ groups the receptors in the pygidial cirri (only those on the right body side could be reconstructed), and in the pygidium proper (**Figure 5-2**A). CRs, like all penetrating ciliated neurons described in this study have ipsilateral projections. CR^{head} neurons do not project to the supraesophageal ganglion (brain) of the larva, but rather project through the dorsal root of the circumesophageal connective down to the VNC, terminating at around the 2nd segment (**Figure 5-2**B). CR^{head} neurons make synaptic contacts all along their length, although fewer synapses are found in the head. $CR^{segmental}$ are for the most part still developing, and thus do not reach the VNC (**Figure 5-2**C). The only exception to this is the interparapCR neuronal pair in the first segment. This neuron does reach the VNC through the parapodial nerve (nerve 2 in annelid neuromorphology) and then turn ipsilaterally down to make synaptic contacts (**Figure 5-2**C). $CR^{pygidium}$ neurons also project ipsilaterally to the VNC up to 2nd segment and make extensive synaptic contacts there (**Figure 5-2**D). In total 45 CRs make 840 synapses[24].

The CR-mediated startle circuit

As shown in Chapter 1, the initiation of the startle response has a short latency. Thus, the underlying circuit activated by the CRs likely follows the most direct pathway to the ciliary bands and muscles activated in the response. To identify meaningful CR targets, an initial set of neurons was formed with neurons having 3 or more synapses from at least 2 CRs. Additional neurons with

[24] Database queried on January 16th, 2019.

at least 1 synapse from any given CR were included if the bilateral pair was already in the initial set. The downstream targets of the interneurons in the resulting set were reconstructed but only included in the startle circuit if they belonged to the motoneuron class, either controlling cilia or muscles. In this manner, a sensory motor circuit was defined that could explain how the startle response is initiated upon CR activation (**Figure 5-3**).

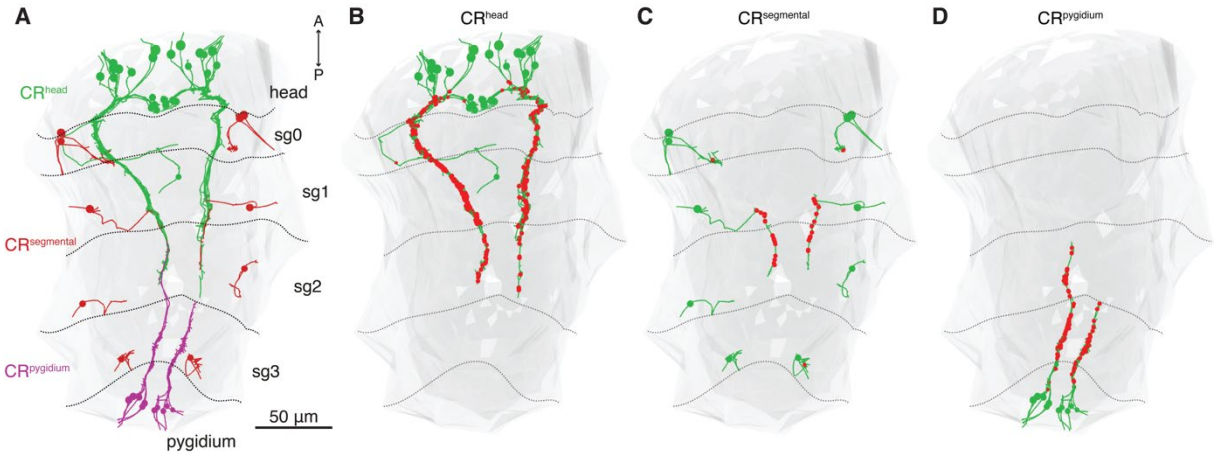


Figure 5-2 CR neurons project ipsilaterally to the VNC and make extensive chemical synaptic contacts. (**A**) Overview of CRs grouped by location. Three groups were formed: CR^{head} (green), CR^{segmental} (brown), and CR^{pygidium} (magenta). (**B**-**D**) Synapses (red dots) mapped to the different CR groups: CR^{head} (B), CR^{segmental} (C), and CR^{pygidium} (D). Segmental boundaries are marked with dotted lines and defined as in **Figure 2-1**. sg0-3: segment 0-3. Ventral view in all panels.

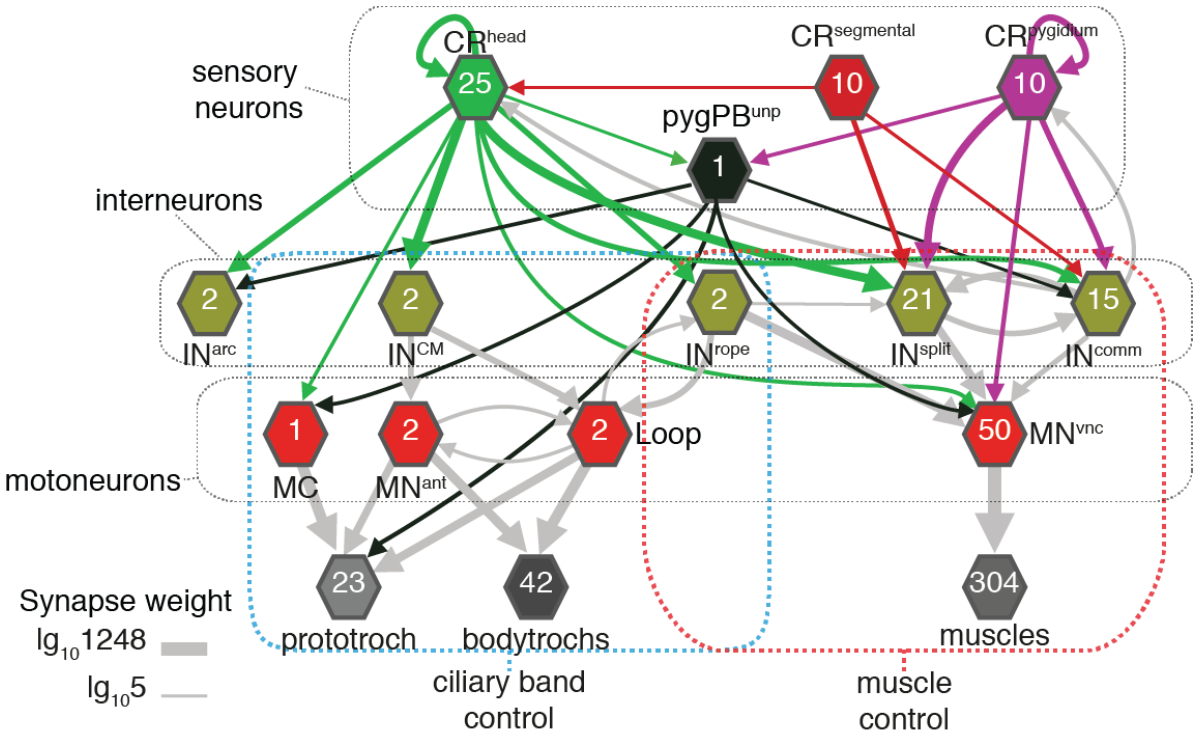


Figure 5-3 The CR circuit potentially initiating the startle response. Wiring diagram showing the three groups of CRs and the downstream circuit. In all figures in this chapter the following convention is applied: group of cells are represented with hexagons, with the number inside indicating the number of cells in each group. Arrows symbolize the synaptic contact a group of cells makes onto another group. Arrow thickness is proportional to the logarithm of the number of synapses. For clarity, in this circuit only interactions with 5 or more synapses are displayed. The circuit is divided into a module for ciliary band control and another one for muscle control (see main text for details).

CR neurons direct synaptic partners

CRs directly target other sensory neurons (including pygPB^{unp}), interneurons and motoneurons (**Figure 5-3**). CRs within each group target each other more extensively than between each group. The different CR groups have both common and exclusive synaptic targets, but CR^{head} neurons have the most diverse set of targets.

CRs in the episphere target a novel and unique set of interneurons

The CM interneurons

CR^{head} neurons, and no other CR group, make extensive synaptic contacts with a novel type of bilateral interneuron called CM (IN^{CM}) (**Figure 5-4**A-B). This is the interneuron type that receives most synapses from CR^{head} neurons on both body sides (**Figure 5-4**C). IN^{CM} are pseudounipolar neurons with a soma located in the first trunk segment that sends neurite projections towards the episphere and end before reaching the brain plexus. The downward branch is much shorter. CRs synapse on IN^{CM} along its whole length (**Figure 5-4**D).

The Rope interneurons

A second novel interneuron type targeted only by CR^{head} but not by the other CRs is IN^{rope} (**Figure 5-4**A-B). The soma of this neuronal type is located in the episphere, but below the brain plexus. It sends an ipsilateral neurite projection down across the whole VNC spanning the three segments. Even though the synaptic strength of CR^{head} neurons to IN^{rope} is much weaker than that to IN^{CM} neurons (**Figure 5-4**C), a similar number of CRs target both interneuron types (**Figure 5-4**D-E). Moreover, the weakness in the CR-> IN^{rope} connection is bilaterally symmetric, and thus it is not likely due to a sampling problem. All synapses from CRs to IN^{rope} are made at the proximal end of the neurite (**Figure 5-4**E).

The Arc interneurons

The last interneuron type targeted exclusively by CR^{head} neurons is IN^{arc}, a cell type forming part of the visual eye circuit (Randel et al. 2014; Randel et al. 2015). This neuron is also located in the episphere and its neurite crosses the midline before projecting downward to the VNC (**Figure 5-4**A-B). CR^{head} on both sides of the body target IN^{arc} with a similar number of synapses as seen for IN^{rope} neurons (**Figure 5-4**C). The synapses are located also in the proximal part of the IN^{arc} neurite (**Figure 5-4**F). pygPB^{unp} also targets both left and right IN^{arc} neurons (**Figure 5-4**C).

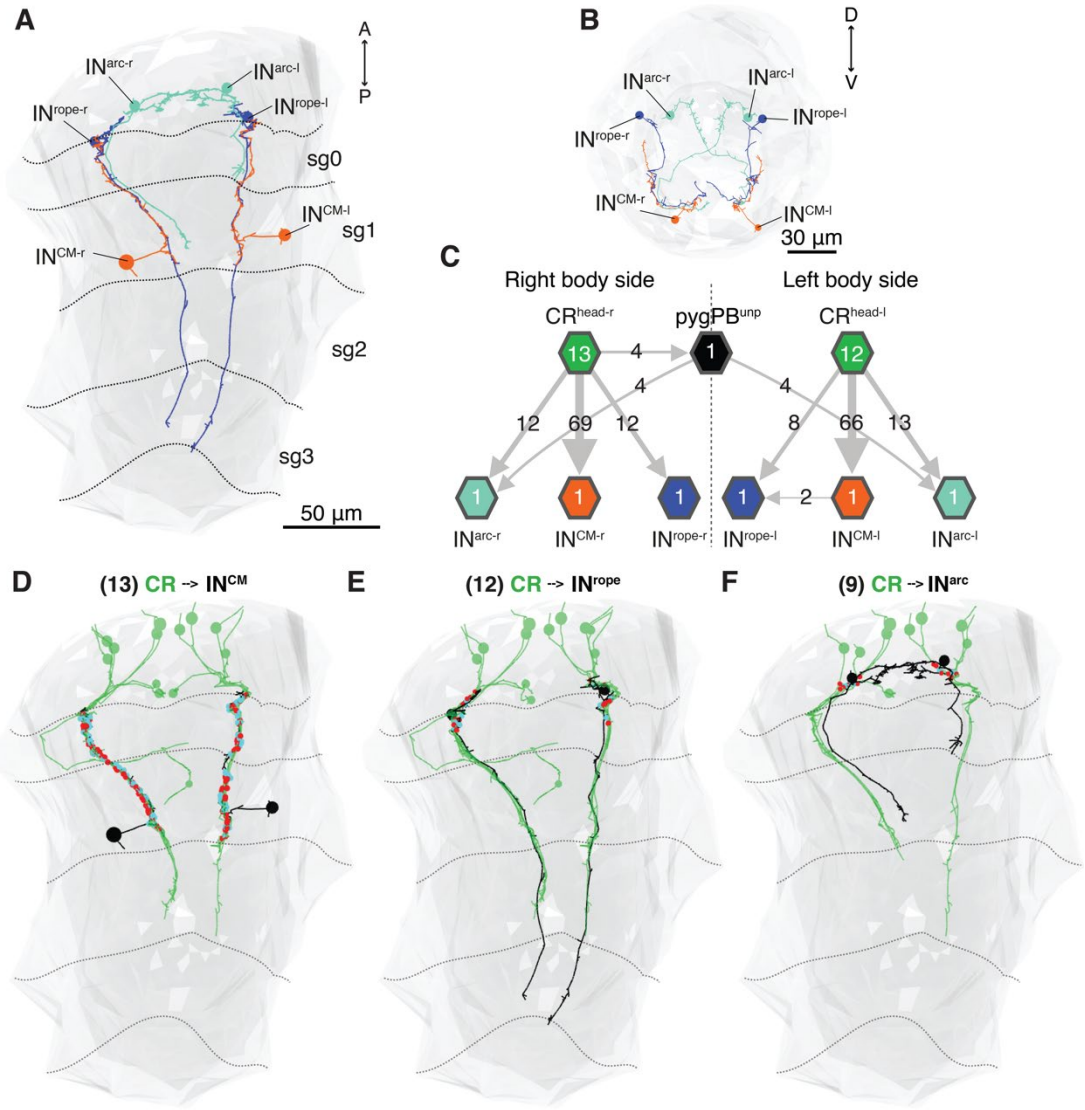


Figure 5-4 Arc, Rope, and CM interneurons are targeted by head CRs. (**A**-**B**) Overview of IN^{CM}, IN^{arc}, and IN^{rope} morphology. Ventral (A) or apical (B) view. (**C**) Wiring diagram showing synaptic interactions between CR^{head} neurons and the interneurons IN^{CM}, IN^{arc}, and IN^{rope}. Only interactions with 2 or more synapses are shown. Numbers on arrows indicate the number of synapses in each case. Remaining symbols as defined in **Figure 5-3**. (**D**-**F**) Mapping of synaptic contacts (red/cyan dots) of CRs onto IN^{CM}(D), IN^{rope}(E), and IN^{arc}(F). Only CRs making synapses are shown (number of cells indicated on top of each panel). Segmental boundaries in A, D, E, and F are marked with dotted lines and defined as in **Figure 2-1**. sg0-3: segment 0-3. Ventral view.

Split and commissural interneurons are targeted by all CR groups

The Split interneurons

In contrast to the three IN cell types found targeted only by episphere CRs, at least two types of trunk-specific interneurons were found to be also targeted by CRs in the parapodia and in the pygidium. The first group is formed by 21 neurons named Split interneurons (IN^{split}) as each has a neurite splitting into an upward and a downward projection (**Figure 5-5**A). These neurons like IN^{CM} are thus pseudounipolar ipsilateral interneurons. The length of each neurite 'arm' differs between these neurons, some of them extending past segmental boundaries. The difference in length may be in part due to neuronal reconstruction factors, but it certainly does not entirely explain the diversity in length. The degree of small branches also varies between neurons.

Both left and right CR neurons from the three anatomical groups target split interneurons in the first and second segments, while CR^{pygidium} neurons also target those in the third segment (**Figure 5-5**B-E). The CR->IN^{split} network is left-right symmetrical, even though not all cells could be matched to a bilateral pair. Synaptic contacts of CRs onto IN^{split} do not show any evident spatial clustering (**Figure 5-5**C-E). Synaptic interactions between IN^{split} and pygPBunp are virtually absent (**Figure 5-5**B).

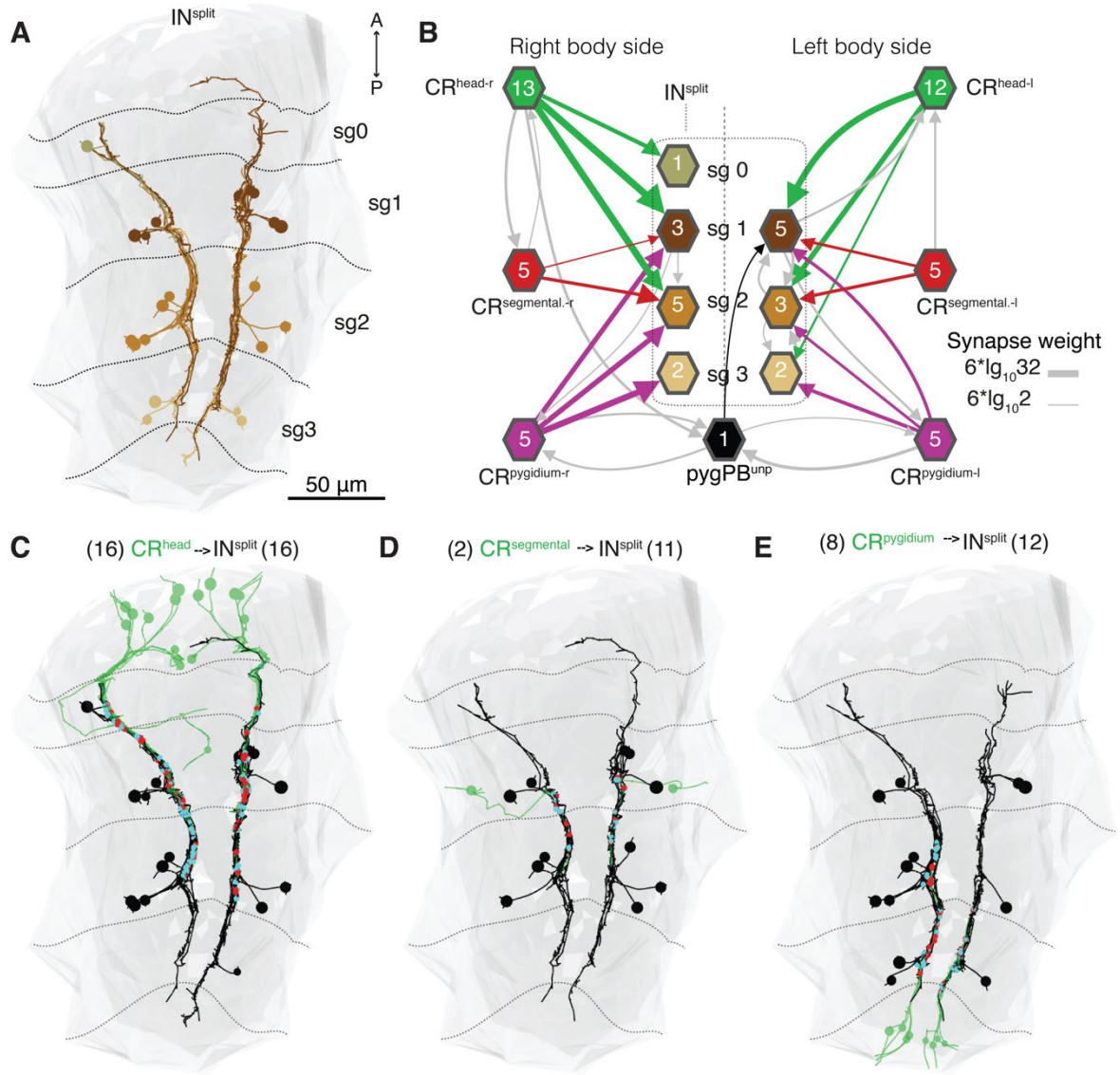


Figure 5-5 CRs in episphere trunk and pygidium target Split interneurons, a group of segmentally iterated pseudounipolar ipsilateral interneurons. (**A**) The 21 IN^{split} neurons targeted by CRs. Note the segmental clustering of these neurons and their pseudounipolar morphology. (**B**) Wiring diagram of synaptic interactions between CRs and IN^{split}. Interactions with less than 2 synapses were filtered out. Network symbol convention as defined in **Figure 5-3**. (**C**-**E**) Mapping of synaptic contacts (red/cyan dots) of CRs onto IN^{split}. Only neurons in each group making synaptic contact are shown. Number of cells in each group is indicated above each panel. Segmental boundaries in A, C, D, and E are marked with dotted lines and defined as in **Figure 2-1**. sg0-3: segment 0-3. Ventral view.

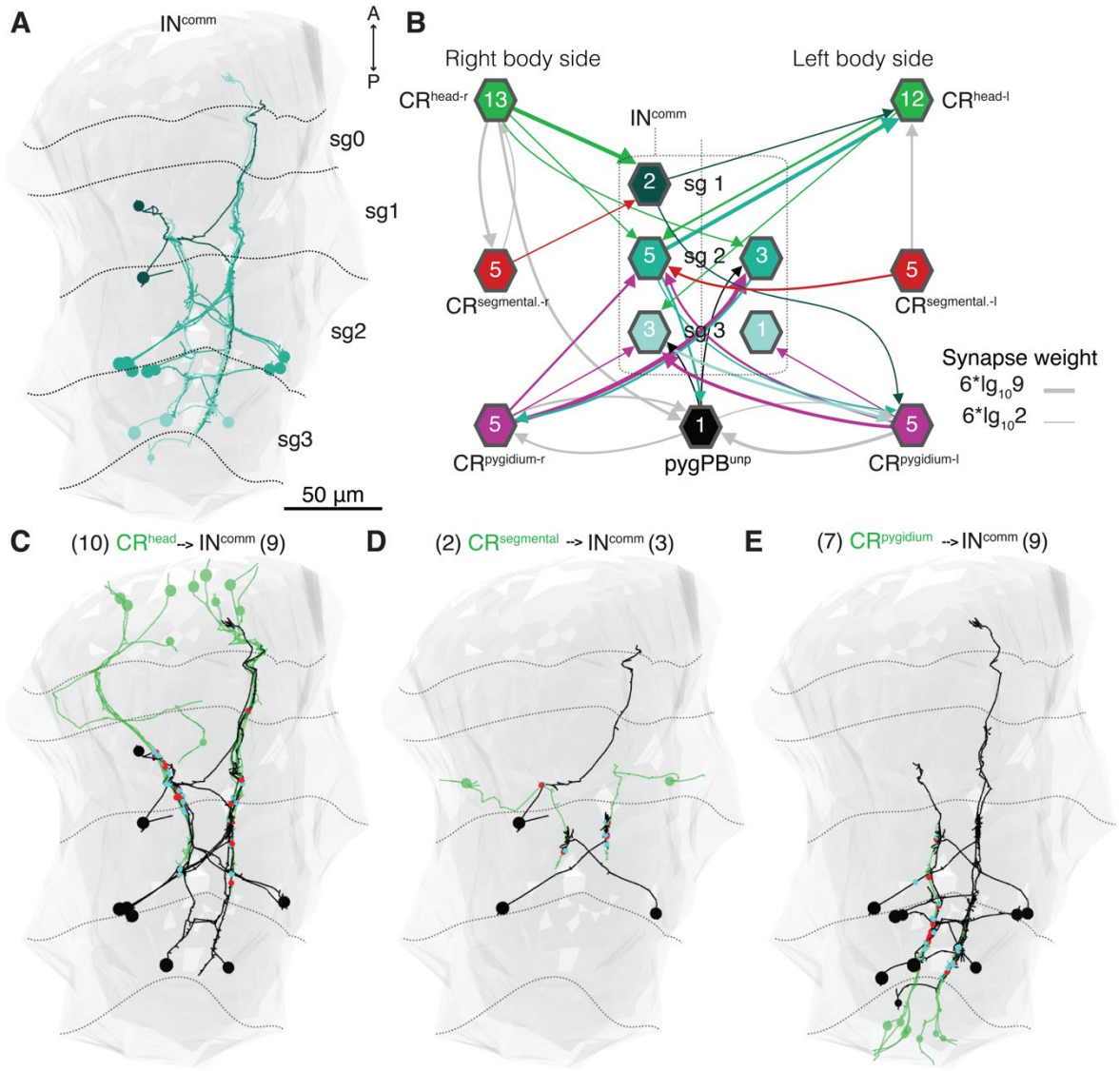


Figure 5-6 A diverse set of commissural interneurons is targeted by CRs. (**A**) Morphology and location of the IN^{comm} neurons targeted by CRs. (**B**) Wiring diagram showing the synaptic interactions between CRs and IN^{comm}. Interactions with less than 2 synapses were filtered out. Network symbol convention as defined in **Figure 5-3**. (**C**-**E**) Mapping of synaptic contacts (red/cyan dots) of CRs onto IN^{comm}. Only neurons in each group making synaptic contact are shown (number of cells in each group is indicated above each panel). Segmental boundaries in A, C, D, and E are marked with dotted lines and defined as in **Figure 2-1**. sg0-3: segment 0-3. Ventral view.

The Commissural interneurons

The second group is formed by 15 trunk commissural interneurons, named here IN^{comm} (**Figure 5-6**A). The only feature uniting these neurons in morphological terms is a neurite crossing the body midline. Aside from this, IN^{comm} showed a wide diversity in terms of neurite projection morphologies. For instance, some of them decussate almost perpendicular to the anterior-posterior axis, while others do so at an oblique angle. Some of these cells project anteriorly or posteriorly after decussating, while others bifurcate in the contralateral side of the VNC (**Figure 5-6**A).

Although CRs in all three groups target at least one of these neurons, the overall connectivity pattern is neither left-right symmetrical nor segmentally repeated (**Figure 5-6**B-E). This is probably due to the asymmetry in the number of IN^{comm} on each side of the body, which may hint that not

all the IN^{comm} neurons could be found. Despite the paucity of this data set, a few network motifs can be distinguished: CR^{head} and $CR^{pygidium}$ target IN^{comm} neurons on ipsilateral and contralateral sides. Moreover, at least some IN^{comm} neurons synapse back onto these CRs, thus forming feedback network motifs not seen among the other interneuron classes in the circuit (**Figure 5-6**B). $pygPB^{unp}$ shows sparse connections to this interneuron population (**Figure 5-6**B).

A circuit for CR-initiated ciliary control

The first layer of interneurons targeted by CRs show a distinct pattern of connectivity that lead to the ciliary and muscle effectors implementing the startle response (**Figure 5-3**).

A converging and left-right symmetric pathway for ciliary control was found from CR^{head} neurons targeting directly and indirectly the ciliomotor neurons involved in ciliary closures (**Figure 5-7**A). Firstly, CR^{head} neurons on both body sides directly target the MC (**Figure 5-7**A-D), a cholinergic ciliomotor neuron triggering prototroch closures (Verasztó et al. 2017). A couple of synapses are also made from one CR^{head} directly onto the prototroch (**Figure 5-7**A, D).

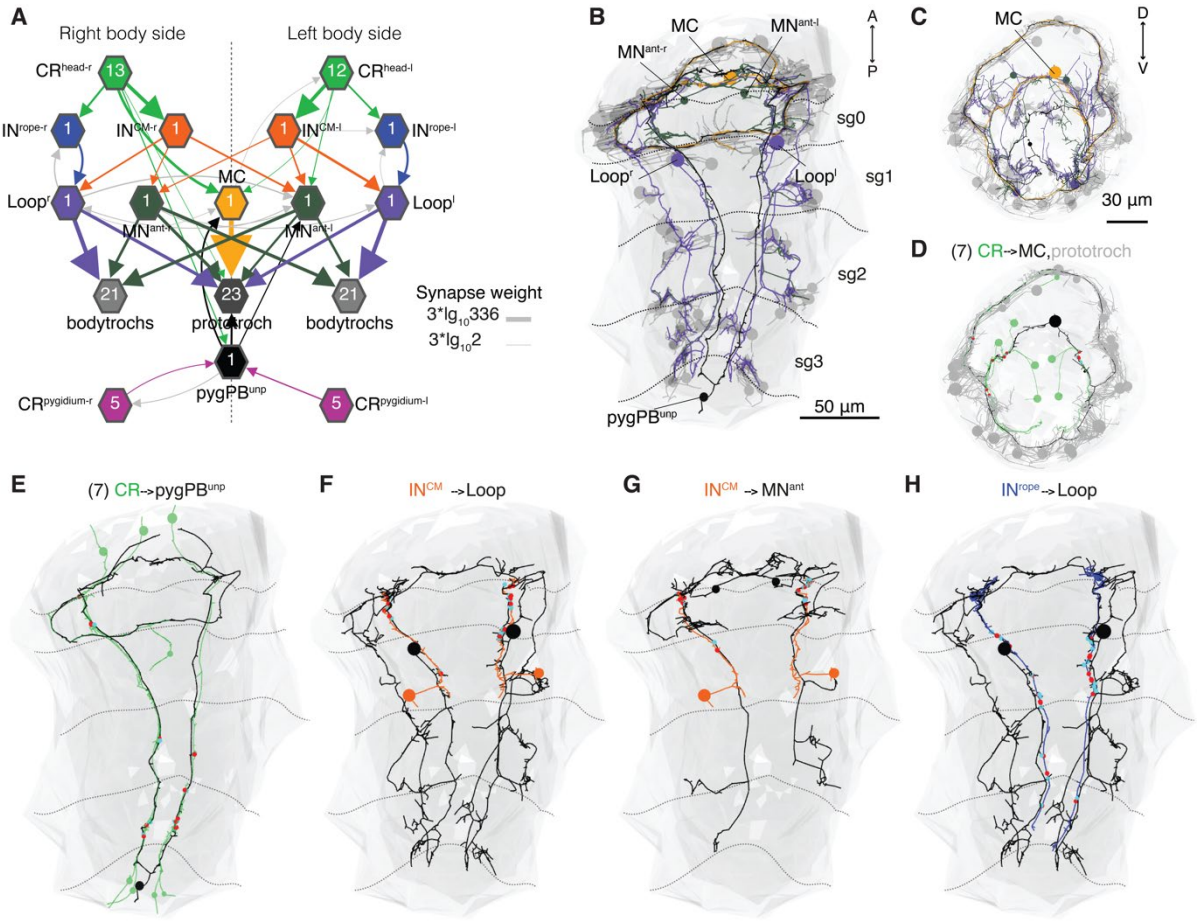


Figure 5-7 A circuit for CR-initiated ciliary control. (**A**) Wiring diagram highlighting the pathway from CRs to ciliary bands. Network symbol convention as defined in **Figure 5-3**. Each neuron group is sorted by segment and body side. Interactions with less than 2 synapses were filtered out. (**B**-**C**) Overview of neuronal morphology of ciliomotor neurons in the CR startle circuit: MC, Loop and MN^{ant}. (**D**-**H**) Mapping of synaptic interactions (red/cyan dots) from CRs to MC (D) or $pygPB^{unp}$ (E); from IN^{CM} to Loop (F) or to MN^{ant} (G); or from IN^{rope} to Loop (H). Segmental boundaries in B, E-H are marked with dotted lines and defined as in **Figure 2-1**. sg0-3: segment 0-3. Ventral view in B, E-H, and apical view in C and D.

An indirect pathway to control the prototroch and the rest of the ciliary bands was found through IN^{CM} and IN^{rope}, the interneurons targeted only by CR^{head} neurons (**Figure 5-7**A). IN^{CM} on each body side directly synapse onto ipsilateral Loop ciliomotor neurons, a pair of cholinergic cells innervating both prototroch and bodytrochs (Verasztó et al. 2017). Loop neurons show an activity tightly correlated with ciliary band closures, and thus they are hypothesized to have an analogous function to the MC cells (Verasztó et al. 2017). IN^{CM} also target MN^{ant} neurons, a putative head ciliomotor neuron type innervating all ciliary bands (Verasztó et al. 2017; Randel et al. 2015).
Interestingly, MN^{ant} have both ipsilateral and decussating neurite projections, and both are targeted by IN^{CM} on each body side (**Figure 5-7**A, G). IN^{CM} makes synaptic contacts onto MN^{ant} and Loop neurons using the distal part of its neurite, at the start of the circumesopagheal connective (**Figure 5-7**F-G).
IN^{rope} neurons on each side of the body also target ipsilateral Loop neurons (**Figure 5-7**A). IN^{rope} neurons synapse on the proximal part of both the ascending and the descending neurite projections of Loop neurons (**Figure 5-7**H). Loop neurons synapse back onto IN^{rope} thus forming a feedback motif (**Figure 5-7**A).
The only neural pathway found from pygidial CRs ($CR^{pygidium}$) to the ciliary bands was through the biciliated sensory cell $pygPB^{unp}$ (**Figure 5-7**A,E), which innervates the prototroch and also targets MC and MN^{ant}(**Figure 5-7**A; (Verasztó et al. 2017)). CR^{head} neurons also target $pygPB^{unp}$(**Figure 5-7**A, E). The role of $pygPB^{unp}$ in ciliary band control is so far unclear. $CR^{segmental}$ neurons do not show synaptic pathways to the ciliary bands.

The CR circuit for musculature control

The pathway leading from CRs to the musculature also consists of a feed-forward pathway through the first interneuron layer, and through direct connections to VNC motoneurons (MN^{VNC}) (**Figure 5-3**). The modules for muscle and ciliary control partially overlap at the interneuron level. While Split and Commissural interneurons only target musclemotor neurons, Rope interneurons target both the Loop ciliomotor neuron (**Figure 5-7**A) and several VNC motoneurons (**Figure 5-3, Figure 5-8**A-B, **Figure 5-9**). A total of 50 different types of musclemotor neurons are targeted by all neurons in the CR circuit. These cells are left-right symmetric and segmentally iterated (**Figure 5-8**B).

Pattern of connectivity of interneurons onto VNC motoneurons

Although IN^{rope}, IN^{split} and IN^{comm} interneurons target all the VNC motoneurons in the circuit, the connectivity and the targeted set drastically differs between each group (**Figure 5-8**-**Figure 5-9**).
The intersegmental nature of Rope interneurons is explained by the numerous synapses it makes onto VNC motoneurons in all three segments (**Figure 5-9**A). Each rope neuron synapses onto both contralateral and ipsilateral motoneurons in each segment (**Figure 5-8**A). The overall left-

right symmetry and segmented nature in the connectivity pattern is also observed between Split interneurons and the downstream motoneurons (**Figure 5-8**A). Even though Split interneurons are segmentally arranged, their motoneuron targets are not only located in the same segment, but also in the more anterior and posterior segments (**Figure 5-8**A, **Figure 5-9**D-F). However, most of these interneurons only target ipsilateral VNC motoneurons (**Figure 5-8**A). Commissural interneurons show a marked segmental and left-right asymmetry in the synaptic connections onto motoneurons (**Figure 5-8**A, **Figure 5-9**B-C). Moreover, the set of motoneurons targeted by these group is relatively small and greatly overlaps with the set targeted by IN^{rope} and IN^{split} (**Figure 5-8**C). In fact, most of the MN^{VNC} in the circuit are targeted by Rope interneurons, thus showing its potentially important role in initiating the startle response.

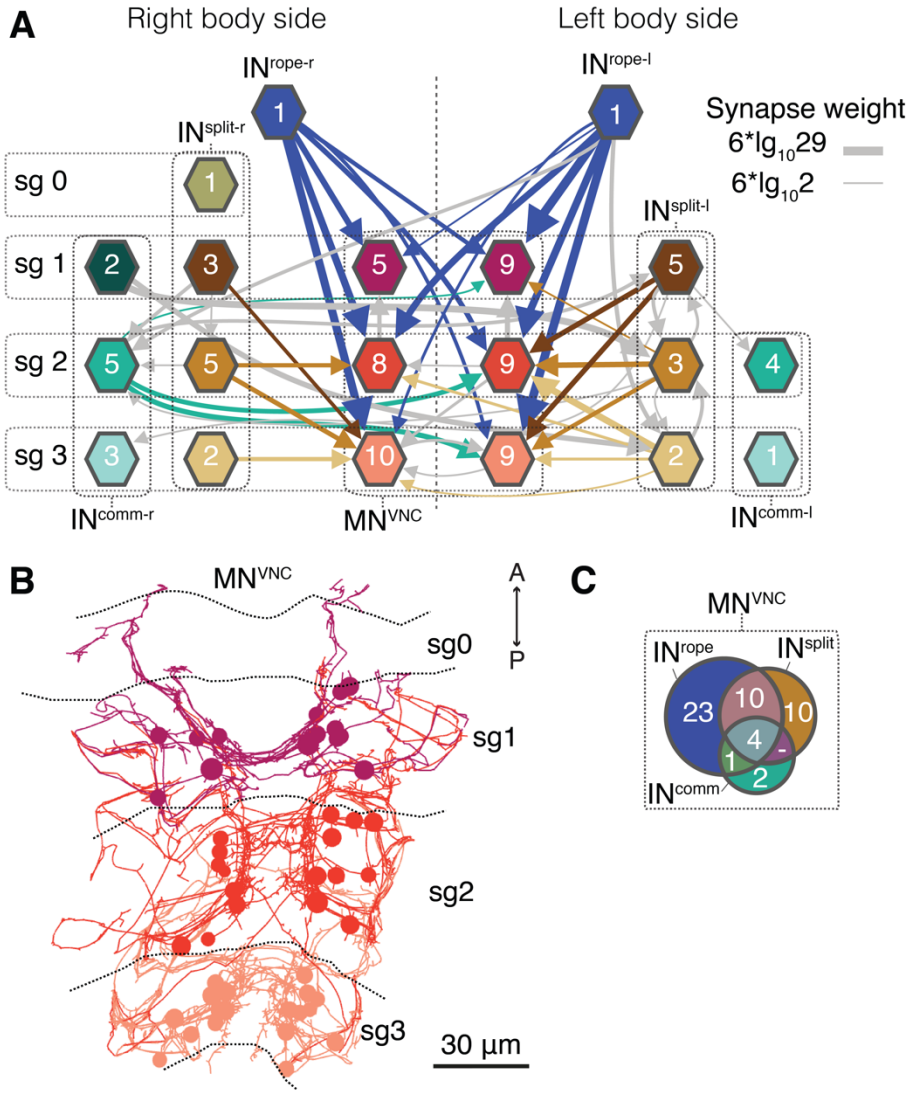


Figure 5-8 Interneurons in the CR circuit target several motoneurons in the VNC innervating whole body musculature. (**A**) Wiring diagram showing the synaptic interactions of IN^{rope} , IN^{split} and IN^{comm} with VNC motoneurons (MN^{VNC}). . Network symbol convention as defined in **Figure 5-3**. Each neuron group is sorted by segment and body side. Interactions with less than 2 synapses were filtered out. (**B**) Electron microscopy reconstruction of the VNC motoneurons in the CR circuit. Neurons are colored by segment and follow code in panel A. Segmental boundaries are marked with dotted lines and defined as in **Figure 2-1**. sg0-3:segment 0-3. (**C**) Venn Diagram showing the extent of overlap in MN^{VNC} targeted by IN^{rope}, IN^{split}, and IN^{comm}.

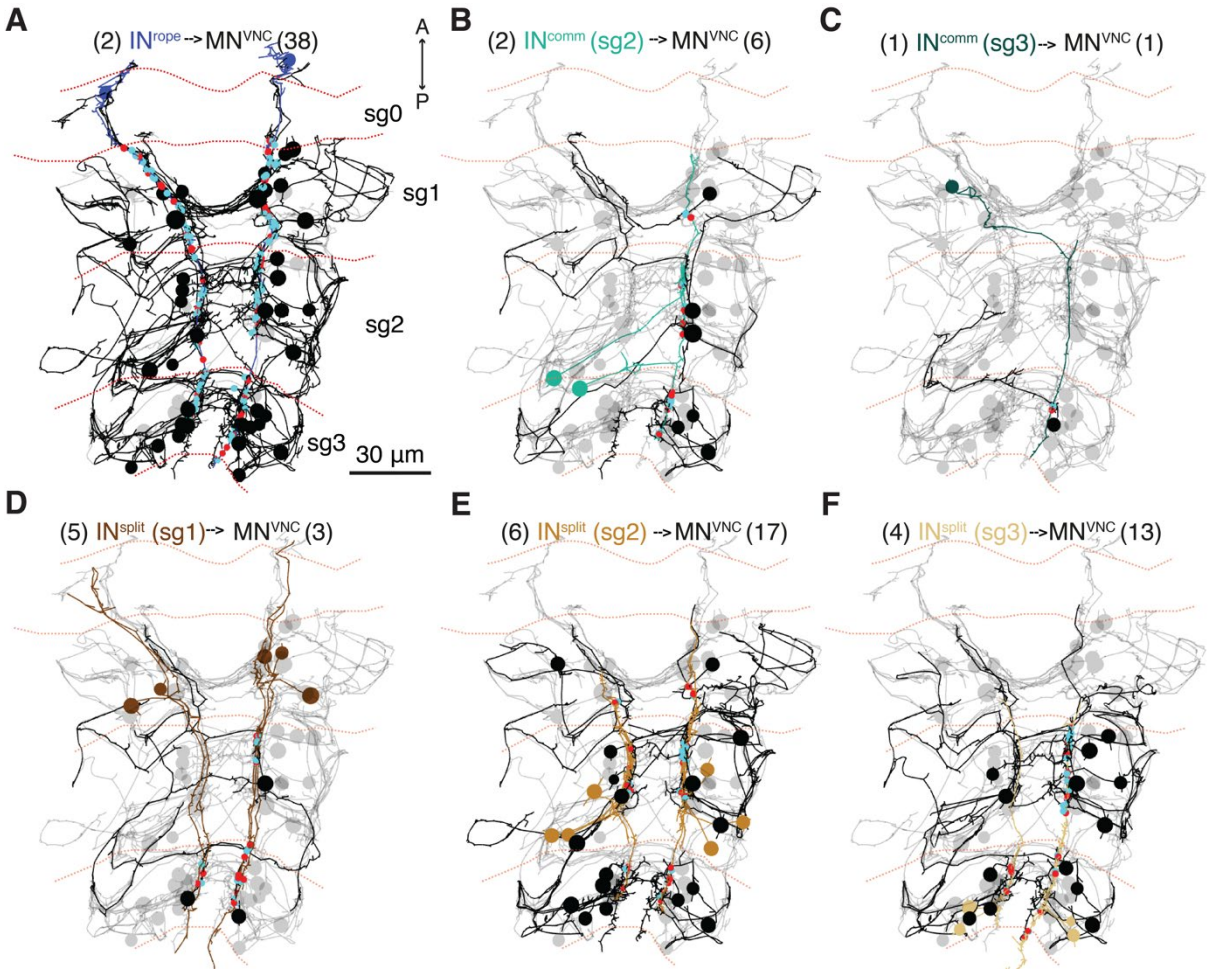


Figure 5-9 Rope, Split and Commissural interneurons synapse onto distinct sets of VNC motoneurons. (**A**-**F**) Mapping of synaptic contacts between Rope, Split and Commissural interneurons and VNC motoneurons in the CR circuit. Synapses are shown as red/cyan dots. All 50 MN^{VNC} are shown in each panel in light grey, and only those neurons targeted by the interneuron(s) displayed are shown in black. (**A**) IN^{rope} synapses onto 38 MN^{VNC}. (**B**-**C**) IN^{comm} in 2nd (B), and 3rd segment (C) together target 7 MN^{VNC} on the left side. Commissural interneurons in the 1st segment do not target any motoneuron. (**D**-**F**) IN^{split} in each segment target various MN^{VNC} in all segments. Segmental boundaries are marked with red dotted lines and defined as in **Figure 2-1**. sg0-3:segment 0-3. Ventral view in all panels.

Muscles innervated by VNC motoneurons

The 50 motoneurons in the CR circuit innervate somatic muscles in the whole larva **(Figure 5-10**). The muscles belong to at least five different classes: longitudinal, transverse, oblique, parapodial and mouth muscles (**Figure 5-10**A-B). Dorsal and ventral longitudinal muscles (sometimes called ventrolateral, or dorsolateral muscles) cause contractions in the antero-posterior axis, while transverse muscles do so in the dorso-ventral axis. Parapodial muscles include both chaetal and acicular muscles in both notopodia and neuropodia. Additional muscles lining the forming mouth and forming part of the axochord (Lauri et al. 2014) were also targeted by the MN^{VNC} set. This group is not necessarily functionally related and was thus only merged for convenience. A more thorough description of the larval musculature is part of another doctoral thesis and it is beyond the scope of this study (S.Jasek and G.Jékely).

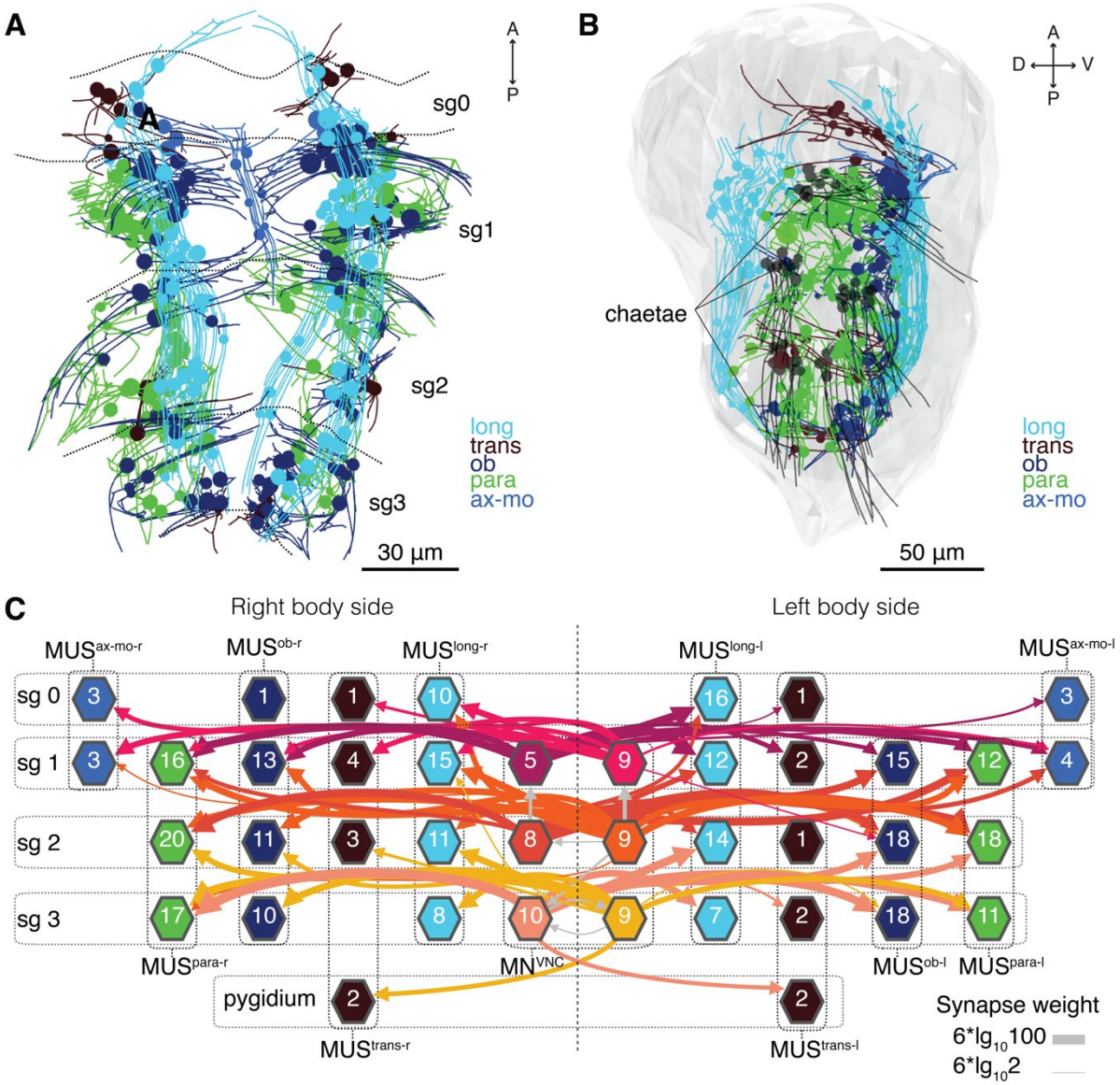


Figure 5-10 Muscles in the CR circuit. (**A**-**B**) EM volume reconstructions of muscles targeted by MN^{VNC} in the CR circuit. Muscles are colored by type. 5 muscle types are distinguished: longitudinal (long), transverse (trans), oblique (ob), parapodial (para), and a mix of axochord, mouth and chin (ax-mo) muscles. Ventral (A) or side (B) view. (**C**) Wiring diagram showing the synaptic interactions between MN^{VNC} and the different muscle types (MUS). Network symbol convention as defined in **Figure 5-3**. Each neuron group is sorted by segment and body side. Interactions with less than 2 synapses were filtered out. Segmental boundaries are marked with black dotted lines and defined as in **Figure 2-1**. sg0-3:segment 0-3.

VNC motoneurons target each of these muscle classes in distinct patterns. Longitudinal and oblique muscles are targeted exclusively by contralateral motoneurons in the same and in posterior segments (**Figure 5-10**C). Transverse muscles show a similar pattern, albeit supported by fewer synapses. Parapodial muscles are innervated by both contralateral and ipsilateral MN^{VNC} neurons. Splitting parapodial muscles into chaetal and acicular muscles did not show a different pattern (not shown).

Diversity of VNC motoneurons in the CR circuit

The VNC motoneurons in the CR circuit belong to at least 10 morphologically different types not previously described at the nectochaete stage (**Figure 5-11**). Thus, they were given names evocating

their overall morphology. IN^{rope} targets 8 out of the 10 MN types, most of them also targeted by IN^{split} neurons, despite the fact the latter interneurons target a smaller number of motoneurons (**Figure 5-11**A, **Figure 5-8**C). IN^{comm} neurons only target four out of the 10 types, but they exclusively target the bow motoneuron type (MN^{bow}). Of note, both CR^{head} and CR^{pygidium} neurons directly make synaptic connections onto the crab motoneurons (MN^{crab}) (**Figure 5-11**A). In fact, this motoneuron type is targeted by other important cells in the circuit including Rope, Split and pygPB^{unp}.

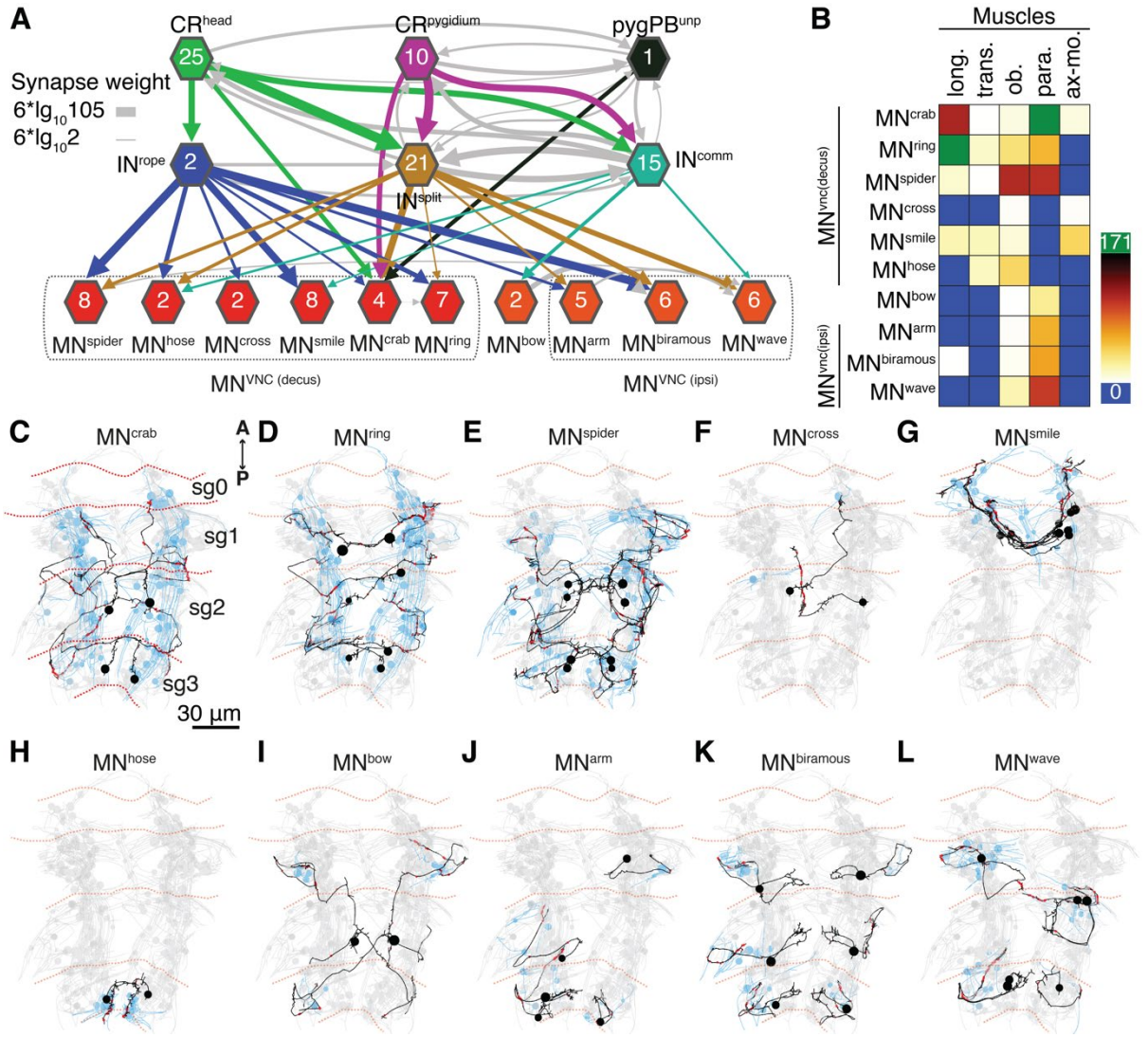


Figure 5-11 Multiple types of VNC motoneurons are part of the CR circuit. (**A**) Wiring diagram showing the synaptic interactions of CRs, pygPB^{unp} and the interneurons with the different VNC motoneuron types in the CR circuit. Network symbol convention as defined in **Figure 5-3**. VNC motoneurons are sorted by type. (**B**) Connectivity matrix of VNC motoneurons types and muscle types. See **Figure 5-10** for definition of muscle types. Color scale is shown to the right from 0 synapses (blue) to 171 synapses (green). (**C**-**L**) Electron microscopy reconstruction of the 10 different motoneuron types and the muscles they innervate. Motoneurons are colored in black, and the muscles targeted are colored in blue. Muscles in the CR circuit not targeted by the particular group of motoneurons are shown in light grey. Motoneuron synapses are shown as red dots. MN$^{VNC(decus)}$: decussating VNC motoneurons; MN$^{VNC(ipsi)}$: ipsilateral VNC motoneurons. Segmental boundaries in C-L are marked with red dotted lines and defined as in **Figure 2-1**. sg0-3:segment 0-3.

As it may have been inferred from the previous sections, the MN^{VNC} set consists of decussating and ipsilateral motoneurons (**Figure 5-11**C-L). Decussating motoneurons target most of the muscle types, while ipsilateral motoneurons target mainly parapodial and oblique muscles (**Figure 5-11**B). The contralateral motoneurons MN^{crab}, MN^{ring} and MN^{spider} contribute most of the neuromuscular synapses and target the greatest number of muscles (**Figure 5-11**B-E). MN^{bow}, the exclusive target of IN^{comm}, has both an ipsilateral and decussating neurite projections (**Figure 5-11**I).

Some of the decussating and all ipsilateral motoneurons are iterated in at least two of the three segments, such as MN^{crab} in the former, and $MN^{biramous}$ in the latter category (**Figure 5-11**C-L). The muscles targeted by ipsilateral motoneurons are restricted to the same segment while decussating motoneurons target muscles in the same and in adjacent segments.

The left right coordination of muscle contraction

Perhaps one of the most important hallmarks of the startle response in *Platynereis* regarding muscle control is the remarkable left-right symmetry in the extent and timing of parapodial elevation (Chapter 1). The wiring diagram of the CR circuit shows different left-right symmetry motifs (**Figure 5-12**).

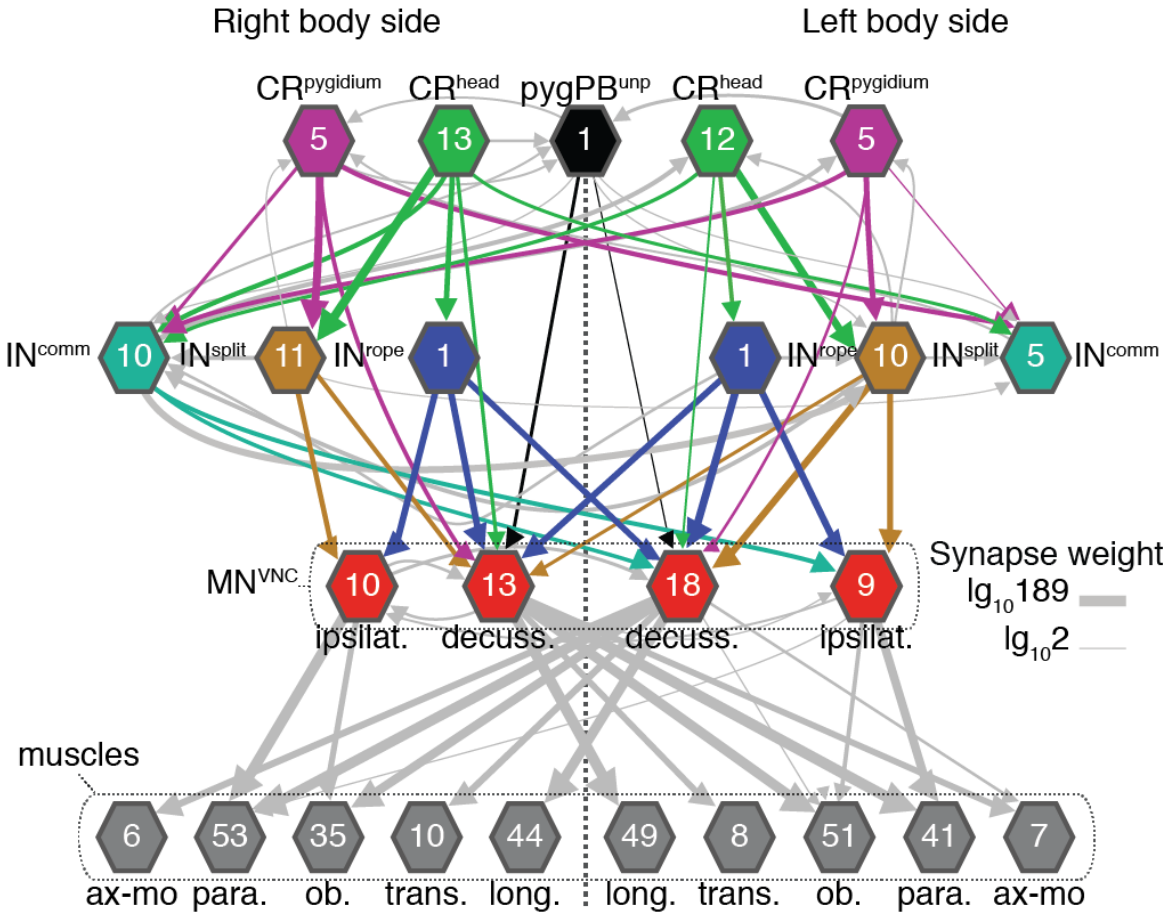


Figure 5-12 Circuit motifs for left-right muscle coordination. Wiring diagram of the CR circuit showing the different neuron types and their synaptic connections sorted by body side. VNC motoneurons (MN^{VNC}) are grouped by ipsilateral (ipsilat.) and decussating (decuss.) types. Network symbol convention as defined in **Figure 5-3**.

As shown before, CR^{head} and $CR^{pygidium}$ target the decussating motoneurons MN^{crab} (**Figure 5-11**A). This connection is present on both the left and right body sides (**Figure 5-12**). pygPBunp also bilaterally targets this motoneuron type (**Figure 5-12**). IN^{rope} was also shown to target VNC motoneurons on the same side and on opposite body sides (**Figure 5-8**A). Grouping the motoneuron set into the decussating and ipsilateral motoneurons identified in the CR circuit further shows that Rope interneurons does not differentially target one type over the other, targeting ipsilateral and decussating motoneurons on the same side, and contralateral neurons on the opposite body side (**Figure 5-12**). IN^{split} neurons target mostly motoneurons that lie on the same body side (**Figure 5-8**A), including both ipsilateral and decussating motoneurons (**Figure 5-12**).

Muscles activated during the startle response

The startle response involves the elevation of parapodia, a behavior that is expected to require at least the activation of parapodial musculature. The startle circuit from CR neurons shows synaptic pathways not only to parapodial muscles, but also to most other major muscles classes (**Figure 5-10**). This possess the question as to whether the other muscle types also contribute to the startle behavior. Only functional analysis could determine which muscle types are activated during the behavior.

To tackle this challenge, tethered larvae injected with mRNA encoding the dual reporter tdTomato-P2A-GCaMP6s (used in Chapter 2) were stimulated with a vibrating filament while the musculature was imaged for detecting any change in fluorescence during the startle response. The signal could not be quantified due to the drastic movements of the animal in all three axes. Only a descriptive account of the results is presented.

Vibrations of the filament strong enough to trigger the startle response on tethered animals led to an increase in GCaMP6s fluorescence in longitudinal muscles on both body sides (**Figure 5-13**A-B). Bilateral activation of oblique muscles in the first segment also occurred upon stimulation (**Figure 5-13**C-D). Parapodial muscles also show an increase in GCaMP6s signal during stimulation (**Figure 5-13**E-F). In this case, the elevation of the parapodia could be directly confirmed, and thus the occurrence of the response verified. Muscles in the $MUS^{ax\text{-}mo}$ category also were activated upon stimulation (**Figure 5-13**A-B). For technical reasons, transverse muscles could not be clearly identified and imaged. These results indicate that all the muscle types in the CR circuit are activated during the startle response.

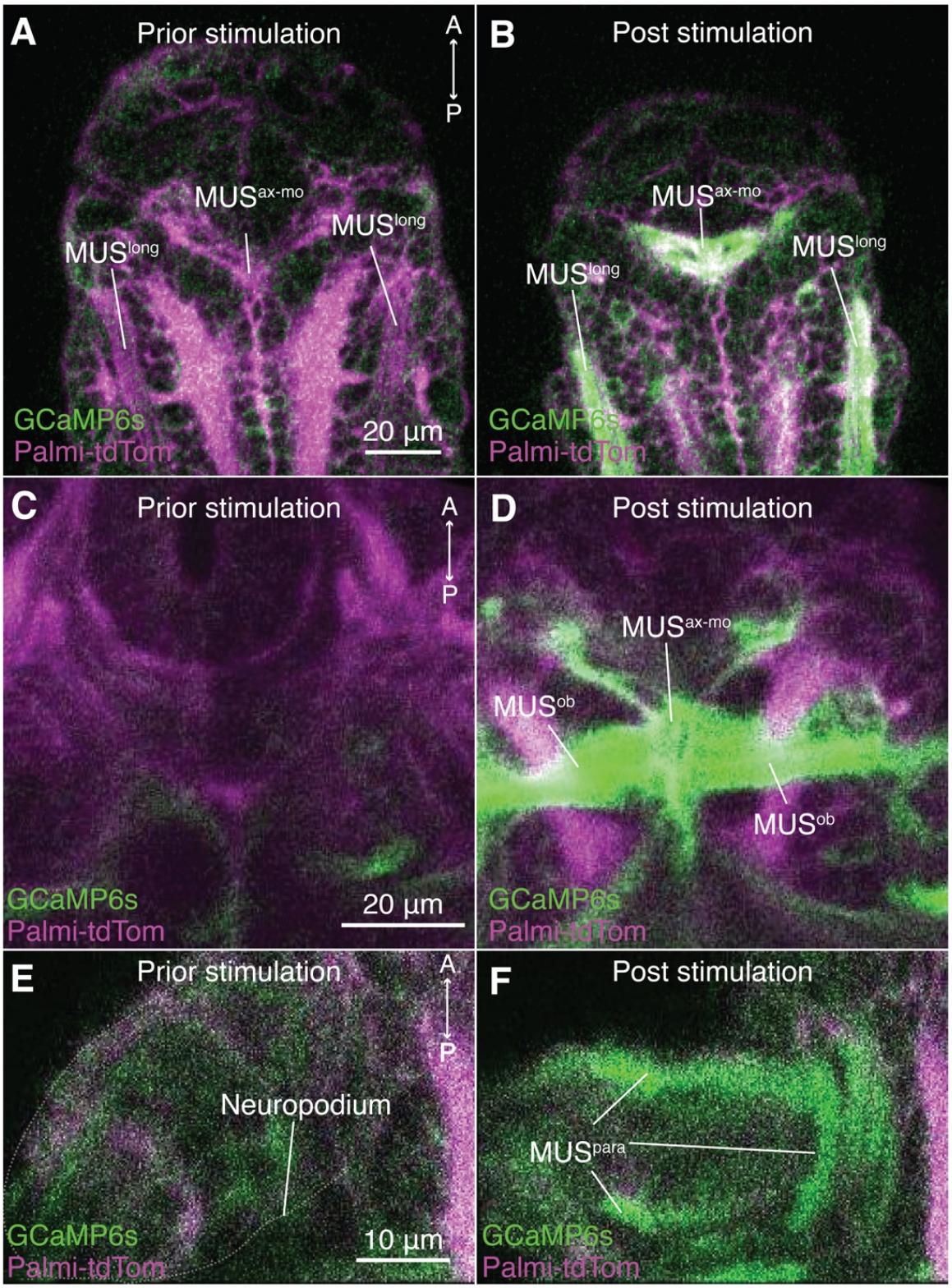


Figure 5-13 Muscles activated during the startle response. (**A**-**F**) Merged confocal microscopy snapshots of GCaMP6s and Palmi-tdTomato channels showing nectochaete larva prior (A, C, E) and after (B, D, F) anterior stimulation with a vibrating filament. (**A**-**B**) Plane at the level of the ventral nerve chord showing longitudinal muscles on both body sides (MUS^{long}). Note the increase in GCaMP6s fluorescence upon stimulation. Chin muscles (part of the $MUS^{ax\text{-}mo}$ category) also show an increase in fluorescence. (**C**-**D**) A deeper optical section at the level of the forming stomodeum in the first segment. Upon stimulation the movement of the larva brings the oblique muscles (MUS^{ob}) and axochord muscles ($MUS^{ax\text{-}mo}$) in focus (D). These muscles show high GCaMP6s signal. (**E**-**F**) Close up view of the neuropodium on the first segment. Upon stimulation, the extremity is elevated and muscles inside it increase in GCaMP6s fluorescence. All panels show a ventral view.

Discussion

Assessing the quality of the reconstruction

The CR circuit adds to the growing list of circuits reconstructed at the synapse level. The reconstruction was challenged by the problems that are idiosyncratic to this technique: including sample size, material loss, tracing and synapse annotation errors. Material loss potentially affected the reconstruction of the pygidial CR circuit, as the CRs on the pygidial cirrus on one side could not be unambiguously identified. It may also have prevented the accurate reconstruction of the single pair of Arc interneurons, as no consistent pathway could be found downstream to these cells. Other groups with higher number of cells were more robust to loss of material. On the other hand, new data sets will be needed when more specific conclusions on the contributions of particular neurons are needed, as the loss of sections in some parts of the stack may have led to the switch of the identity between neurons within one class (as neurons of the same group often projected next to each other).

The use of an arbitrary synapse cut-off also could have caused failure to include weakly connected partners, but the identification in many cases of bilateral cell types indicates that although conservative, the cut-off did allow to recover most of the relevant targets. CRs show both weak and strong interactions with different groups of cells. The bilateral symmetry in the strength of most of these interactions suggests that even when they are weak, the interactions observed are real and have a physiological basis (Jarrell et al. 2012; Schikorski and Stevens 1997).

Thus, despite problems and limitations, the number of features mentioned above indicate that the reconstruction can be trusted and can be used to formulate the first hypotheses on the neuronal implementation of the startle response mediated by activation of CRs.

CR-dependent pathways for ciliary control

Direct and indirect pathways converging on ciliomotor neurons

A synaptic pathway was found from CRs on the episphere (CR^{head}) to the two main candidate neurons controlling ciliary arrests in *Platynereis*: the MC and the Loop ciliomotor neurons (Verasztó et al. 2017). The MC cell has been conclusively shown to be responsible for most of the prototroch closures. A number of CR^{head} neurons directly synapse onto this neuron and thus may activate it to induce prototroch closures upon mechanical stimulation. The weak bilateral CR->MC interaction may reflect a low activation threshold, or a high neurotransmitter release probability that may make fire MC even when not all CRs are activated. This in turn could explain the low stimulus threshold needed to trigger closures upon anterior stimulation.

CRs in the episphere did not directly innervate Loop neurons—which are thought to elicit ciliary closures of the bodytrochs—even when the neurites of both cell types overlapped. Instead, two

parallel interneurons pathways converging onto Loop neurons could explain how bodytrochs are arrested upon anterior stimulation. The first pathway would require CRs to activate the single pair of CM interneurons—the most strongly connected of all direct CR targets—which in turn synapse onto Loop neurons and could activate these cells to drive bodytroch arrests. CM interneurons may need to integrate the signal from a bigger pool of CRs before firing. Alternatively, the high synaptic weight may reflect the need to ensure IN^{CM} fires more reliably.

The second pathway through the single pair of Rope interneurons is synaptically much weaker, but it nonetheless involves a similar number of CRs in the episphere as those targeting CMs. Thus, Rope interneurons may also integrate the CR signal to then drive the activation of Loops on each side of the body. The two parallel pathways may work in different contexts or be activated at different stimulus thresholds (see below).

No direct or indirect pathway to MC or Loop neurons could be found from CRs in other parts of the body, thus further supporting the differences observed in the profile of ciliary arrests upon anterior and posterior stimulation. Although it may require further corroboration, this observation means that there is not a specialized mechanism for $CR^{pygidium}$ neurons for triggering ciliary arrests on their own.

Possible mechanisms for synchronous ciliary band closures

Prototrotroch and bodytroch synchrony

Synchronous ciliary band closures observed upon anterior stimulation could rely on the network motif between CR^{head} neurons and MC and Loop neurons. The motif suggests that in principle it would suffix to activate even a single CR^{head} neuron (like hCR1 or hCR2) to arrest all the ciliary bands. The action of a single neuronal population also explains why the bodytroch never fails to close when the prototroch closes upon anterior stimulation. Additional synaptic complexities are probably required to ensure such reliability, but the core system could be straightforward. The presence of an interneuron layer between CRs and Loops also could explain the slight delay between prototroch and bodytroch closures. Thus, unlike it was hypothesized in Chapter 1, the putative pacemaker neurons thought to coordinate ciliary closures are not part of the startle circuit. This latter system may rather be used for less pressing stimuli, such as changes in internal states driven by neuroendocrine factors (Williams et al. 2017). CRs shortcut this circuit and target MC and Loops via a dedicated pathway.

Bilateral synchrony in ciliary arrests

The observed left-right body synchrony of ciliary arrests could be a consequence of the bilateral symmetry in the network motif between CR^{head} neurons and the cholinergic ciliomotor neurons as long as the stimulus is bilaterally synchronous, too. Based on the proximity of each CR to its bilateral pair, and on the small size of the larva it is likely that the stimulus always reaches the

sensory dendrite of CRs on both sides at practically the same time. Nonetheless, the presence of other motifs in the circuit suggest that it needs a way to ensure both sides of each ciliary band are arrested. This could be the function of the ciliomotor neuron MN^{ant}, an additional target of CM interneurons that is also thought to induce ciliary arrests (Verasztó et al. 2017). Each MN^{ant} innervates bodytrochs on both body sides, which as discussed in Chapter 1 cannot coordinate arrests using gap junctions as they do not form a continuous band of cells. Thus, MN^{ant} could serve as an alternative pathway to ciliary arrests ensuring bilateral synchrony when the stimulus itself fails to synchronize both body sides.

CR-dependent pathways for muscle control

Calcium imaging showed that extensive muscle contraction occurs during the startle response. Not only parapodial, but also most of the main categories of fast-contracting (i.e. striated) muscles (Brunet et al. 2016) showed increased GCaMP6s fluorescence. Thus, the startle response in *Platynereis* requires a system to control the musculature of the whole body.

Direct and indirect pathways for whole-body muscle contraction

Direct pathway to the Crab motoneurons

The CR wiring diagram showed various possible neuronal mechanisms by which the muscles can be activated. A direct pathway from either CR^{head} or $CR^{pygidium}$ neurons to a VNC motoneuron type, the MN^{crab} is in principle sufficient to activate longitudinal muscles in all segments. On the other hand, it would not explain the complete contraction of all the other muscles that a full-fledged response requires. Preliminary experiments have shown that upon weaker anterior stimuli, longitudinal muscles are activated even when there is no elevation of parapodia (unpublished observations). Thus, this direct pathway may have a low-activation threshold, perhaps as a way to prime muscle contraction in the case additional input from indirect pathways is received.

Indirect pathway: The Rope interneurons as command-like interneurons

The first indirect pathway to muscle contraction involves the intersegmental Rope interneurons. Each of the two Ropes targets both ipsilateral and decussating VNC motoneurons in all three segments, thereby potentially leading to the full-muscle contraction characteristic of the response. This connectivity pattern strongly resembles that of interneurons playing critical roles in other startle circuits, such as the M cell in fish, or the LG/MG giant fibers in crayfish escape circuits (Korn and Faber 2005; Edwards et al. 1999). Ropes may thus have an analogous function, and similar firing properties. For instance, after activation, a single spike of Rope interneurons may be all what is needed to drive a fast and robust activation of VNC motoneurons across segments, as it is observed upon activation of the M cell or other command-like interneurons in startle circuits (Nicol 1948; Zottoli 1977; Thomas and Wyman 1984). Ropes may thus need to have a mechanism of coincident detection in place to be activated only upon synchronized input (König et al. 1996).

This input would take the form of a more abrupt water disturbance that univocally indicates the presence of an incoming threat.

Indirect pathway: The Split interneurons act as a population to drive whole-body muscle contraction

The wiring diagram indicates that only water disturbances sensed by CRs in the episphere would activate the IN^{rope} pathway. Ropes may thus not be the main pathway upon stimulation from other sides of the body, as it is the case in crayfish, where two different giant fiber systems work depending on the site being stimulated(Edwards et al. 1999). $CR^{pygidium}$ neurons do not target this interneuron and thus may rather elicit whole body muscle contraction through an alternative pathway. This may be the function of the population of segmentally iterated interneurons here named Split interneurons. As a population, IN^{split} neurons control VNC motoneurons in all segments and could in principle cause the contraction of the whole musculature involved in the response. The use of different converging pathways based on the site of stimulation is a well-known circuit design principle in other startle responses (Hale et al. 2016). Startle response pathways need to be tailored to elicit an adequate response based on the site of stimulation.

In *Platynereis,* the only observed significant difference between parapodial elevation elicited upon anterior and posterior stimulation is the threshold and the latency of its activation. This could be thus explained by the Split interneuron pathway having a higher activation threshold than Rope interneurons. This may be achieved in part by the different nature of both pathways: unlike Rope interneurons, a single Split interneuron, or even all Split neurons in one segment may not be able to make VNC motoneurons to spike; instead, the Split population across all segments may be required as a whole to drive a full response. Such mechanism is supported by the fact that CRs target Split in all segments, and Splits synapse on each other, perhaps reinforcing their activation. It could also explain the longer time needed to initiate parapodial elevation upon posterior stimulation relative to anterior stimulation.

Partially overlapping circuits to drive whole-body muscle contraction

Adding to the complexity of the model, is the fact that not only CRs in the pygidium, but also those in the episphere and in the segments synapse onto IN^{split} neurons. Thus, the Rope and Split interneuron pathways may not only be required when the larva is stimulated in different sites, but they may also act in parallel or in series to each other. Rope and Split interneurons may act in series to ensure a full contraction of all the musculature. In fact, the VNC motoneuron population targeted by Rope and Split interneurons does not fully overlap. The intensity, and perhaps nature of the stimulus may alternatively dictate which neuronal pathway is chosen during startle response. If as hypothesized, Rope interneurons are more sensitive than Split interneurons, the former pathway may be activated at lower stimulus intensities than the latter upon anterior stimulation.

The activation of parallel pathways according to the nature of stimulus is another emerging principle from startle circuits (von Reyn et al. 2014; von Reyn et al. 2017).

Both parallel and serial mechanisms may be at work during the startle response; the pathway through the Split interneurons may be used by all CRs to drive a WideE-type response, but CR^{head} neurons may additionally use Rope interneurons to shorten initiation latency and increase sensitivity. If the startle response triggered by anterior and posterior stimulation are implemented by partially overlapping circuits, different habituation profiles could be expected as in the *C.elegans* response to anterior and posterior stimuli (Chen and Chalfie 2014). Moreover, artificial inactivation of Rope interneurons could lead to WideE elevation upon anterior stimulation with more similar temporal parameters to the WideE-type response elicited by posterior stimulation.

Additional mechanisms for intersegmental muscle contraction

The intersegmental nature of Rope interneurons provides a simple and commonly observed solution to trigger synchronous contraction along the anteroposterior axis. Split interneurons across different segments summing their inputs on motoneurons across the VNC also provide a possible mechanism for synchronous intersegmental muscle contraction. Nonetheless, additional network motifs could contribute to ensure the tight synchrony in WideE parapodial responses.

One such mechanism could be at the level of the VNC motoneurons. The motoneurons in one segment targeted muscles not only within their own segment, but also in the anteriorly adjacent segment. This is mainly the result of the innervation pattern of MN^{crab} and MN^{spider}, two of the motoneuron targets of both Split and Rope interneurons that are serially repeated in the 2^{nd} and 3^{rd} segment and innervate all muscles classes likely important for the response. Thus, these motoneurons can provide bi-segmental coordination and thus resulting in the short elevation delays between neighboring segments.

Network motifs for coordinated bilateral muscle contraction

Synchronous bilateral contraction

Another hallmark feature of the startle response not commonly addressed in other startle circuits is the bilateral synchrony in muscle contraction. In crayfish and in *Drosophila* synchronous bilateral muscle contraction is thought to be achieved by electrical coupling between homologous giant neurons on both body sides (Phelan et al. 1996; Edwards 2017; Glantz and Viancour 1983). As Split and Rope interneurons are exclusively ipsilateral interneurons, an additional mechanism must be at work in *Platynereis*.

As suggested for explaining bilateral synchrony of ciliary arrests, one possible mechanism to achieve synchronous bilateral muscle contraction could rely on the symmetry in the wiring of the Rope and the Split interneuron circuits. Under the simplest model, a coincident stimulation on left and right receptors would lead to synchronous activation of the circuits necessary for whole body

muscle contraction on each body side. In other words, both body sides work independently but the effect is an apparent temporal coordination.

Although intuitively possible, the experiments performed were not necessarily controlled to create uniform stimuli. Moreover, as shown in Chapter 1, lateralized stimuli also triggered a bilaterally synchronous response. Thus, other mechanisms are required to ensure bilateral synchrony even upon nonsynchronous stimulation. Rope interneurons on each side of the body target VNC motoneurons on the same side and on the opposite side of the body. This motif would ensure that even in the extreme case where only one side is activated (from the anterior side), the normal bilateral response would be displayed. Split interneurons only showed a partially similar connectivity motif. Therefore, it is not clear how this interneuron population could ensure synchronous bilateral contraction.

To rule out symmetry in the stimulus as the main explanatory factor, the circumesophageal connectives on only one side of the body could be ablated, thereby inactivating the CR afferents and the Rope on that side of the body. If anterior stimulation in the ablated animal still elicits a normal response, then Ropes may be using a mechanism such as the one suggested to ensure bilateral synchrony.

Extent of muscle contraction

A mechanism in which a single Rope neuron activates VNC motoneurons on both body sides would also explain the symmetry observed in the extent of muscle contraction (expressed in the apparently similar angle at which parapodia are elevated on each side of the body). However, ipsilateral VNC motoneurons, which innervate parapodial muscles, are not targeted by Rope interneurons on the opposite side. The only group innervating both ipsilateral and decussating VNC motoneurons from the opposite side are the enigmatic IN^{comm} neurons. In *Drosophila,* ablation of Even-Skipped(Eve)--expressing commissural interneurons alters the symmetry in the extent of bilateral muscle contraction (Heckscher et al. 2015). The authors of that study suggest Eve^{+} commissural interneurons are conserved across bilaterians, including *Platynereis* (Denes et al. 2007). It may thus be that at least some of the IN^{comm} neurons are Eve^{+} and regulate bilateral symmetry of muscle contraction during the startle response.

Neuronal pathways for coincident and independent control of muscle and cilia

As discussed above, the CR wiring diagram has separate pathways to control muscles and cilia, each accounting entirely or in part for the observed features of the startle response. For instance, a clear prediction of the circuit is that MC and CM interneurons are needed to trigger closures independently of parapodial elevation upon weak anterior stimulation.

On the other hand, some of the behavioral features of the response include correlated events of ciliary and muscle activity that also require explanation. One of them is the similitude in latency in

the initiation of ciliary arrests and parapodial elevation upon anterior stimulation. IN^{rope} was the sole interneuron type in the circuit targeting both ciliomotor and musclemotor neurons. These neurons thus could explain how the timing of ciliary closures and muscle contraction can be coordinated to the millisecond. In contrast, the CR wiring diagram did not reveal a direct pathway to explain the synchrony in bodytroch ciliary arrest and initiation of parapodial elevation upon posterior stimulation. It is actually not clear why ciliary closures occur at all upon posterior stimulation.

Novel neuronal types in the CR circuit

The neurons downstream of CRs included previously unreported neuronal cell types. Thus incidentally, the CR circuit provides one of the first views of the anatomical diversity of interneuron and motoneuron types in the nectochaete larva. Such diversity is consistent with the complex molecular fingerprint in *Platynereis* VNC (Vergara et al. 2017). It will be crucial to reconcile the morphological with the molecular diversity to assign the specific cell type identity and neurotransmitter expression to each of the neurons in the startle circuit. So far, parallels based on morphology alone can be drawn to previous studies on annelid VNC neurons.

Rope interneurons as precursors to giant neurons

The Rope interneurons are so far only known to be targeted by CRs and have as main targets VNC motoneurons. Thus, their main function could be restricted to drive a robust whole-body muscle contraction upon mechanical stimuli. As noted, Rope neurons may functionally resemble giant neurons in other startle circuits. It is early to postulate any evolutionary similarity between these interneurons, but the repeated appearance of giant-like neurons in startle circuits requires a closer comparison, perhaps at the molecular level.

A developmental comparison could be made between Rope interneurons and the giant neurons in adult *Platynereis* worms (Smith 1957). These fibers drive escape reflexes involving parapodia elevation responses (Horridge 1959). Adult *Nereis* has a system of three sets of giant fibers, one pair of paramedial, another of lateral, and one single medial fiber, each activated by stimulation from different locations (Bullock 1948). The fibers are formed by anastomosis of several neurons (Smith 1957; Nicol 1948). The medial giant is formed by numerous cells lying in the ventral side of the sub-esophageal ganglion (Sigger and Dorsett 1986b). As this fiber is activated upon anterior stimulation, it is tempting to speculate that Rope neurons are but one of the precursors of this fiber. Genetically labelling the Rope neurons would be a way to trace their development and link them to giant fibers in the adult.

Interneurons types

In a similar logic, the Split interneurons may bear some developmental relationship with the paramedial fibers, which are sensitive to posterior stimulation, and are formed by segmentally

repeated units (Sigger and Dorsett 1986b).On the other hand, anatomical studies in the adult VNC identified pseudounipolar cells such as the Split or the CMs (see Figure 5 in (Smith 1957)). Those bifurcating cells also showed neurite "arms" of different size. A few of the more complex morphologies of commissural interneurons were also identified in that report, specifically the long decussating ascending interneurons. No physiological information could be found about these cells and whether they have any role in the startle reflex of the adult. In any case, the report by Smith adds support to the morphologies of the neurons reconstructed in the larva.

Motoneuron diversity

Smith also reports 7 types of either decussating or ipsilateral motoneurons with similarities to those reported here. One of them, the 4B motoneurons, are very similar in morphology to the crab motoneurons (MN^{crab}) (Sigger and Dorsett 1986a), namely in having a neurite running anteriorly along the VNC before decussating and branching into an axon innervating dorsal and ventral longitudinal muscles (see Figure 21 in (Smith 1957)). The 4B motoneurons have a giant axon themselves, are contacted by the lateral and medial giants (but not by the paramedial giant) and may be part of the fast motoneuron system controlling startle responses in adult worms (Wilson 1960; Sigger and Dorsett 1986a; Horridge 1959). The importance of these motoneurons is perhaps reflected in the fact that they were targeted by most of the sensory and interneuron types in the CR circuit.

Proposed role of pygPB^{unp} in the CR circuit

Evidence was shown in Chapter 2 that the biciliated sensory cell pygPB^{unp} also serves as a hydrodynamic receptor. The CR circuit showed that this cell has many targets in common to CRs. These targets include not only the previously reported interaction with the MC cell and with the prototroch (Verasztó et al. 2017), but also the Crab motoneurons and some of the Commissural and Split neurons. Nonetheless, the pygPB^{unp} circuit falls short in explaining all the features of the startle response. Thus, its role in this behavior may be additive, or even antagonistic to that of the CRs.

Limitations, unknowns and outlook circuit

The circuit provides a rich source of hypotheses that could be tested using a range of approaches. The circuit diagram may be particularly informative in the case of the startle response, as no neuropeptide modulators are likely to act in the short time required to initiate the response. Despite this, the full understanding of the response has to await the acquisition of functional data on the components identified as part of the CR circuit.

No sign known for most of the interactions

One of the main caveats of the circuit presented here is the lack of information on the neurotransmitter identity from almost all the cells, from the CRs to the motoneurons. Some

immunostaining reports against monoamines have failed to show expression in CRs, as well as *in situ* hybridization of cholinergic markers (Starunov et al. 2017). However, a targeted analysis is required to determine the neurotransmitter expression in these cells. Motoneurons such as the MN^{crab} make extensive contacts with the longitudinal muscles, which are positive for acetylcholinesterase; and thus these neurons are probably cholinergic (Denes et al. 2007). If this is confirmed it could mean that the interaction between the serotonergic $pygPB^{unp}$ and Crab motoneurons is inhibitory, as suggested for its interaction with the cholinergic ciliomotor neuron MC (Verasztó et al. 2017). The cholinergic cell marker (Vesicular Acetyl choline Transporter (VAchT) is actually expressed all along the VNC at the nectochaete stage (Simionato et al. 2008), and thus the other motoneuron types may be cholinergic as well. It is currently not possible to infer which neurotransmitter is expressed in the interneurons participating in the circuit.

Future directions

The future interrogations to the circuit could follow a similar path to that marking the history of the research on startle circuits. The identified neuron concept provided an experimental framework to understand the more complex escape responses in fish or in crayfish. Thus, like the M-cell or the LG/MG cells were considered to be the core of escape responses, the Rope interneurons in the CR circuit could be set at the center in the effort to understand how the larva makes the startle response. A useful first step in this direction would be to find a good molecular marker to label this cell. This would open calcium imaging, ablation and genetic manipulation of neuronal activity. Crab motoneurons may be also an important cell to identify and analyze what effect their absence or their activation has on the animal. However, a finer-grained description of the muscle types and subtypes activated during the response will be required to understand the different functions of the MN^{crab} and the remaining motoneurons. Faster recordings methods such as spinning disk confocal microscopy or light sheet microscopy will be needed to determine the order of muscle contraction.

Conclusions

The molecular, neuronal and ecological basis of the startle response in the *Platynereis* nectochaete larva

In this study, the startle response in the planktonic larva of the marine annelid *Platynereis* was described and quantified. The startle response was found to be elicited by both touch and hydrodynamic stimuli. Although simple at first glance, the response is a fast locomotor behavior that involves the synchronous and bilateral activation of all the ciliary bands and of the musculature across all the segments of the animal. The startle response was also found to be differentially modulated according to the stimulus direction.

Penetrating uniciliated cells with a collar of microvilli were found by calcium imaging to detect the hydrodynamic stimuli eliciting the response. These cells—now called CRs—express the Polycystin-2 ortholog PKD2-1, and the Polycystin-1 homolog, PKD1-1. The genome editing technique CRISPR was utilized to generate frame-shift mutations in both *PKD2-1* and *PKD1-1*. Homozygote and trans-heterozygote mutants in either of the two genes were not able to display the startle response upon mechanical stimuli. This defect was interpreted as a consequence of CRs not functioning properly to detect the mechanical cues. The drastic effect of mutating *PKD1-1* and *PKD2-1* suggests that these molecules play a crucial role in the sensory function of CR cells.

The *PKD2-1* mutants were used to address the ecological relevance of the startle response. Comparison of survival rates between mutants and wildtype larva when co-incubated with a rheotactic predator revealed that the lack of a startle response in the mutants significantly compromises survival. Thus, the startle response is a behavior that evolved to escape or to defend against planktonic predators.

The identification of the sensory receptors triggering the startle response by physiological, genetic and anatomical approaches, allowed to use connectomics to reconstruct a potential neuronal circuit driving the initiation of the startle response. The synapse-level circuit revealed dedicated neuronal pathways to separately drive the coordinated activation of ciliary bands and muscles. CRs were found to target motoneurons directly or indirectly through a layer of novel types of interneurons. These parallel pathways eventually converge on muscle and cilia and are thought to get activated according to site of stimulation or the nature of the mechanical stimulation. Additional motifs in the CR circuit suggest mechanisms for achieving intersegmental and bilateral synchronization of cilia and muscle activation.

A neuronal model of a predator avoidance and defence behavior

Based on these findings a set of models is proposed to explain the ecological relevance of different behavioral features of the startle response, and how these are implemented at the neuronal level (**Figure 0-1**).

The first conceivable scenario occurs when a swimming *Platynereis* larva detects a predator—in this case a rheotactic predatory copepod—immediately ahead of it (**Figure 0-1**A, left schematic). In this scenario, the larva detects the predator first by sensing the water disturbances it generates (blue waves). This hydrodynamic stimulus would be very weak, which could either mean the copepod is barely moving and not actively attacking the larva, or it is rather far from the larva. The weak stimulus would cause the larva to immediately stop swimming and sink. The larva would not elevate the parapodia as this may betray its location. At the neuronal level, the weak stimulus would cause the activation of the PKD1-1/PKD2-1-dependent mechanotransduction cascade in only a few of the CR neurons in the episphere (**Figure 0-1**A, right schematic). This weak signal from the CRs will lead to the activation of only neurons with low threshold, which in this model include the MC ciliomotor neuron, and the CM interneuron. MC activation will cause the immediate ciliary arrest in the prototroch, while CMs will indirectly cause ciliary arrest in the bodytrochs by activating the Loop and the Ant ciliomotor neurons. A weak signal may also directly activate MN^{crab}, leading to longitudinal muscle contraction.

A similar scenario may occur when the copepod starts to approach the larva from the front (**Figure 0-1**B, left schematic). The response in this case would also involve the immediate pause in swimming, but the approach may also cause the elevation of the parapodia at a low angle (the LowE-type response). The neuronal implementation under this scenario would be essentially the same as before, perhaps only slightly differing in the number of CRs cells activated. The pathway explaining the low-angle elevation of the parapodia could be indirect and currently unknown, as the response has a very long latency (**Figure 0-1**B, right schematic).

A third scenario would take place when the rheotactic predator has detected the larva and it is rapidly approaching from the anterior (**Figure 0-1**C, left schematic). Upon detecting the abrupt and strong hydrodynamic stimulus, the larva would simultaneously stop swimming, abduct all the parapodia and extend all the spiny chaetae. Many CRs in the episphere and segments would be activated by such stimulation. This will generate a strong signal that will activate most of the pathways in the CR circuit (**Figure 0-1**C, right schematic). On one side, all the ciliary control pathways will be activated including the MC, CMs, and due to the stronger signal, the Rope interneurons. The activation of this last neuron will also cause the concomitant activation of VNC motoneurons in all segments on both body sides. In parallel, and if the attack is imminent, a perhaps synergistic activation of Split interneurons across all segments will reinforce the activation of VNC motoneurons ensuring a robust parapodial elevation response.

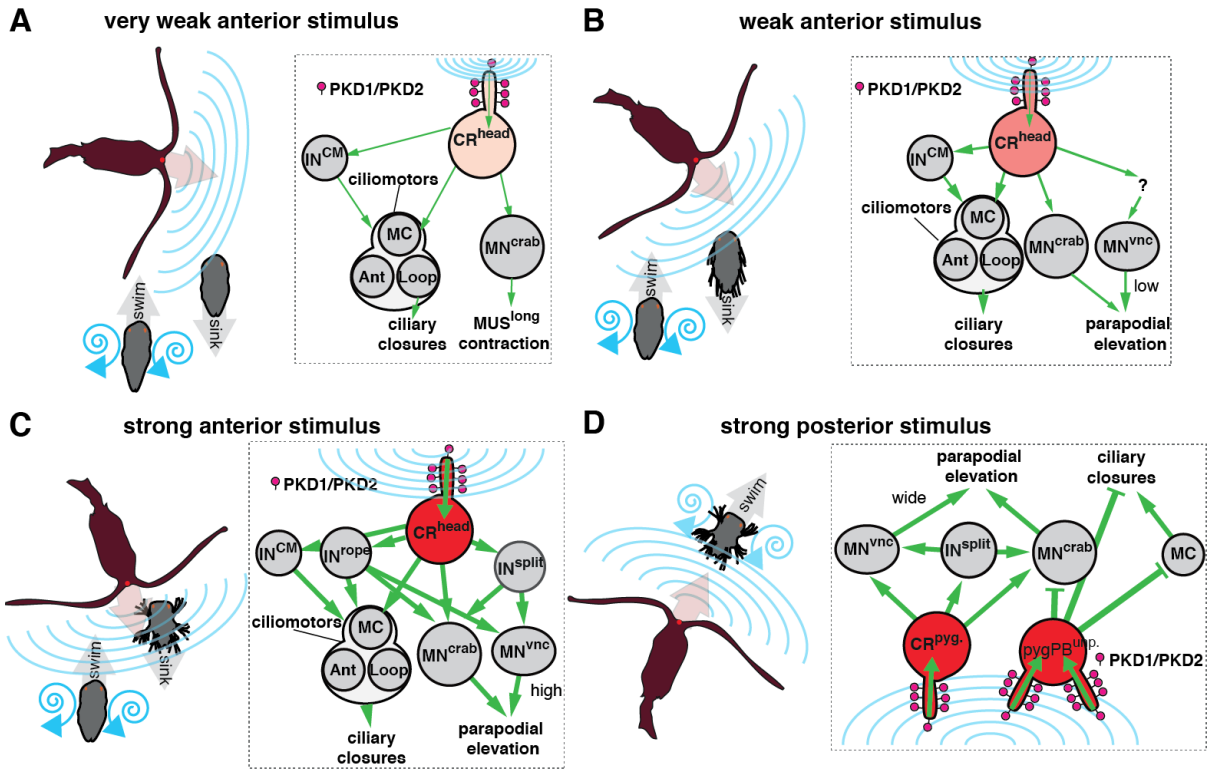


Figure 0-1 Model of the activation of the CR circuit in different predator encounter scenarios eliciting the startle response. (**A**-**D**) Four different scenarios (left schematic in each panel) that differ on the direction of approach of a predator and the distance from the larva. A defined group of neurons and molecules in the CR circuit (right schematic in each panel) is hypothesized to get activated in each scenario. Water displacement is shown in left schematics as blue arcs or spirals. Green arrows in right schematics indicate the direction of the activation of the circuit. Green bars indicate inhibitory interactions. Line width is a subjective measure of the strength of the signal. CRs and pygPB^{unp} are colored in different tones of red to indicate strength of activation. See main text for explanation of the models.

The fourth and final scenario involves the fast approach of a predator from the posterior side (**Figure 0-1**D, left schematic). Upon detecting strong water disturbances from the posterior, the larva may display the full elevation response. Contrary to a frontal attack, the larva would not stop swimming, but instead keep swim away from the threat, perhaps even at a faster pace. Strong posterior stimulation would lead to the activation of CR^{pygidium} neurons and pygPB^{unp} (**Figure 0-1**D, right schematic). The PKD1/PKD2 complex may be working in the transduction of the mechanical signal in both cell types. While a strong signal from pygidial CRs would lead to the direct activation of Crab motoneurons, and to the population of split interneurons, the pygPB^{unp} may actually rise the threshold of activation of both the MC and the Crab motoneurons by direct inhibition. This would ensure that the first action of the larva upon posterior stimuli is to swim faster away from the threat, and only display the full elevation response upon an imminent attack. In conclusion, these four models of the implementation of the CR circuit in different ecologically relevant scenarios provide the first set of testable hypotheses that will be refined as more functional data are accumulated. Comparative analyses to other species displaying a similar response can be a productive avenue to distinguish motifs in the circuit that are product of adaptive evolution from those that are due to the evolutionary history of the species (Dumont and Robertson 1986). The

degree of plasticity and conservation of certain behavioral features can also be gauged from a comparative approach, as performed for startle responses in other organisms (Hale et al. 2002)

Appendix

Materials and Methods

General Materials

Reagents

- PTW: PBS (Max Planck institute in house recipe) + 0.1% Tween
- LB medium (made by the in-house culture facility)
- Natural sea water from the North Sea (Bremerhaven, Germany).
- Ampicillin 100 µg/ml
- Tween 20 (P2287, Sigma)
- Distilled Water, nuclease-free (10977-035,Gibco)
- $MgCl_2$ (CAS 7791-18-6, Sigma)

Equipment

- Benchtop centrifuge 5424 (Eppendorf)
- Benchtop centrifuge 5415 R (Eppendorf)
- Thermomixer compact (Eppendorf)
- Gel electrophoresis chamber Peqlab and Jena Analytik
- Stereomicroscope source lamp CL1500 ECO (Carl Zeiss GmbH)
- Stereomicroscope Stemi 2000 (Carl Zeiss GmbH)

Platynereis dumerilii culture

P. dumerilii worms from the Tübingen strain and the CRISPR mutants generated in this study were cultured as previously described (Hauenschild and Fischer 1969), but the artificial aeration method was omitted. Animals were cultured at 18-22°C in a 1:1 or 1:2 mixture natural sea water (NSW, North Sea): artificial sea water (ASW, Tropic-Marin). All the water was passed first through a system of filters and sterilized with UV light (Grünbeck GmbH). Animals were fed once per week with a mixture of *Tetraselmis* algae and fish flakes, and if older than 2 months, additionally fed once per week with spinach and once with *Artemia* nauplii. Water in the culture was generally exchanged every second week.

The fertilization protocol followed was described previously (Dorresteijn 1990), but the antibiotic treatment was not necessary due to the filtering system used in the facility. All batches were made in NSW and kept at 18°C in a 16h-8h light/dark cycle.

Sanger sequencing

All sequences in this study were determined by Sanger sequencing using labeled ddNTPs (BigDye® Terminator v3.1 Cycle Sequencing Kit, ThermoFisher Scientific) and obtained by the in-house sequencing facility (run by Detlef Weigel, Christa Lanz and Andrea Belkacemi).

Gel electrophoresis

Unless otherwise noted, agarose gel electrophoresis was performed in 1% molecular biology-grade agarose (Carl Roth GmbH) gels using 0.5X TBE as running buffer (ThermoFisher Scientific). TBE buffer was reused except for recovery of gel bands or for running RNA samples. DNA samples were run at max. 140V and RNA samples at max.120V using an electrophoresis power supply (BioRad).

Polymerase Chain Reaction (PCR)

Unless otherwise noted, PCR was performed using the following components and concentrations:

PCR components	Final concentration
5x HF green buffer (Thermo Fisher Scientific)	1x
Forward primer 5µM	0.5 µM
Reverse primer 5µM	0.5 µM
dNTPs 10mM (Thermo Fisher Scientific)	2mM
DNA template (e.g.gDNA,Plasmid)	1/10th of final vol.
Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific or New England Biolabs)* 2U/ul	0.02U/µl

*Either Polymerase was used interchangably with Thermo HF buffer with no appreciable differences.

Microinjection of *Platynereis* zygotes

The microinjection method followed was originally described in (Ackermann 2003).

Injection setup

After fertilization, eggs were placed at 14-15°C using a temperature controller (BadController V, Luigs & Neumann) for 50 minutes. Eggs were de-jellied afterwards and microinjected for ca. 1h at 30 min using pre-pulled injection capillaries with 0.5 µm inner diameter (Femtotips II, Eppendorf) driven by a Femtojet pump (Eppendorf) at variable pressure. Injection needle was controlled with a motorized 3-axis micromanipulator (Unit MRE/MLE Mini 25, Luigs & Neumann). Mini/Midiprep-purified plasmid constructs were precipitated with ethanol overnight prior to preparing the injection aliquot and resuspended in nuclease-free water (Gibco®, ThermoFisher Scientific). This solution was centrifuged at maximum speed for >30min at 4°C, and from it the injection aliquot at the desired concentration was prepared. mRNA and sgRNA solutions were injected directly from the aliquots prepared after *in vitro* transcription. All injection solutions were centrifuged for > 30 min before needle loading, in order to minimize clogging due to larger aggregates.

in vitro mRNA synthesis from PdumIVT plasmid

To synthesize mRNA from plasmids derived from the *Pdum*IVT parent plasmid (described in (Randel et al. 2014)), a PCR fragment was amplified (see primers in **Table 0-4**), using the following program:

Temperature	Time	Repeat
98°C	3 min	1 time
98°C 72°C	10 sec x*15sec	35 times
72°C 10°C	10 min ∞	1 time

x=predicted #Kb ORF

The sequence of the purified product was confirmed using the amplification primers as sequencing primers. 200 ng of the purified fragment was used for *in vitro* transcription using the mMESSAGE mMACHINE T7 ULTRA kit (Ambion). The resulting capped and polyadenylated mRNA was purified with the RNeasy mini kit (Qiagen), eluted in 25 nuclease-free water (Gibco®, ThermoFisher Scientific), aliquoted and stored at -80°C until use. One of the aliquots was diluted 10 to 50 times and the resulting concentration was measured with a spectrophotometer (NanoDrop, ThermoFisher Scientific). The integrity of the mRNA was assessed by agarose gel electrophoresis using denaturing conditions. In brief, loading dye containing formamide (R0641, ThermoFisher Scientific) was added to the mRNA aliquot in a 1:1 ratio and then denatured in a thermocycler at 70°C for 10 min. The aliquot was put in ice-cold water until loading on the gel.

Oligo annealing

Lyophilized oligos were dissolved in annealing buffer (10mM Tris, pH7.5-8.0, 50mM NaCl, and 1mM EDTA) to 100 µM. 1 µl of each oligo was added to 48 µl annealing buffer and incubated in the Thermocycler following this protocol:

Temperature	Time	Repeat
95°C	5 min	1 time
95°C	1 min	-1°C per cycle (repeat 95-X times)
X°C	30 min	1 time
(X-1)°C	1 min	-1°C per cycle (repeat X-26 times)
Hold at cold T	∞	

5 µl of the annealing reaction was mixed with 45 µl ligation-grade water. The concentration of this solution was measured with a spectrophotometer (NanoDrop, ThermoFisher Scientific) to calculate the amount needed to add to the ligation reaction (usually a molar ratio of > 10 of annealed oligo to vector was used).

DNA Ligation and transformation

Ligation reactions were always performed with T4 DNA ligase (EL0014, ThermoFisher Scientific) for 2hr at RT, or for 24hr at 4°C. The following formula was used to calculate the amount of insert to add to have a defined insert: vector molar ratio:

$$MassI = \left(\frac{LengthI}{LengthI}\right) * MassV$$

Where $MassI$ and $MassV$ are the mass in ng of the insert and vector, respectively, and $LengthI$ and $LengthV$ are the length in base pairs of the insert and vector, respectively. Ratios from 1 to 6 were used to clone promoters, while annealed oligos usually required up to 10:1 insert to vector ratio. 100 ng plasmid were used for all ligations in a final volume of 20 µl. At the end of the incubation period the ligation reaction was placed on ice and 10 µl added to electrocompetent cells (TOP10 cells, in-house stock) thawed on ice. After electric shock, cells were incubated in S.O.C. medium (15544034, ThermoFisher Scientific) for 45min at 37°C, and plated overnight in LB plates with antibiotic at 37°C.

Figure assembly

All figures were assembled with Adobe Illustrator CC (Adobe Inc.). Adobe Photoshop CS5 (Adobe Inc.) was used to color SEM micrographs presented in Chapter 2 and Chapter 3. Scale bars were added in Illustrator based on calibrated images taken in FiJi.

Primer list

Table 0-1 Primers used for amplifying gene fragments for probe synthesis

Gene name		Primer sequence (5'-3')
ASIC8016	ISH1_Fwd	TCACTCTTTGCAATTACAACCAGT
	ISH1_Rev	CGGAGAGGATTTTCAAAAGAAGGA
ENaC417306	ISH1_Fwd	GACAGGATCGACAGGACTTAATCT
	ISH1_Rev	GACCATCTTCCTCAGACAAGAGTT
	ISH2_Fwd	AAGTGTGTTCCTGTGTGGAGACT
	ISH2_Rev	GGAAGAATATGTTCACTCGCAATA
ENaC2547	ISH1_Fwd	GAAGTCTATGGGGAATACCACAAC
	ISH1_Rev	AACTCGATCATTGTGAATAACGAA
ENaC415688	ISH1_Fwd	TGCTGATCTTTGGTGCAGAC
	ISH1_Rev	CTTACGGGAGCAGCTTTGAC
	ISH2_Fwd	TGGAGCTACTAATGAACTCCAAAA
	ISH2_Rev	TATCTGAATGGTGAACTCAAATGC
PKD1-1	ISH1_Fwd	TGTCTTTGTAACTGTTGGTCTGGT
	ISH1_Rev	ATGTTCCTCAGAGATTCCTTCATC
PKD1-2/LOV-1	ISH1_Fwd	AGGTTTTATAAATGACCGGAACAA
	ISH1_Rev	TAGCACAGTTACGGAGAGTAGACG
	ISH4_Fwd	GGCTTTATTCGACAACAAGAAAAT
	ISH4_Rev	AGCCATCTGTTGCATATGAAGTAA
PKD2-1	ISH1_Fwd	TCCTCGTCATCATAATATCCATTG
	ISH1_Rev	CTCTCTTTGTTGAGTTGGTCCTTT
PKD2-2	ISH1_Fwd	GACACTTACTCAGCTGTCAAGGAA
	ISH1_Rev	CTCGGATCCGGTATACATCTTTAG
NOMPC	ISH1_Fwd	TTACAACTTGATGGAGGACAGAAA
	ISH1_Rev	GCCAGCCATGAAGATAACTAAGAT
	ISH2_Fwd	ATGATGTCCGACACTTATCAAAGA
	ISH2_Rev	TGCACATATTAAATTCATTGCTGTT
	ISH3_Fwd	GCACAAAAGAAGTCAAAAGAGACA
	ISH3_Rev	TATGTTTCAGGAAAAGTTTGACCA
	ISH4_Rev	TCTTTGATAAGTGTCGGACATCAT
Piezo	ISH1_Fwd	GGCGAGTGTTCTTCCTTGAG
	ISH1_Rev	TCTCGGCCAGTCTTTCTTGT
	ISH2_Fwd	TTCTGCGTCTCCAAACTGTG
	ISH2_Rev	CCGTACTCCTCCACGTCATT
	ISH3_Fwd	TGATGGACTGGATCTGGACA
	ISH4_Fwd	CCTGCCCAGTTGTGGTACTT

	ISH3-4_Rev	TGGTGAGGGCTAACAATTCC
	ISH5_Fwd	ATCAATTCCAGTTTTTCCAAGAAG
	ISH5_Rev	TCCTTGAAGTTGTCAAACTGGTAA
	ISH6_Fwd	GAAGACAACACAGTTTCTCAGACC
	ISH6_Rev	CGCATAGTAGTCCTTCCGTTCTAT

Table 0-2 Oligos used for insertion of diverse protein tags

Primer name		Primer sequence (5'-3')
3xHA tag oligo adaptors	Forward	CGCGCCATTACCCATACGATGTTCCAGATTACGCTGGCTATCCCTATGACGTCCCG GACTATGCAGGTTCCTATCCATATGACGTTCCAGATTACGCTTTGCGAT
	Reverse	CGCAAAGCGTAATCTGGAACGTCATATGGATAGGAACCTGCATAGTCCGGGACG TCATAGGGATAGCCAGCGTAATCTGGAACATCGTATGGGTAATGG
Palmitoylation tag oligo adaptors	Forward	GATCCGGCGCGCCATCTGTGCTGTATGAGAAGAACCAAACAGGTTGAAAAGAAT GATGAGGACCAAAAGATTCCGCGGA
	Reverse	CGCGTCCGCGGAATCTTTTGGTCCTCATCATTCTTTTCAACCTGTTTGGTTCTTCT CATACAGCACAGATGGCGCGCCG
P2A oligo adaptors	Forward	CCGGTGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTG GAGGAGAACCCTGGACCTATGGC
	Reverse	GGCCGCCATAGGTCCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGGCTG AAGTTAGTAGCTCCGCTTCCA

Table 0-3 Primers for promoter amplification

Gene name		Primer sequence (5'-3')
PKD1-1	Forward	ATGGATCCGGAAGCCTGATAACAAAGTGAGTAGGAA
	Reverse	AAGGCGCGCCGCTCCTCCTCCAAGTGGCTCTACCATCTCCGTCTGTATC
PKD2-1	Forward	ATGGATCCAATAAAATTTGAACCGAGGCCAATGGA
	Reverse	AAGGCGCGCCGCTCCTCCTCCGGGGCGTGACATCAGCCAGTAGTTGGT
NOMPC	Forward	ATGGATCCAATCCGAAAGGCTCCGCCCAATATT
	Reverse	AAGGCGCGCCGCTCCTCCTCCTCCCTCCCTCACAGCCAGCAT

Table 0-4 Primers for template generation for *in vitro* transcription (IVT) using *PdumIVT* plasmid

Primer name		Primer sequence (5'-3')
IVT primers for	Forward	GATCCCCCTCGGATCCTAATACGACTCACTATAGGGAGATTTGATGTTTAC AGGGC

PdumIVT plasmid	Reverse	gaattcgagctcggtacctcgcgaatgcatctagatCCAATTTTCTCTTAAACAACTCC

Table 0-5 Primers for CRISPR (i.e. genotyping, gRNA cloning)

Primer name		Primer sequence (5°-3')
Cas9-GFP ORF	Forward	ATGGCGCGCCTAGACAAGAAGTACAGCATCGGCCTGG
	Reverse	TTAACCGGTCTTGTACAGCTCGTCCATGCCGAG
PKD2-1_truT2 gRNA	Forward	TAGGGCCACCCGTCAGACTG
	Reverse	AAACCAGTCTGACGGGTGGC
PKD1-1_truT1 gRNA	Forward	CACCTGTGTCAACTTTACCCCA
	Reverse	AAACTGGGGTAAAGTTGACACA
PKD1-1_T5 gRNA	Forward	CACCAGGCGGATATCAGTGAAGCG
	Reverse	AAACCGCTTCACTGATATCCGCCT
NOMPC_T3 gRNA	Forward	CACCTCTCCCTCACAGCCAGCATC
	Reverse	AAACGATGCTGGCTGTGAGGGAGA
PKD1-1_truT1 IVT	Forward	ttaatacgactcactatagTGTGTCAACTTTACCCCA
PKD1-1_T5 IVT	Forward	ttaatacgactcactatagAGGCGGATATCAGTGAAGCG
NOMPC_T5 IVT	Forward	ttaatacgactcactatagTCTCCCTCACAGCCAGCATC
PKD1-1 PCR genotyping	Forward	TCAAACTGGTCAAGATTAAATTCCAGA
	Reverse	TCTATTTCACTAATGTTGTTCCTGATG
PKD2-1 PCRgenotyping	Forward	CCCTTTGTGAGCAGGAGATGCCCTGC
	Reverse	CATGACCTGAGTGTAGTAGTACATGGT
NOMPC PCR genotyping	Forward	ACCAACCAATAAAGTAAACTATCAACCTAT
	Reverse	ACTTGACAATGTCGTCTCTTGAGTAG
PKD1-1_T5 PCR sequencing	Forward	TATGAGACTGAATGCACAATAGAGTTT
PKD1-1_truT1PCR sequencing	Reverse	TAAGTGAAGGTCACATACTCGTCAGT
PKD2-1 PCR sequencing	Reverse_1	AGGATGACCAAGAAGACCAAGTAG
	Reverse_2	AACCTTCAAATATGTTCACTACAATCC
NOMPC PCR sequencing	Reverse	TGATTAGGTTTGACAAAAAGATAA

Table 0-6 Penetrating ciliated cells found in the electron microscopy volume of a nectochaete larva

		Neuron name	# Cilia	Microvilli Collar	# Microvilli	Fully penetrating (TEM) (-/+ if reaching cuticle, but not crossing)	Microvilli cross cuticle
Prostome	**Episphere**	MS1	1	+	13	+	-
		MS2	1	+	13	+	-
		hCR1r	1	+	10	+	+
		hCR1l	1	+	10	+	+
		hCR2r	1	+	10	+	+
		hCR2l	1	+	ND	+	ND
		hPUc1r	1	+	7	+	-

		hPUc1l	1	+	8	+	-
		hPUc2r	1	+	7	+	-
		hPUc2l	1	+	7	+	-
		hPUc3r	1	+	8	+	-
		hPUc3l	1	+	9? (+1 partially shared with PUc2l)	+	-
		hCR3r_MC1_r-triad	1	+	10	+	+
		hCR3l_MC	1	+	10	+	+
		hPU1r	1	-	NA	+	NA
		hPU1l	1	-	NA	+	-
		hPBc1r_triad	2	+/-	C1: ca.7,C2:?	+	-
		hPBc1l_triad	2	+	C1:9,C2:6	+	-/+
		hPU2l	1	-	NA	+	NA
	(developing) Antennae	antCR1r	1	+	10	+	+
		antCR1l	1	+	10	+	+
		antCR2r	1	+	10	-	-
		antCR2l	1	+	10	-	-
		antCR3r	1	+	10	+/-	+/-
		antCR3l	1	+	10	-	-
		antCR4r	1	+	10	+/-	+/-
		antCR4l	1	+	9-10	-	-
		antCR5r	1	+	10	+/-	-
		antCR5l	1	+	10	+/-	-
		antPUc1r	1	+	10	+	+/-
		antPUc1l	1	+	9	+	+/-
		antPUc2r	1	+	8	+/-	+/-
		antPUc2l	1	+	9	-	-
		antPUc3r	1	+	7-8	-	-
		antPUc3l	1	+	8	+/-	+/-
		antPUc4r	1	+	7-8	-	-
		antPUc4l	1	+	9	-	-
		antPUc5r	1	+	7	-	-
		antPUc5l	1	+	7	-	-
		antPUc6r	1	+	8	-	-
		antPUc6l	1	+	8	-/+	-
		antPUc7r	1	-/+	7(branched)	-	-
		antPUc7l	1	+	8	-	-
		antPUc8r	1	+	7	-	-
		antPUc8l	1	+	8	-	-
		antPUc9r	1	+/-	7	+/-	-
	Do	MS3l	1	+	12	+	-

		MS3r	1	ND	ND	+	ND
		MS4	1	+	12	+	-
		MS5r	1	+	13 (one double)	+	-
		MS5l	1	+	9	+	-
		doCRunp	1	+	10	+	+
	Dorsal sensory organ	dsoCR1r	1	+	10?	+	+
		dsoCR1l	1	+	10?	+	+
		dsoCR2r	1	+	10	+	+
		dsoCR2l	1	+	10	+	+
		dsoCR3r	1	+	10	+	+/-
		dsoCR3l	1	+	10	+/-	+/-
		dsoCR4r	1	+	10	+/-	+/-
		dsoCR4l	1	+	10	+/-	+/-
		dsoPU2unp	1	?	?	+/-	-
		dsoPUc1unp	1	+	9	+	+
		dsoPU1unp	1	-	-	+	NA
		dsoPU3unp	1	-	NA	+	NA
Segmented trunk	**(developing) Head cirrus**	hCirrusCR1r	1	+	10	+/-	+/-
		hCirrusCR1l	1	+	10	+/-	+/-
		hCirrusCR2r	1	+	10	-	+/-
		hCirrusCR2l	1	+	10	+/-	+/-
		hCirrusPUc1r	1	+	8 or 9	-	-
		hCirrusPUc1l	1	+	7?	-	-
		hCirrusPUc2r	1	+	8 to 11?	.	-
		hCirrusPU1l	1	-	NA	-	NA
		hCirrusPU2l	1	-	NA	-	NA
	Ventral side	ventraltrunkPU1_sg1r	1	-	NA	+	NA
		ventraltrunkPU1_sg1l	1	-	NA	+/-	NA
		ventraltrunkPB1_sg1r	2	-	NA	c1:+,c2:+	NA
		ventraltrunkPB1_sg1l	2	c1:+,c2:-	c1:6	c1:+,c2:+	-
		ventraltrunkPB1_sg2r	2	-	NA	c1:+,c2:+	NA
		ventraltrunkPB1_sg2l	2	-	NA	c1:+,c2:+	NA
		ventraltrunkPB1_sg3r	2	-	NA	c1:+,c2:+	NA
		ventraltrunkPB1_sg3l	2	-	NA	c1:+,c2:+	NA
		ventralParaPU1_sg1r	1	-	NA	+	NA
		ventralParaPU1_sg1l	1	-	NA	+	NA
		ventralParaPU2_sg1r	1	-	-	+	-
		ventralParaPU2_sg1l	1	-	NA	+	NA
		ventralParaPB1_sg1r	2	-	NA	c1:+/-,c2:+	NA
		ventralParaPB1_sg1l	2	-	NA	c1:-,c2:-	NA
		ventralParaPB2_sg1r	2	-	NA	c1:+,c2:+	NA

		ventralParaPB2_sg1l	2	c1:+,c2:-	c1:11	c1:-,c2:+/-	-
		ventralParaPU1_sg2r	1	-	NA	+	NA
		ventralParaPU1_sg2l	1	-	NA	+	NA
		ventralParaPU2_sg2r	1	-	NA	+	NA
		ventralParaPU2_sg2l	1	-	NA	+	NA
		ventralParaPU3_sg2r	1	-	NA	-	NA
		ventralParaPU3_sg2l	1	-	NA	-	NA
		ventralParaPU4_sg2r; SN	1	-	NA	-	NA
		ventralparaPU4_sg2l; SN	1	+	9 or 10	-	-
		ventralParaPB1_sg2r	2	-	NA	c1:+,c2:+	NA
		ventralParaPB1_sg2l	2	-	NA	c1:+,c2:+	NA
		ventralParaPB2_sg2r	2	-	NA	c1:-,c2:-	NA
		ventralParaPB2_sg2l	2	-	NA	c1:-,c2:-	NA
		ventralTeloPUc1r	1	+/-	4 or 6	+	-
		ventralTeloPUc1l	1	+	10	+	+/-
		ventralTeloPUc2r	1	+/-	5 or 6	+	-
		ventralTeloPUc2l	1	+/-	9	+	+/-
		ventralTeloPB1r	2	-	NA	c1:+,c2:+	NA
		ventralTeloPB1l	2	-	NA	c1:-,c2:-	NA
		ventralTeloPB2r	2	-	NA	c1:-,c2:-	NA
		ventralTeloPB2l	2	-	NA	c1:-,c2:-	NA
		ventralpostParaPUc1_sg1r	1	+	11	+	-
		ventralpostParaPUc1_sg1l; SN	1	+	10	+	-
		ventralpostParaPUc2_sg1r; SN	1	+	9	+	-
		ventralpostParaPUc2_sg1l; SN	1	+	11	-	-
		ventralpostParaPUc1_sg2r; SN	1	+	9	+/-	-
		ventralpostParaPUc2_sg2r; SN	1	+	8	+	-
		ventralpostParaPUc1_sg2l; SN	1	+	9	+	-
		ventralpostParaPUc2_sg2l; SN	1	+	9	+	-
		ventralCirrusPB1_sg2r	2	-	NA	c1:+,c2:+	NA
		ventralCirrusPB1_sg2l	2	-	NA	c1:+,c2:+	NA
		ventralCirrusPB1_sg3r	2	-	NA	c1:+,c2:+	NA
		ventralCirrusPBc1_sg3l	2	-	c1:8,c2:6-9	c1:-/+,c2:+/-	-
	Dorsal side	dorsalParaPUc1_sg1r	1	+/-	8 or 9	+/-	-
		dorsalParaPUc1_sg1l	1	+/-	8 or 9	+/-	-
		dorsalParaPU1_sg1r	1	-	NA	-	NA
		dorsalParaPU1_sg1l	1	-	NA	-	NA
		dorsalParaPB1_sg1r	2	-	NA	c1:+/-,c2:+/-	NA
		dorsalParaPB1_sg1l	2	+/-	c1:5;c2:6	c1:+/-,c2:+/-	-

		dorsalParaPB2_sg1r	2	-	NA	c1:-,c2:-	NA
		dorsalParaPB2_sg1l;	2	-	NA	c1:-,c2:-	NA
		dorsalParaPU1_sg2r	1	-	NA	+	NA
		dorsalParaPU1_sg2l	1	-	NA	+	NA
		dorsalParaPU2_sg2r	1	+	NA	-	NA
		dorsalParaPUc1_sg2l	1	+	7	+/-	-
		dorsalParaPB1_sg2r	2	-	NA	c1:+,c2:+	NA
		dorsalParaPB1_sg2l	2	-	NA	c1:+,c2:+	NA
		dorsalParaPB2_sg2r	2	-	NA	-	NA
		dorsalParaPB2_sg2l	2	-	NA	c1:+/-,c2:+/-	NA
		dorsalTeloPUc1r	1	+/-	5	+	+
		dorsalTeloPUc1l	1	+/-	7 or 8	+/-	-
		dorsalTeloPU2r	1	-	NA	-	NA
		dorsalTeloPU2l	1	-	NA	-	NA
		dorsalTeloPB1r	2	+	c1:10,c2:10	c1:+,c2:+	+
		dorsalTeloPB1l	2	-	NA	c1:+,c2:+	NA
		dorsalTeloPB2r	2	-	NA	-	NA
		dorsalTeloPB2l	2	-	NA	-	NA
	Parapodia	interparapodPUc1_sg1r	1	+/-	7 or 8	+/-	-
		interparapodPUc1_sg1l	1	+/-	4 or 5	+/-	-
		interparapodCR1_sg1r	1	+	10	+	+
		interparapodCR1_sg1l	1	+	10	+	+
		interparapodPU1_sg1r	1	-	NA	-	NA
		interparapodPU1_sg1l	1	-	NA	-	NA
		interparapodPUc1_sg2r	1	+	8	+	+/-
		interparapodPUc1_sg2l	1	+/-	7	+/-	-
		interparapodCR1_sg2r	1	+	10	+	+
		interparapodCR1_sg2l	1	+	10	+	+
		interparapodPU1_sg2r	1	-	NA	-	NA
		interparapodPU1_sg2l	1	-	NA	-/+	NA
		interparapodPUc1_sg3r; SN	1	+	10	+	-
		interparapodCR1_sg3r	1	+	10	+/-	-
		interparapodCR1_sg3l	1	+	10	+/-	-
		interparapodPM1_sg1r	3	-	NA	+	NA
		interparapodPM1_sg1l	3	-	NA	+	NA
		interparapodPM1_sg2r	4	-	NA	+	NA
		interparapodPM1_sg2l	4	-	NA	+	NA
		interparapodPM1_sg3r	1 (three basal bodies)	-	NA	-	NA

		interparapodPM1_sg3l	1(two basal bodies)	-	NA	-	NA
		neuropodPUc1_sg1r; SN	1	+	9	+	+
		neuropodPUc1_sg1l; SN	1	+	9	+	+
		neuropodPUc1_sg2r; SN	1	+	9	+	-
		neuropodPUc1_sg2l; SN	1	+	9	+	+
		neuropodPUc2_sg2r; SN	1	+	9	+	-
		neuropodPUc2_sg2l; SN	1	+	11?	+	-
		neuropodPUc1_sg3r; SN	1	+	8	-	-
		neuropodPUc1_sg3l; SN	1	+	11 to 13	-	-
		neuropodPUc2_sg3r; SN	1	+	11	-	-
		neuropodPUc2_sg3l; SN	1	+	9	-	-
		notopodPUc1_sg1r;SN	1	+	9	+	-
		notopodPUc1_sg1l	1	+	9 to 11	+	-/+
		notopodPUc1_sg2r	1	+	9	+	-
		notopodPUc1_sg2l	1	+	10	+	-
		notopodPUc1_sg3r; SN	1	+	9 or 10	sections missing to determine	-
		notopodPUc1_sg3l	1	+	10	-	-
		notopodPUc2_sg3l	1	+	10	-	-
	Glands	spinPU1_sg2l	1	+	9	+	-
		spinPU2_sg2l	1	+	10	+	-/+
		spinPU3_sg2l	1	-	-	+	-
		spinPB1_sg2l	2	+/-	hard to assess	c1:+,c2:+	-
		spinPB2_sg2l	2	c1:+,c2:-	c1:8	c1:+,c2:+	-
		spinPB3_sg2l	2	c1:+,c2:-	c1:7	c1:+,c2:+	-
		spinPU1_sg3r	1	-	-	-	-
		spinPU1_sg3l	1	-/+	5 or 6	-	-
		spinPU2_sg3r	1	+	10 or 11	+	-/+
		spinPU2_sg3l	1	+	11	+	-/+
		spinPU3_sg3r	1	+	8	+	-/+
		spinPU3_sg3l	1	+	10 or 11	+	-/+
		spinPB1_sg3r	2	-	-	c1:+/-,c2:+/-	-
		spinPB1_sg3l	2	-	-	c1:+,c2:+	-
		spinPB2_sg3r	2	c1:+,c2:+	c1:8,c2:7	c1:+,c2:+	-
		spinPB2_sg3l	2	c1:+c2:+	c1:8c2:7	c1:+,c2:+	-
		spinPB3_sg3r	2	-	-	c1:+,c2:+	-
		spinPB3_sg3l	2	-	-	c1:+,c2:+	-
Pygidium	**Ventral**	pygPBunp	2	+	C1:10 (5 unique, 1shared, 2 split) ,C2:10(7	+	+

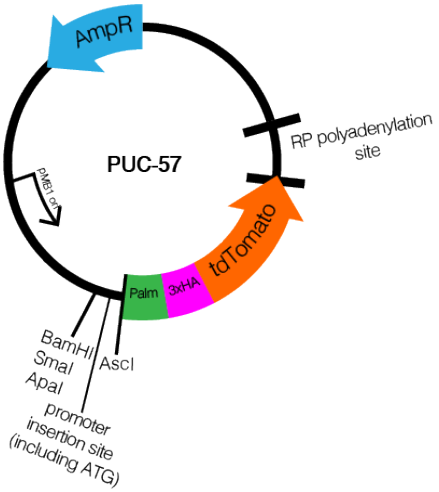
					unique, 1 shared,2 split but only one of each forms part of collar)		
		ventralpygCR1r	1	+	10	+	+
		ventralpygCR1l	1	+	10	+/-	+
		ventralpygCR2r	1	+	10	+	+
		ventralpygCR2l	1	+	10	+	+
		ventralpygCR3r	1	+	10	+	+
		ventralpygCR3l	1	+	10	+	+
		ventralpygPU1unp	1	-	NA	+	NA
		ventralpygCRunp	1	+	10	+	+
		ventralpygPB1r	2	-	NA	c1:+,c2:+/-	NA
		ventralpygPB1l	2	-	NA	c1:+,c2:+/-	NA
	(developing) pygidial Cirri	pygCirrusMSPUc1r	1	+	12	+	-
		pygCirrusMSPUc2r	1	+	10 or 11	+	-
		pygCirrusMSPUc3r	1	+	12	+	+/-
		pygCirrusPU1r	1	+/-	NA	-	NA
		pygCirrusPU2r	1	-	NA	+	NA
		pygCirrusPU3r	1	-	NA	+	NA
		pygCirrusPU4r	1	ND	?	+	?
		pygCirrusPU5r	1	ND	?	+	?
		pygCirrusPU6r	1	-	NA	+/-	NA
		pygCirrusPU7r	1	+/-	5?	+/-	-
		pygCirrusPU8r	1	-	NA	+/-	NA
		pygCirrusPU9r	1	-	NA	-	NA
		pygCirrusPU10r	1	+	not possible to determine	+	NA
		pygCirrusCR1r	1	+	10	+	+/-
		pygCirrusCR2r	1	+	10	+	+
		pygCirrusCR3r	1	+	10(hard to count)	+	+
		pygCirrusPBc1r	2	+	c1:10,c2:11	c1:+,c2:+	-

Plasmid maps

Palmi-3xHA-tdTomato_PUC-57

This plasmid was generated by restriction cloning. plasmid database ID (Jékely Lab): 614

Plasmid map



Sequence

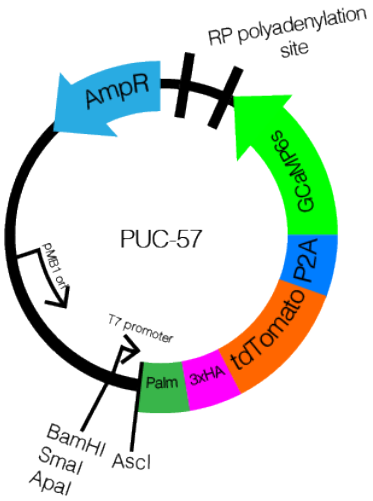
> **Palmi-3xHA-tdTomato_PUC-57**

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGT
CGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACT
GTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGT
GAATTCGAGCTCGGTACCTCGCGAATGCATCTAGATCCAATTTTCTCTTAAACAACTCCAGACCAGTAAGTTACATGTTTTTTTTTTATTGGAAAGATGTTACAGACTGATGTTTGTAAAAGTGTTTCTTTTAACC
GGTCTTGTACAGCTCGTCCATGCCGTACAGGAACAGGTGGTGGCGGCCCTCGGAGCGCTCGTACTGTTCCACGATGGTGTAGTCCTCGTTGTGGGAGGTGATGTCCAGCTTGGTGTCCACGTAGTAGTAGCC
GGGCAGTTGCACGGGCTTCTTGGCCATGTAGATGGTCTTGAACTCCACCAGGTAGTGGCCGCCGTCCTTCAGCTTCAGGGCCTGGTGGATCTCGCCCTTCAGCACGCCGTCGCGGGGGTACAGGCGCTCGGT
GGAGGCCTCCCAGCCCATGGTCTTCTTCTGCATTACGGGGCCGTCGGGGGGGGAAGTTGGTGCCGCGCATCTTCACCTTGTAGATCAGCGTGCCGTCCTGCAGGGAGGAGTCCTGGGTCACGGTCACCAGACC
GCCGTCCTCGAAGTTCATCACGCGCTCCCACTTGAAGCCCTCGGGGAAGGACAGCTTCTTGTAATCGGGGATGTCGGCGGGGTGCTTCACGTACGCCTTGGAGCCGTACATGAACTGGGGGGACAGGATGTC
CCAGGCGAAGGGCAGGGGGCCGCCCTTGGTCACCTTCAGCTTGGCGGTCTGGGTGCCCTCGTAGGGGCGGCCCTCGCCCTCGCCCTCGATCTCGAACTCGTGGCCGTTCATGGAGCCCTCCATGCGCACCTTG
AAGCGCATGAACTCTTTGATGACGGCCATGTTGTTGTCCTCGGAGGAGGCGGTGCCGGAGCTGCCGCTGCCGGTGCTGCCGGTGCCATGCCCCAGGAACAGGTGGTGGCGGCCCTCGGAGCGCTCGTACTGT
TCCACGATGGTGTAGTCCTCGTTGTGGGAGGTGATGTCCAGCTTGGTGTCCACGTAGTAGTAGCCGGGCAGTTGCACGGGCTTCTTGGCCATGTAGATGGTCTTGAACTCCACCAGGTAGTGGCCGCCGTCCT
TCAGCTTCAGGGCCTGGTGGATCTCGCCCTTCAGCACGCCGTCGCGGGGGTACAGGCGCTCGGTGGAGGCCTCCCAGCCCATGGTCTTCTTCTGCATTACGGGGCCGTCGGGGGGGAAGTTGGTGCCGCGCA
TCTTCACCTTGTAGATCAGCGTGCCGTCCTGCAGGGAGGAGTCCTGGGTCACGGTCACCAGACCGCCGTCCTCGAAGTTCATCACGCGCTCCCACTTGAAGCCCTCGGGGAAGGACAGCTTCTTGTAATCGGG
GATGTCGGCGGGGTGCTTCACGTACGCCTTGGAGCCGTACATGAACTGGGGGGACAGGATGTCCCAGGCGAAGGGCAGGGGGCCGCCCTTGGTCACCTTCAGCTTGGCGGTCTGGGTGCCCTCGTAGGGGC
GGCCCTCGCCCTCGCCCTCGATCTCGAACTCGTGGCCGTTCATGGAGCCCTCCATGCGCACCTTGAAGCGCATGAACTCTTTGATGACCTCCTCGCCCTTGCTCACTAGGCGCGATCGCAAAGCGTAATCTGGAA
CGTCATATGGATAGGAACCTGCATAGTCCGGGACGTCATAGGGATAGCCAGCGTAATCTGGAACATCGTATGGGTAATGGCGCGTCCGCGGAATCTTTTGGTCCTCATCATTCTTTTCAACCTGTTTGGTTCTT
CTCATACAGCACAGATGGCGCGCCGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAC
GAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA
GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGT
CAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC
AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACG
GGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCA
TAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCC
ACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT
TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG
TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCG
TATCACGAGGCCCTTTCGTC

Palmi-3xHA-tdTomato-P2A-GCaMP6s_PUC-57

This plasmid was generated by restriction cloning. plasmid database ID (Jékely Lab): 657

Plasmid map



Sequence

> T7::Palmi-3xHA-tdTomato-P2A-GCaMP6s_PUC-57

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGT
CGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACT
GTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGT
GAATTCGAGCTCCCAATTTTCTCTTAAACAACTCCAGACCAGTAAGTTACATGTTTTTTTTTTATTGGAAAGATGTTACAGACTGATGTTTGTAAAAGTGTTTCTTTTAATTAACCTTCGCTGTCATCATTTGTACA
AACTCTTCGTAGTTTACCTGACCATCCCCATCGATGTCTGCTTCCCTGATCATTTCATCAACCTCTTCATCTGTTAACTTCTCTCCAAGGTTTGTCATCACGTGGCGAAGCTCTGCTGCACTGATGTAGCCATTGCC
ATCCTTATCAAACACACCGAACGCTTCTCTAATTTCTTCTTCCGTGTCCCTGTATTTCATTTTTCTTGCCATCATTGTCAGGAACTCAGGGAAGTCGATTGTGCCGTCACCGTCGGCATCTACTTCATTGATCATGT
CCTGCAGCTCTGCTTCTGTGGGGTTCTGCCCCAGAGACCGCATCACCGTCCCCAGCTCCTTGGTTGTTATTGTCCCATCCCCGTCCTTGTCAAATAGGGAGAAAGCCTCTTTAAATTCTGCGATCTGCTCTTCAGT
CAGTTGGTCCGGCAGGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGT
TGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGATGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTC
ACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCACCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGC
TCGACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGCTCCCTCCGGTTCCGCCCTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCAT
GTGATCGCGCTTCTCGTTGGGGTCTTTCGAAAGTTTGGACTGCACGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTAGGCGAGCTGCACGCCGC
CGTCCTCGATGTTGTGGCGGATGTGGAAGTTCGCCTTGATGCCGTTCTTCTGCTTGTCGGCCTTGATATAGACGTTCTCGAGTGAGCTCAGCCGACCTATAGCTCTGACTGCGTGACCTGTCTTATTCCACTTAC
GACGTGATGAGTCGACCATGGTGGCGAGATCCTTATCGTCATCGTCGTACAGATCCCGACCCATTTGCTGTCCACCAGTCATGCTAGCCATACCATGATGATGATGATGATGAGAACCTGCGGCCGCCATAGGT
CCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGGCTGAAGTTAGTAGCTCCGCTTCCACCGGTCTTGTACAGCTCGTCCATGCCGTACAGGAACAGGTGGTGGCGGCCCTCGGAGCGCTCGTACTGTTC
CACGATGGTGTAGTCCTCGTTGTGGGAGGTGATGTCCAGCTTGGTGTCCACGTAGTAGTAGCCGGGCAGTTGCACGGGCTTCTTGGCCATGTAGATGGTCTTGAACTCCACCAGGTAGTGGCCGCCGTCCTTC
AGCTTCAGGGCCTGGTGGATCTCGCCCTTCAGCACGCCGTCGCGGGGGTACAGGCGCTCGGTGGAGGCCTCCCAGCCCATGGTCTTCTTCTGCATTACGGGGCCGTCGGGGGGGGAAGTTGGTGCCGCGCATC
TTCACCTTGTAGATCAGCGTGCCGTCCTGCAGGGAGGAGTCCTGGGTCACGGTCACCAGACCGCCGTCCTCGAAGTTCATCACGCGCTCCCACTTGAAGCCCTCGGGGAAGGACAGCTTCTTGTAATCGGGGA
TGTCGGCGGGGTGCTTCACGTACGCCTTGGAGCCGTACATGAACTGGGGGGGACAGGATGTCCCAGGCGAAGGGCAGGGGGCCGCCCTTGGTCACCTTCAGCTTGGCGGTCTGGGTGCCCTCGTAGGGGCG
GCCCTCGCCCTCGCCCTCGATCTCGAACTCGTGGCCGTTCATGGAGCCCTCCATGCGCACCTTGAAGCGCATGAACTCTTTGATGACGGCCATGTTGTTGTCCTCGGAGGAGGCGGTGCCGGAGCTGCCGCTGC
CGGTGCTGCCGGTGCCATGCCCCAGGAACAGGTGGTGGCGGCCCTCGGAGCGCTCGTACTGTTCCACGATGGTGTAGTCCTCGTTGTGGGAGGTGATGTCCAGCTTGGTGTCCACGTAGTAGTAGCCGGGCA
GTTGCACGGGCTTCTTGGCCATGTAGATGGTCTTGAACTCCACCAGGTAGTGGCCGCCGTCCTTCAGCTTCAGGGCCTGGTGGATCTCGCCCTTCAGCACGCCGTCGCGGGGGTACAGGCGCTCGGTGGAGG
CCTCCCAGCCCATGGTCTTCTTCTGCATTACGGGGCCGTCGGGGGGGAAGTTGGTGCCGCGCATCTTCACCTTGTAGATCAGCGTGCCGTCCTGCAGGGAGGAGTCCTGGGTCACGGTCACCAGACCGCCGTC
CTCGAAGTTCATCACGCGCTCCCACTTGAAGCCCTCGGGGAAGGACAGCTTCTTGTAATCGGGGATGTCGGCGGGGTGCTTCACGTACGCCTTGGAGCCGTACATGAACTGGGGGGACAGGATGTCCCAGGC
GAAGGGCAGGGGGCCGCCCTTGGTCACCTTCAGCTTGGCGGTCTGGGTGCCCTCGTAGGGGCGGCCCTCGCCCTCGCCCTCGATCTCGAACTCGTGGCCGTTCATGGAGCCCTCCATGCGCACCTTGAAGCGC
ATGAACTCTTTGATGACCTCCTCGCCCTTGCTCACTAGGCGCGATCGCAAAGCGTAATCTGGAACGTCATATGGATAGGAACCTGCATAGTCCGGGACGTCATAGGGATAGCCAGCGTAATCTGGAACATCGTA
TGGGTAATGGCGCGTCCGCGGAATCTTTTGGTCCTCATCATTCTTTTCAACCTGTTTGGTTCTTCTCATACAGCACAGATGGCGCGCCATTTTTTAAGTCTTCCAAGAGACACGAAAACGTGTGCACACCAGAAA
ATTTCTCAAACAGAAAGAAAGATTTGGGGGCCACAGCTTCCTGAACATGTGGGTTTGAGACACCAAAATAATTTCCACCTAAATTATTAACAATTTGTTTATAAATAGCCCTGTAAACATCAAATCTCCCTATAGT
GAGTCGTATTAGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCAT
AAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC
GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAG
GAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA
ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC
ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATA
TATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA
TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT
TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGAT
CAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTG
TCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACT
TTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT
CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA
TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCC
TTTCGTC

Promoter sequences

The promoter sequences shown below were cloned into the plasmid **Palmi-3xHA-tdTomato_PUC-57**. The plasmid database ID (Jékely Lab) is shown in each case:

NOMPC (6319bp)_Plasmid ID: 654

>NOMPCp

ATCCGAAAGGCTCCGCCCAATATTTACACGGAGTGCTGGTTAAACCTATTGAAAATGGTATTTAATTGGCTCAAAACTTTCAGACTTGTTATGAGGGAGCAATAACTACATCTCCGGGGGGATGAGAGGGGGC
GTTATTAATTTTTAATATGATGTCTTGCATGTCCATTTTTCTGCCAGTTTTGAAATAAATGCGGGAGTAAGCCTCTTGTTAAGATGGTGTTCTTTAATTGCACCCGCGCTTTTGTGAAATAGAATATTTCCTGTCG
AGAAAAAGTTGAGACTTTTGGCGAAAAAAAGTTGTTCGCCATTTAACACTGAAAAAAGTGTGTACGTCGCTTTAGAATAGAAAGATGTTCGTTTTCAAATGTATCGTAATGTTAAAATTGCGATAAAAAGTTTCT
TGCTCCTTTGTACAAAAAATATTTTTTCCTGGTGGTAAGATCCTGAGCCTCTCTCTTCCAACTATCGTAGATCAGGACAGCCTCTAAAGGTCATCCCCTAGAATTGACTATGTAACTTGTGAGAAACATTGGAGAA
AGGTAGTTGGCTCAAAACTCAGTCTCACTCAGACATACGTAATTGACCTGTTTCACTTCTGTGACTCTGCAGATATAGATCAACTGATAGATCTATAGATAGGCCCAGATTTGTGACAGCAGACCTGAATATATA
TGACTGACTGGTTACTGACTGGGTCCATGTGATTCCCCTTTGGTTCTGTGGCCTGTCGAAAGTGTTCAATAAACAAATCGTTGTGTACGTGCTGTAATGGGGATGGCCAACAAAGCCACATCGATTTGCTTATG
GCCAATTTAATTATGAGCAGAGCAGTAGGTTGTGATTGGGGGAGCTGGCCTATCACCCGTGCAAGCTAGCTTCCGGTTCACCATGGCAACCAGTGATGCCAGAGCAGTAAGCACTGCGCATGTCTGAGTCAGC
TCTGCCCCCCCCCATCACTTTCACCCTTGCAAAATTCCATAGCATCATTTTATACCAATGAACGTTACATTCCAACACTAAGCACTTCACTATAAAAGGAAACCTCTACTGAACTGAAAAGATAACTTGCTGGATTTT
CAAAAAGTTGTTTAAATAGTTATATGAACTTTTCTCTACACTAACTGTTGGACATATTATTACTGACATTGTGTGATAGTGGAATAACTTGTTTCCATTGTTTTTTGGAGAATTCAGGGACATAGCCATGACTTTCA
CTCGGACATCATTTTAGAAGGATTTACAGTCTAGTTCTAAGGCTCAGAATATAAACTTTTATGACAATGAAGATTAAAACGGATTAGCGATGTCTTCCACTGCAAAAGAACAAAATGAAACAGACAAGACTCCTC
CAGATAAAGCTGTTGACGAGAAAAAGGAGGAAGAAAAAAAGGTTGAGGAAGCCGCGAAACCCACACCATCGGACAAAGAGTCTAAACCGGAGAACGCGGACAACAAAGCCTCAGCCAGTGCCCAGCCACAAC
CGCCTAGACTGGAAAAAGAGCTTCCATTAGAACCAAAGGAAATACCAAAGACAGAATCAGCAGCACAAGACACTAATCAAAATAACAACAAAGATACAGTCACTTCCAAATCTACAGCACCGCCAGCAGCATTA
GATGATTCTTCCAAAAAGAAAGAACCAGAAAACAAAGATGCAACCACAGAGCTTCCAAAACAAGATGAAACAAAGAAGGTGGACTCACCAACAGAAAAAGGAACGTCTGACACCAAACAGCCAGAAAAGAAAG
ATGGAGACCCTAAGCCAGTAGAAAACAAGGAGCAACAACCTACTCCGCCAGCGACAGAAAAGCCGAAATCTCCTCCTAAACCTTCACCAGCAGAAGATAAGGATAAGAAATCTCCACCGACACCAGAAGATAAA
TCAAAACAACCAGAAGCAGATAAATCAACACCTCCAACAGATAATGTAGTAAAATCAACAGAGGCTGAAAAAGTTCAACAAGCAACTGCTCCAGAAAAGAAGGTTGAAGCAAAAGATGCAGAAAAGGATAAAA
AGACTCCACCAACTGGTCTCCAACCTATCCGCCAGCAATACCACCGAGAAAGAATGTCAGCCCAAATCCAGGAAGGAAAATTCCTGTTCCATCTCACCCGCAACCACCGAAGAAGACCCCTTTGAGTACCCAACCA
GAAAGAAGAACCCCACCGACTCCTCCGACAAACAGAAAGACTCCTCCAACCTCCAGAAAAGAACCTCCGACACCTCCCGCTTCCAGGCGTCAACAACCGCCGACGCCCCAAACCAGACGTAATCCGACGTCCTCC
AATGCCGGTAAACGATCTTCCTCTAGTTCACGACAGTCTTCCGCAGAAAAGCAATCGAAACAAAATGGCAAAGCTAATGAGGATAAAACTGACAACAAAACTGAACCTAAAGCGGGCAATAAGCCTGATAACAA
AACGGTTAATAAGATTGATAACAAAACGGAAAACAAGACTGGAGTTGCAGCCACAGGACAAAACTCGAACAATAACACAAATGGTCCAAAACCGACAATTCCGGCGAACGACAAAAATGTAAAGACTGCAACT
GACAAAAGGCCTCCAGTTTCTGCAACAGACAAATCGAATGATGCTGCTAATGCGAAGAATAATCCTCAAGATAAAACCAAGATTCCAAAATTGACAGAAGGAACAGCTCCAACAACAGCAACTGATGCTAAGTC
AACAAAAGGTGATCCCAAGAATGCCTCCAATGTTAAAGACTCCCATCAACCTAAACCTGATGCTAAACATGATAAAGATCCTAAACATGATGCTAAACAAAATGACGCTAAACATGATGACAACAAAGCTTTGCC
GCACAAAGATGACAAGAACAAAAAGTCGCACGAAGCGACTCAGAAATCTTCCAAAGGCGAAAACGAGGGAAATATCAGGAGTGACGCCGAGTTAAACGGTGCCCCCTTAGACCCTGAAAATCCTGAAAAACAG
GCTCCACCATCCAGTCGTCTTCTTAATTTGGCTCAAAAGGGAGAATGGTCAATTCTGGAACAGTCACTTCGAACTCTGGACAAAGGAAATCCTCAAATATTTGTTGTGGCAAGGTAAAGATACAAGTCGTTTATG
AAAGTGGATAATTATTGTTGTGAATTTAGTTGAGGATATTAGTAGATCGCATTCACTTAAAATGGGGGCAAATTTAATATATATCAGTTACTTGCACCGTCATCAACACCTGAAAAGTCCTTGAAAAAGGGTTAT
ATACAGCTGTGGAATTACGATTGCTTAGAGACAGGAACCAAAATGTTCATTGTGAGGTCATATTCTTGTACTTAAATGCTGTATAATTCTTAATGACACATATCTGGCAAGTGATATCATATCCTACACATGTCAT
CAGAAAGTCAATGAGTCATTCAACTTGAACTCACCATTAAAGTCACATGTTTCCAAATGACACATTTATAGTTAATTTCAGTGGGAAAGATGGGATATATCATCATTGATGTCTTCAAGCAGTTGTATATAAGGA
CTCATAATTCATTGGCTGTGATTCGTATACATTATCGAAAGACAAAATTCTTTGAAAATCACACATCACAGTTTAAACCAAATTGCATTTAGCCTTCTTGTTCTTTATGGGTAATGATAAGGTCGAAGTCGCCCAA
CTGTTAATAGTTATGACATAATTATAGGTCATGCTAATATGTCTGAATTTCATGTTTCATACATGCATAAATTCCAAAACATAAGCTGGGTTATGGGGTAGAAAGAATATTCATGAAGTCATTTTTTAAAACGCCC
TTTTCACCTGTCAACATTTGTGTCATGCAGGTCATGCCAATAGCCCAATTAAATTAGAACTTGATTTAGCCTTAATTTTCATGGTCACAATCTACTGAACCATGATGTTAGTTGGCCATGCTGTAAATAACATCACC
TTTTTCCAAACTAGATCTAAAGTGAAAAAAGGAATATGCCAAAAATCGCTTGGAGGAAAATAAAAAATCGAGGTAGACATAGGTTAATTATGGAGGGCTAGATTTAAGCATTCGTGTAAAAAGAAGTCTTTAGT
ATTGAACAACGGTTATGCCCCAAAGGCTTCAATTTTCCTTTCAATAAATACAGGGTAGCATGGATTACTGTTGGACAAAATAGATAGGCCCAGATTTGTGACAGCAGACCTGAATATATATGACTGACTGGTTAC
TGACTGGGTCCATGTGATTCCCCTTTGGTTCTGTGGCCTGTCGAAAGTGTTCAATAAACAAATCGTTGTGTACGTGCTGTAATGGGGATGGCCAACAAAGCCACATCGATTTGCTTATGGCCAATTTAATTATG
AGCAGAGCAGTAGGTTGTGATTGGGGGAGCTGGCCTATCACCCGTGCAAGCTAGCTTCCGGTTCACCATGGCAACCAGTGATGCCAGAGCAGTAAGCACTGCGCATGTCTGAGTCAGCTCTGCCCCCCCCCAT
CACTTTCACCCTTGCAAAATTCCATAGCATCATTTTATACCAATGAACGTTACATTCCAACACTAAAGCACTTCACTATAAAAGGAAACCTCTACTGAACTGAAAAGATAACTTGCTGTGTATATGTAATAACTAAC
ACTCGTGGATGTGCTTTATCATTCATTCGAGGACCTGCCAATACTAACATTCGTGGAGGTGTTACAGATTTCAGTTTTAAGGTCAAAATTAGATTTAGGGTTAGTTTTAGCCAAGTTTAGGGTTAGTATGTTCGT
AATTACACCTATACAAATTCGTGAATGACAATATTTGACACCTCCACGAATGCCTAATTCGTGTACATGATCGTTTTTGCAGGTTTACAAATTTGTCTACTTGCAAAATATTGAAAAATATATACCTCGATTTTGAA
CAGTATGAAAGGGCCACTAATTTCGATCAGGACTTATAAACAGTGTATCAAAAGTATACATTGAATTCAACACAAAGCTTTATATCAGTTGAGGGACGTTCCAACAGAGGACCAGATCTAACCAGAAGAAGGTC
AAGTGCCCGATTTCGTAATTACATCTTTTTTCAGCATTACTCTAAAAGTTGATAAAAGCTATAGCGCTAGTGAAAATTTTGTTGAAACTGTTTACAAACTTTGGATGGGATTTTTAAAAAGAGTAGGATAACCATC
TAAGTGTTGAAAAGCTTTACTCAAAATGGTTGGGTAATAGCTGGCGCCATAGTGCTGCTAAAAATGGTCATGTTCTCAAGTACCAATGCCATTGTCATGTTCTCAGTCACTAATGCCGTTAGCCAAAAGAGAAG
AAAACTCTGTCATGTTCTCAAGCACTTATGCCATTGTCATGTTTTCTAGTCCTGATGCCATCAACCAAATAGAAGAAATCACACTCATGTTCTAAAGCACTATAATGCCAACAGCCAAAAAGAAGAAAACAGTCAT
GTTCTAAAGCACTAATGCCATCACCCAAAAAGGAGAAATCACTGTCATGTTCTTAGGCACTAAAGCCATTGTCATGTCTTAAAGCACTTATGCCATAAGCAAAAATGAAGAAATCACTATCATGTTCTCAAGTACT
AAAGCCATTGGCATGTACCCTTCGGGTCTCGGGAATATACTGGGGTCTCGGGGAGTATACCGGGGATGTATGCATCTACCCATGGTATAACCTCATAACATATAACGGAATGATGCTAAGAGCAGTTTTTGCT
CTATAAGTCTCAAATAGTCACTCTCAAGTATTAGTTTCTACATGTAGGCCTATATGTCTGACACTTCAATTAAAAATCACCAACTGTGGAAGTTTCAAACTAATATAAGAGTGTCTCAGCAATCAAGGGTCCATTT
CTAGAGTGATGGGATGATTCTTACCATCATTCCGTCACACTCGATATGAACCCCTGAATACTGAGTTTGAAAATTGAGCAATGGTGTTTTTGCTGTTTGAATTGTTCTCTTCACTATCAGATAAATATATAACCCA
AAAGCTAGTACCAAGAATGATTAAATTGAGACTTATAGAGGAAAAAACTGCACTTCGCGTGGATTTCATGTCTTGAAGGTTCACTGACAATATTTTGAAATCAACCAACTTAACCAACCAATAAAGTAAACTATC
AACCTATCAATCAACCAAAAAACCAATCAATCAATTAATCAATCAATGAAGTGATAACGTTAGCAGACCAGCTATTTGCTAGTTTTATGTTGCTAATATATATATGTATCACAATGATCTTTGCCTTCGTGATCATT
TGTTTGTCATTGAAGTGGGCCCAATAGTCTCTAATAAATCTGATGCAGTAACCAAAATATTTGTTCAATTGTAGGACACTGGAGTTACACCCCTGATGCTGGCTGTGAGGGAG

PKD1-1 (2567bp) _Plasmid ID: 655

>PKD1-1p

GGAAGCCTGATAACAAAGTGAGTAGGAAAAAAAAAAAAGGAGTTTTGAAAATGAAGGAACTAAAAAAATAAGGAAAACGTTAGATTATGATAGGAGACGGTCTCTTGGGATGTGCCTTTGCACACTTTTCTTA
CACATTTTCCATAACTCTCATTTCATTACACTTTTAATCGCTTTGCTTCTCCTGGATTATTTTCTGACTTGCTTACGTTTCAGTTTTTGTCCTGGCAACAACCGAGCTGGACGGCAGGGGCCACCTGGATGTCTGGC
TGTAACTCAGTGAATGTTGACAAAGCCACATTAGGTAAACTGCCCTTGCTCATGCCCTTCCTTTTTCCTATATTTGTAGAGATACATTCACTTACTGGGCAAGTCAATATAGACATGTGTGTGGGAATAATCTCTT
GTGTTTTATATTTTCAGAAGCCATTGGACTCCTCTGACCTAGAATCTCATCAGTCTTGTTATTCAATTTTAGTCCTTGCAATTAATTGTTTGTTAGATCTACATTAAAATTGTATTTTAGAATCAGGTGGTGCTATG
TGTAGTGTACATGTATTATTGTATAAATTGTTTCTTTATATTACTAGTGTTGCACTGAATTCGTTTTCATCCAATATTTAATCATGTTAATATATCAGCAATAAAGATATAAACTGTTAAAATCCCTGCTCAAATAAA
TTTCCCTAGAAATTTTAATATATACATTGATTGATTCCTATTCCCCTCATGACATTTGTTAATCCTTGTGAGCAGCAAGGACTACCTTAGAGGACTGCACATCATATCCAATTATCCATGCACCTGTATAACATGTC
GTCTCATATCTTTATAAACCTCATGAAACTGTTGACATTTTATTTCTTTGCGGAAGTAAACATAATATGTTTAGAAGCATTCATATACTGACTGGAATATTTCTCATACTTTTGTGTAAGTCATCTGCAACTTTTTAT
TTTTTTATTAATACATGTTGGCAATATCTATTAGTTTGCAAAAATTGATAGGATATTGCATCTCCAAATCTAACAAACCAATTTTAACCATGTTGTTCGTTTTATTGTTTTGCTGTGTAAAAAATGTTCATACTAATA
AAATGAAACATTTGCACCACTGGTGCGGTAAATTTAGAAGTCATGTTATGGGGCTGAGCGAGTTTAAGTGTAAAATATGTAAAAGAAGAACTTGTAATATGATATATCAAATTGGAAACATGCTTTGTTTCTTA
CTGCACTGCAATAAACATTTAAAAGAGAAAAACATGCCTTGCCCTTGTGGTTTCTTGTTTCATAATGTTTGACCAAAGTGACCGAATTGCGTAAATCTAATTTACCTCCCAATATCAAATTATATACTTTTTGGAC
TAACTTCAGCCAAGTCAATCAAATTTGATTTTTAAAGTTCATTATCATCACCTGAACCTAGTAAAGCGAAGTTACAGAGCTGGTATGTTTCGATTTATGTGATAGGGCATACAATGTCAGGGATCTGTTCACTACA
AGTATTACTGGAAAAACCAAGTCAGCTGTACAAAAGCGCAGAAGTAGATTGCAGTTGTAAGTTTGGTGCCAAACTGAACTAACATATATGTAAACAATAGCATATCAATTCAGAACAATTGAAAGAACCCAAAT
TAGAGAACAGAATAGATTAAGCTAGGTGCCAACTAAGTCTTACACAAAACTGAAGTCAAGTCTATACTTTAAGCCAGTAAAAGACCCTTCAATTTGAAAAAAAAAATTAATGAAATTCTGACCAAAACTTAATGA
GAAAAACGGACTTTTTACCCCATCTTCAGGGTCATGAAAAAACCCTCTATACTTATTAGTATCATACTGATACTTCTTTAACATCTAAAAGGAGTCTTAAGCTAAAGATTGATATAATATGTGGGTAAGGAAGCTC
TTAACTAGCCAATTCAACCCACCCTACAAAAAGTGTCCGACTCAAGGTTTCTTAGTTTCTCAGTTTCTCCAACTTTCTTGATCCCCTGGTGTGGAGAAATTCGTGGTGAGCCTTCCTTAGTGTCTGACTCAAGTAA
CCTATACCCCTATAAAGCTCTTTCAACACTTGACACTCTGATGGTAAAACTGACGTCAGAGTTATTAATAGCAAACAACCTCACAGAACTTGATGGATGATTTGAGATCCTGTATTGCACCAGATTCAAGCAATCG
TTTCAACCTGCCAACCTGGGGTGCCCAGGGGTTGGGAGCAACCTGTTGATGACCAAATTGATTTGATTTTCTCCTGTTCCAATTGCTTACGTCATCTGGCAGAGGTGAGTTGTCAACAAGGGGATGTTCTGTGCC

AAGTAATTAAATTTCCTGCAGGGGAAACCGATGAATAATAGAATGGATCTTAGCAATGAGCGTGACTGGGGTCCACAGGGATCTGAGGGTGTTGCACTCTTCTATCTCAGCAGGAGTTTGTAGAGATGTTGTT
GCTAGGTTACGGTACCCCCCTCCATCCAGAAGGGGATGCTCCGCTGAAAGGAAGTCGAAAAAATCTTCAGAATTCGACGAGCCTCGATCAGGGCGATACAGACGGAGATGGTAGAGCCACTT

PKD2-1 (1473bp) _Plasmid ID: 656

>PKD2-1p

AATAAAATTTGAACCGAGGCCAATGGAGCGTGACGCTTATCTTGACGTCAGACCTGTTGCCCGGCAGCCGCACCTTCTGAAACTGCCCAGCATCTCCGCTCCGCATCCCACACCTCAACGGCATCTTCAAGAGGT
AAGTGTTGCTTTAAGAAGTTTTCAAAGATTTGAAAGAATTTTTCTGTTGCACCTTTAAAAAACTTTTAAGCCATCTACTTTAAAAATTTCAATTACATTGTAATTTTTTTTTTTTCAATAGGCTCACTTTAAGAGTGC
TTAAAAAAATGTTTTATTTTTATCTGGTAATTTATTATTTTGAATTTTAGAATTTTGCAAATTGTTACTAAAGCCATCAATATTTTAAAAATTTCAATTACGTCGTAAGATTTTTAAATAGGTATACTTTAAGAGTGC
TTTCAAAGAAATTTTTTATTTTACCTGCTTTATTATTTTGAATTTTAGAATTCTTGCAAATTTCAAACAATTTGTGCTAACAATTTGTGCACATGTTTTCGTGTCCTTTTGAAGTGATCTCTTGCTGAAAAATGAAGA
TGCTGCAAAGTTTTTCATATATATATTCATATATATATATTTATCAAGTTGGTGACAATTATTCATCTGAATTTAATAGTGTGAAAATATGAAAGTTCAAAGCTTGTGACATATTGAATAGATTAAAAGCAAAAGA
TGCCAAAAGAACCAAGAAAATATTGAGAGAAGTAGACTGATTTTTTTTTTCTTGGCAATTAAGAAGAAGACAGACAGAAAGCTTCATTTTGTGAAGAGCTACTAGTGTTTTTTAGAATAGCGAAGTAGAAGGTT
CAAAATAATATAATCAATAAAAGCATTCTAACATCAATGCACAACATATGAATATGTCTGCCAGTTATCAATATGTTGAATAAGTAATAGCACCTGCAGCTTGCTGCTAGGAGACTGTGCATAATTGAAAGTTTAA
ATTGATGTGGAAGTGTATAGTTGTATTGATGGGCAGAACAATTGAGCTGTTTTGCAAGGTAAGAAGCAGACCTGTGCATTTGCATTCTTTGTGAGGTCATTTGTCACGTAGAAATGTAGAAGTAGGAGATAAT
TCAGAATGCTATGAGGAGGATGTTGGAATGACTTTTCTTCAAAGTGAATTTACTATTTCACTGAAGTTGCAAAGTGAGATTATTTTTTCGATCATAGGCACTTGCCCTTGACGGCGTATGCATAATGAAAAGGG
AATGCTGAATCTGAAACGCCCTTTCTCGATTTGGGTTTTTTTTTTATCAACGTGGTTTTTTATGCGCAATTATTGGATAATTTCGCAATTGATCTCATATGATAATATTTTACCTAAGGTTTCATTCTATCTTATGT
GAATGAAATTGTGTTATCAATGATATTATATGTTAAACTCTTTGTCTGTTTCAGATTGACTTATTTGTAGAACGGACCATGCAGCAACCAACTAACTGGCTGATGTCACGCCCC

Sequenced inserts and ESTs used for WMISH probes

NOMPC

>*NOMPC*_ISH1 Plasmid_ID:269
TTACAACTTGATGGAGGACAGAAAGTGGGTATTTGACCTCATGGCACTTGGCAAGATGAACGACAACAAATCTATCCAGGAATTTATTCTGGTATCTCCGGCCCCGGTTGACACTGCTGTCAAGTTGTCGCGTG
CCTTCAGGAATCTCTCAACGAAGGAAAAAGAAAGAGCGAGGGATTTACTGGCCGCTGGTGACTTTTGTGAAGTAATGGCGACAGAGTTGAATGCCATTGCAGCCGGTATCAACGGCCCGGCTATATTACTCAA
ATCAGTGGACAATCGAGGAGTGCCTCTTCCGGATATTCTTATTGAATGCGAACAGAAAGAAGTCGTTGCTGATGCCGTCGTCCAGCGGTATCTGTCAGAAGTGCGGGCGGGTTACCTGAGTCACTGGTCTTCT
CCCAAAATGATGCTGCTCTTCTGTGTCTTACTGTTTGTGCCTCCTGTCTGGCTGGCTTTCTCTCTTCCTAACAAACACATTCGCTTCAACAAAGTGCCTATGATCAAGTTCATGCTCTACTTAGTATCTCACATCTAC
CTTATCATGCTTTTCACTCTGACTATCGTGTATCCGATCTATCCTCTCTGGGAATCTGGAAACTTGATTCCGCACTGGCACGAGTGGCTGCTGCTTTCCTGGTTTTCCGGCTTGTTGGTCAACCAGCTGACGAATC
CAGAAGATCGTCAAGGTTTAGGATGGATCAAGGTCATCGTGATCGCCATCAGTAACTTCGGGGTCACCACTCATCTCATCGCTTTCTCATTCTCTGGCAACGATCGCCTGTACTGCTTGTATGCTCGTAACCAAT
TCTTCGGTCTTGCCCTTATGCTGTGTTTCGTCCAACTCCTGGACTTCTTGTCCTTTCATCAACTGTTCGGACCGTGGGCCATCATCATCCGTGATCTCATGAAGGATCTGACCCGTTTCCTCGTAATCTTAGTTATC
TTCATGGCTGGC

>*NOMPC*_ISH2 Plasmid_ID:270
ATGTCCGACACTTATCAAAGAATTCAGGCGTAGTCTGACACAGAATGGAAGTTCGGAAGAGCGAAGTTGTTCAGGAATATGAACAAAACTTCTGCCACTCCATCTCCATTGAATCTCTTCACAAAACTTTATGTG
TATATAGTCATCCTCATTAAGCATAAAAGCAAAGTATGTAGTCTCACAGATGTCCAGCAGTACATAGACGAAGAAGAGGAGCTGGACCAGGACTCTGAGAATGCAGCAGCTGGGCAAGGGAACTGGAGGAAT
ATGATGAGGAGAAACACACAAGTTGCTCCTGAAGATTTCATGCAAGCTGCCTCGCGGAGCAACGGTCCGACTACAGTGGAAGAAGTCGTCGATTGGAAAGTTGTTGCGGACAGATTTCTTTATCTCCGAGGTC
TCGCAGACGGAAATAATATGGATAGCGACAAAGAAGATGATGATGATGATGACAAGTTTGAGAGCAATGAAAATGCTGCTGAACAGCAATGAATTTAATATGTGCA

>*NOMPC*_ISH3 Plasmid_ID:271
GCACAAAAGAAGTCAAAAGAGACACCCTTGCACATCGCAGCCAGGACGGAGGATGGAGAGAAGTGTGCTGAGATGCTTGTGAAAAGTGGGGCTAATGTGAACGCAACTAAAGAAAATGGAGAAACAGCGTT
GCACATTGCTGGCCGATACGGTCACCTGCAGATGATCTGTGCTTTACTGGCAGAAGGCGCAGATCCAACCTGGCAATCTAAGAGTGGAGAAACGCCTCTTCACATCGCAGTGCGCTACTGCCACTGGGAAGTA
GCGGATGAATTATTACGGTTTGTGAAGAAAGAAAGGTCACCGATTGACGCCACAATGTTGGTTAATGCACCAAACATTGAGGGAGAAACTGCTGTCCATTATGCTGGAGAATTAACAAAGAAACAGGCACATC
ATGAGTTTGAGGACACAGATATGATCAGTTTGCTACTCCAAAATGGGGGAGATACTAACACCCATACAAAGTTGACCCAAGAAACACCGATCCACTACTGCGCCAGGGCAGGCAATGCAGACGTCATGCTTGA
GATAGTGAAGCACCTTGGGCCAAACCAGGTGCAGGTTGCTGTCAACAAACAGGCTAAGAATGGCTGGTCACCCCTTCTAGTGGCAAGCGAACAAGGTCATTTGGACATCTTGAAAATTCTTCTCCAGCATCATG
CAAGAGTGGATGTCTTTGATGAGCATGGCAAAGCTGCATTACATCTCGCAGCAGAGAACGGTCACGAGGAAGTCGCGGATGTTCTTCTGCAACACAAAGCTTTCGTCAATGCCAAGTCCAAACTTGGTTTGACT
CCATTGCATCTAGCTGCCCAATCAGGCTACAACAAACTGGCAAGGCTACTCATAGAGAAGCATGGGGCTACCAATGATGCATTGTCTTCGGCAAAGAAGACCCCACTCCACCTTGCAGCCCAGACTGGTCAGAT
TGAAGTGTGTAGTACCTTGCTGAAGATGAAAGCAGATGCCAACGCCAAGGATGCTCACAGTCAAACGCCTCTCCATTTAGCAGCAGAAAACGACCATTCAGATGTGGTCAAACTTTTCCTGAAACATA

>*NOMPC*_Conn Plasmid_ID:313
TTACAACTTGATGGAGGACAGAAAGTGGGTATTTGACCTCATGGCACTTGGCAAGATGAACGACAACAAATCAATCCAGGAATTTATTCTGGTATCTCCGGCCCCGGTTGACACTGCTGTCAAGTTGTCGCGTG
CCTTCAGGAATCTCTCAACAAGGGAAAAAGAAAGAGCGAGGGATCTACTGGCCGCTGGTGACTTTTGTGAAGTAATGGCCACAGAGTTGAATGCCATTGCAGCCGGTATCAACGGCCCGGCTATATTACTCAA
ATCAGTGGACAATCGAGGAGTGCCTCTTCTGGATATTCTTATTGAATGCGAACAGAAAGAAGTCGTTGCTGATGCCGTTGTCCAGCGGTATCTGTCAGAAGTGTGGGCGGGTTACCTGAGTCACTGGTCTTCT
CCCAAAATGATGCTGCTCTTCTGTGTCTTACTGTTTGTGCCTCCTGTCTGGTTGGCTTTCTCTCTTCCTAACAAACACATTCGCTTCAACAAAGTGCCTATGATCAAGTTCATGCTCTACTTAGTATCTCACATCTAC
CTTATCATGCTTTTCACTCTGACTATCGTGTATCCGATCTATCCTCTCTGGGAATCTGGAAACTTGATTCCGCACTGGCACGAGTGGCTGCTGCTTTCCTGGTTTTCCGGCTTGTTGGTCAACCAGCTGACGAATC
CAGAAGATCGTCAAGGTTTGGGATGGATCAAGGTCATCGTGATCGCCATCAGTAACTTCGGGGTCACCACTCATCTCATCGCCTTCTCATTCTCTGGCAACGATCGCCTGTACTGCTTGTATGCTCGTAACCAAT
TCTTCGGTCTTGCCCTTATGCTGTGTTTCGTCCAACTCCTGGACTTCTTGTCCTTTCATCAACTGTTCGGACCGTGGGCCATCATCATCCGTGATCTCATGAAGGATCTGACCCGTTTCCTCGTAATCTTAGTTATC
TTCATGGCTGGCTTTACCTTGGAACTAGCTACTATTTACCAGCCTGTGTTAGCTCCTTTGATACCAGAGCCTGACAAGGGAACAGGGGATGGTTCTGGAGGAGCAGGAAAAGAGACTCCCATCAGTACTTTTGA
ACTTCTCTTCTTTGCGCTACTTGGTCTCGTTGACCCGGAGAACTTACCCCTTGTCAACAGGAGTCCCCTGTGGAGTGTTACCCTTGCTAAGGCTGTTTTTGGATGTTACTTGACACTGACAATTATCGTGCTTATC
AATCTCCTCATTGCTATGATGTCCGACACTTATCAAAGA

PKD1-1

>*PKD1-1*_ISH1 Plasmid_ID:255
TGTCTTTGTAACTGTTGGTCTGGTTCTCCTTCTCTATCTTGGAGTCATTATTTGGGCAAGGAGGGCGGACAAGAAGGATGTCATCAGGTGGGGAGCTCAACCTCTTGAAGACAACCTGCCATCAGATCAGTACC
ACTACCAGATCGCAGTGTACACAGGGGGCAACAAGAATGCTGAGACTACTTCAAGAGTAAGTCTCATCATCTCTGGGGATGATGGCGACACCGGAGTCAGACGCCTTTGGGATGGCAAGCGAAAGGAACTTG
AAAAGGCCAGCATCAACAACTACATTCTGTCAACAGAAGGATGCCTGGGACCCCTTTCTTTCCTTCGAATCTGGCATGACAACACTGGAAGTGGGAAGAAGAGATCCTGGTTCTTGGACCAGGTCCAGGTGAA
GGACCTCCAGACTGGTGAACTCTTTTGCTTCCTGTGTGAGCGATGGCTTGCTGTTGAACAAGATGATGGATTAGTGGAAAGGATCATCCCTGTAGCAGGACTGAGAGAAATGACTGCCTTCAAACACCTCTTCA
ACTCCTCGGTCAAGAAGAAGCTGGCAAACGACCATCTCTGGGTGTCAGTGGTCAAGCGTCCTACTCGCAGCAACTTCACCAGAGTTCAGAGAGTGTCTGCGTGTGCCGCTTTGCTCTTCCTCACCATGATTACCA
ACTGCATGTTCTTTAAGGCTGATGAAAACAAGGAGAATGTCCGGGCAATCAGCTTCGGTCCATTTGAATTCACTCTGCACCAACTGTTCATCTCTGTGGTGACGACCTGCATCGTGTTCCCTCCATCTTTGCTCAT
CATCACCATCTTCCGACGAGTCAAACCCAAGAGCAATGGTGTGATGCAATCCAACCAGAGGTTCTCCCAGAAGTCCAGCAAGTTCAAGTGGAGAAGGTACAACCCTGGATCATCTCTGTGGGACAACCAGCCAA
GCAGGGCCCAGAGGATGAAGGAATCTCTGAGGAACAT

PKD1-2/LOV-1

>*PKD1-2/LOV-1*_ISH1 Plasmid_ID:281
AGGTTTTATAAATGACCGGAACAATCGTATAATGGGCTTTGCCACTATGAGGCAGCTTAGAGTTAAAAAAGCAAAATGCAATTTAGTTAAACCGATGGACAAAATATTAAGGGAATGTAATGTTGCATACGCCT
TTTACCACGAAGACACTGAAACAAGAGGGGTTGGTTGGGAACCGCTCTACTCAAATAGCACATACAACAATTCAGCTTACGAGTATGTCCATCGATCAGCCAAAAGTCTGGACAGCTTCCCATTCTGGGCGGTG
CACCATGTTTACGGTGGAGGGGGTTATGTGAGGGAATTGCGCGGAAGTACCAACAGGCTCAAACAACACATCAGAGAACTGCATGACGGAGGGTGGTTCGATCACTACACTAGGGCGGTCTTCATAGAATTC
ACAGTCTACAATGCCCAGGTCAACATTTTCACCATCTGTACGCTCGTGGCGGAGTTCTTACCGACAGGTTCCCTCTTCACCAGTTACAGATTTGAGCCAGTGAATCTACTCGGCTACAGCATGGACACGGCCAGC
TTTGAGATCATCTGCCAGATCATCTACATGCTCTATATCTTATTCTTCATAATCAGCGAAGCAAGAGAACTGTACAGGAAGAGAAGCGCCTACTTCACAGAATGGTGGAACTGGGTTGAAATGATGATCATTTTC
CTGTCGTTATCAGGAGCGGTCATCTTCTTCTACCGTCTTGTGATGGCAAGCAAGTTGTCCAAGAAGTTTGAAGAGTCAGGCGGAAACGCTTACATGAAGTTCCAGTACGTTGGATACTGGAACGAGCTCCTCTT
GTACATGATTGGTTGGTTGGTCTTCTTGGCAACCATCAAGTTCCTGCGCTTACTCCGTTTCAACAGGAGAATGTCCATGTTGGCGTCTACTCTCCGTAACTGTGCTA

>*PKD1-2/LOV-1*_ISH4 Plasmid_ID:329
GGCTTTATTCGACAACAAGAAAATGGGAAATGGGAAGTACAAAGCAATAATTGCCGTAAGAGAGCTTCAAGGCTCGGAACTCCAGGCTTATAAAAATGACTCGACTAAGATCCCTAAATTCACCATTCCTCAGG
ACAAAGTCTTCACCACTCCATACAGGACTGAATTCACTTGCATGGGAGCTTTTTTTCTTTGATTCAACCAATGTGGAAAAGGGCTGGCACGGTGAAGGATTGACTGTTGGCTCAAAGACCGTGCCAAGCAAAGT
CCACCTCCAATCAACCCATTTGTCCCTTTTCGGAGCTGGCTTCGCTGTCCAGCCCAACAAGATCGACTGGGCCTTCGTCTTCTCCAACATGGACTTCTTCAAGAATCCGACCCTGTACATTGTGGAAATCATCACTT
GCGTAATATTCATCATTATGCTGATATGGGCAAGGTACAAGGACAAGCAGGATAAAATGAAGCTTGGGGTATGCCCATTGCCTGACAACGAAAAAGAAGATAAATACCTGTATGAGATTCTGGTCTACACAGG
TTTGAGAACCAATGGTGGAACCGAATCGAAGGTGAACTTCGTTATCTCTGGAGAACATGATGAGACAGACGTCCGAACATTTGAGGATCCCGAAAGAAAGATATTCAAACGCGGAGCTGTAGACGCCTTCTTG
ATGGCCACACCACATCCCCTCGGTCACCTGAACTACTTGCGCGTGTGGCACGACAACAAGGGCAAAGGAAAGTACGGGAGTTGGTATCTGCACTACGTCATTGTCTCAGATTTACAGACTCGGCACAGATATTA
CTTCATATGCAACAGATGGCT

PKD2-1

>*PKD2-1*_ISH1 Plasmid_ID:256

TCCTCGTCATCATAATATCCATTGTCTGCATAGCGTTCAACATATATCGGACACTGGCGGTCGGGAAAAAGCTCGACCAGCTGCTTGCGAATCCGAACATGTTCCCGGACTTTGAGTTCCTCAGCTACTGGCAGG
TGCAGTTCAACTCTGCCATTGCCATCACCGTCTTCATGGCTTGGGTCAAGGTGTTCAAGTACATCAGCTTCAACAAGACCATGACCCAGCTCTCCTCGACCCTTGCTGCCTGCGCCAAGGACTTGGCTGGATTTG
CCGTCATGTTCTTCATCATCTTCTTGGCCTTCGCTCAGCTCGGATACCTGATCTTCGGAACCCAAGTCAAGGACTTCTCCAGCTTCGACGATGCCATCTTCACCCTGTTCCGAATCATCCTGGGAGACTTCAACTTC
CAAGAACTGGAGGAGGCCAATCGTATCCTCGGACCCATCTACTTCTTGCTCTACGTCTTCTTTGTGTTCTTCGTTCTGCTCAACATGTTCTTGGCTATCATCAACGACACCTACTCGGAGGTCAAGGCCGAGCTG
GCCAACCAGAAGAGCGAGTTCGAGATGACCGACTACTTCAAGAAGGGCTACGACCGCATGCTGGACAAACTGGCCTTCAAGAGGGACAAGATCGTTGACATCCAGAAGGCCATCAACTCAGCTGACCACAACC
AGGACAAACAACTGGACTTTGACGAGTGGCGTCATGACCTCAAGAAACGTGGCTACGCTGAGGCGGAGATTGAGGCCGTATTTGCCAAGTACGATCTGGACGGAGACAGAGTGTTGGACGAGGGCGAACAG
ATGAGGATGAACCGCGAACTGGAAGGACAAAAGGACCAACTCAACAAAGAGAG

PKD2-2

>*PKD2-2*_ISH1-2 Plasmid_ID:280
GACACTTACTCAGCTGTCAAGGAAGACTTGGCCAACCAGAAGAGTGAATTCGAAATGAGTGACTATTTCAAAAAGGGATACAGCAGAATGAAGGAGAAACTCAACTTCAAGAGAGACAAGATCATGGACATCC
AGAAAGCCATCCAAGCTGCGGACACCAACCAGGACAGGATGCTGGACTTCGAGGAGTGGAGACGAGACATGAAGACGAGGGGTTACGGAGAAGCCGAGATTGAAGCGGTGTTCGCAAAATACGACCAGGAT
GCTGACAGGGTATTGGACGAGAATGAACAGAGTAAGATTAACGCGGATCTTGAGGGACAGAGAGTGGCTGTAAACAAAGAGTATGGGGACCTAGAAGAGCGAGGGAATAGACCTGGGTTCGACTCGGGAG
ATGAGAGTGACGATGAAGATAGCGGAACCAAGTCATCTCGCACTGGTAGGGTCATAAACGGAGTCTCTTACGATGAGTTTACGGTCTTGTTCCGAAGAGTGGACCGCATGGAGCACTCCATTGGAAGCATCGT
ATCCAAAATTGATGCTGTGTTGGTCAAACTGGAAGCAATGGAAAGGGCAAAATCAAAGCGAAGAGAAACAATGAGCAAGTTATTGGACAGCATTACTGAGGCGGATGGAAGTAGCGATGAGTCCAAGAGAG
AACAAATGGAGAAACTTGTCAGAGAAGAATTGGAACGTTGGGATTCCGAATCGTCCATCAGCCAGCCACCAGGGGGAGCCAGGGGTACGTCACCAGGGGGCAGCAGCTCTGGCAACAGTGGGGGTGTGCGG
GCACGCTCCCGACCCATTGCCTCAGTCCAGAACAACAGTGACTACTTTGAAAGGACCAACAGCGGTAGCAATGCATAATTCAGGTTTTCTTTTATCTAGTATTGCCTTGTCTCAGACTCCCGTTCTGTCCCTTGAT
TTTTGTCCTAAAGATGTATACCGGATCCGAG

Piezo

>*Piezo*_ISH1 Plasmid_ID:251
GGCGAGGGTTCTTCCTTGAGAACTCGAATGACGTGGAGGAGTACGGAGGAGGAGACAACAAGGACATCCTGGAGGACGTGGAGGCGGGCACGGAGAACCCCGTTGAAGACTTCCTCTCATGCAAGGATTCG
TACCTGGACTACCTCAAGACGGGGGTGTTCATGTACTTGTTCTGGATGACGCTGGTGTTAGTCTTCATAGCGGGCACCACACGCATCTCCCTCTTCTGCATGGGCTACCTCATGGGCGTCTTCTTCTTCCTCTGG
TTTGGCCAGGAGATGTTGATCAAGCCACTCCGCAAGCTCTTGAGGCTATGGGGCATCCTGCTCGGCTACAATTACTTTGTGTTAATGTGCAAATGTGCTCTGCAGTTGGTTGGCTGTGTGTACATCGATGTGAT
CAGCAACCAGTGCTGGCTTGTGCAGCTCCTCAGCATAGTGTGCATCCAGGCTGAGCAGTACAGGGAAAAACTCATCCCCAAATGCTCCATGAAAGAGGACCAGGCGGGGTTGATCTGGGACAGCTTCTGCTTC
GTCTTCCTTCTCTTCCAACTGAGAATTTATCGCAGCCATTACTTCAGACACGTGGTCACGGAACTGAACATCCAGCAAAAGCTCTCTGCCAGGGGTGCTGAATTGATAAACGAGAACCTGATGAAATCTGTTGAA
GAAAGGAAAGAAGCAGAGAAACAGGTCTTGCTGAAGATCAAACAGAAAATGGAGAGGCTGAAAATGAGACAAGAAAGACTGGCCGAGA

>*Piezo*_ISH2 Plasmid_ID:264
TTCTGCGTCTCCAAACTGTGGAAAGTAGCCTCCGACCTCCTGTCTCGCTTCAGCGACCTCTTGTGGCGAATTGTGGAAGTGCATTTCATCAAAGCCGTTTTCTTCCCCATATTCATGTGTGCAGTCATGGAGGTG
TCTGCCATCAACGTGGTGTTCATGGTGTTCCTCCTGGTGGGCATCCCCATCCCCCAGCTGCACAAGGCCGTCTCCCACGCCTGTCTCGTTTGGTGTGGCCTCGTCATTCTCGTCAAGATGGTCTTCCAACTACAG
ACCATTCAAGACGCCCCCAATGTCTTCAATTCCACCTGCATTAAGTCGAATACGTCAAATGCCACTCCTCTGGAAATACCAAGTGACAATGCCAAGTACCTGGGCTTGCAACAAGCCCCCAACCTCGCTTACTACA
TCAGGAATTACATAGCCATCCTGGTGGTCCTTGCATTTGAGGGGATCATCTGGATCAGACAGGCCCAGCACTACAACAAGCCGAATGTGGTTAAGCCAAAACAAGGAATCATATTCCCTGAAATCAAACGCGCT
GATGCTGACGCCAACCTTCTCAAGTTCTTCAAGTTCTTGCTAATTACTGCTTCTATAAGTTTGGACTGGAGCTGTGTTTCATCATGACAGTGATCACCATGGTGGTGCGCCTGGACGCTTACTCCTTGGTCTACG
GCTTCCTGCTTGGCTTCCTGCTCTGCCTCAGCAGACGTGCGTGCTTCATCTGTTGGCCCATTTACATCATCGTCTTGGCCGTCCTATTAGCCGCCCAGTACCTAGCAGTGCTTGGAGCGCCACCTATATTATGTTG
GGAGTATCCATGGTCGACATATATAAGAGATAGCCTAAAAGTATTCCTATATTTGCCGAGTTATGAACACAGCCCAGATGCCATAAAGCTAATAGCCGATTTCTTCCAACTCGGCTTCGTGTGCATGCAGTGGC
GAGTGTTCTTCCTTGAGAATTCGAATGACGTGGAGGAGTACGG

>*Piezo*_ISH3 Plasmid_ID:254
TGATGGACTGGATCTGGACAGACACGACCTTGGCCCTCACGTCTTGGCTGGAGATGGAGGACATCTACGCCAACATATTCCAGCTCAAGTGTTGGAGAAAGGCCGAGGATGCCTATCCGACCCAGCGCGGTCA
GCTGAGGAAACTGGCCATCAAGTATGGGGTTGGTGGTCTGCTGCTCTTCCTGCTCATCTTCATCATCTGGTTCCCTCTCTTGCTCTCCTCCTTGGCCGACACCGTTCTCTACGTCCCCAACCCCCCAGTGGACTGC
ACTGTGCAGATCAAGATTGGAGGCTATCAACCCACATTCCATATGAGCGCCCAGCAGAACTTCCTCAAGCCCATGAATGAATACAACTACAGTCGCCTGCTCAACATGTACAGCGATAAACCGGAAGCCTTGTCG
GTTCTTGAGCAGTACAAAGCGTCAGACATTACCATCGTGCAAATAAACGGGGAATCAACCGCATTATGGGGCATTAGCCCTCCCAGCAAAGACGCCCTCTTCCACGATCTCATAAATAAAAACACCAATTTGAAA
ATATACTTCTCCATATCATTTAAAAGAGATCCTAAGCCGGGCCTCAACACAGATACGGTCACTGCGCAATTTCCTCATCCTTTAAATCCCTCAAATAAAACGAGGGAGGAGTTCGCATTGATGATCAATGGCAATA
GAACGAAGCCGGTGAATGTCGTTGGCCTGTTTCCACGCTTCATGCGAGTGCCTCCCAGTGAAGCCCCTCAGGAATTGTTAGCCCTCACCA

>*Piezo*_ISH4 Plasmid_ID:265
CCTGCCCAGTTGTGGTACTTCGTCAAGTGCATCTACTTTGGATTGTCCGCCTACCAAATCCGTTCTGGCTACCCAACCAGGATCCTGGGCAACTTCCTCACCAAGAGATTTGACTACACCAACCTCTTCCTATTCAA
GGGATTCCTGGCCATCCCATTCCTCCTTGAGCTGCGAGCTCTGATGGACTGGATCTGGACAGACACGACCTTGGCCCTCACGTCTTGGCTGGAGATGGAGGACATCTACGCCAACATATTCCAGCTCAAGTGTT
GGAGAAAGGCCGAGGATGCCTATCCGACCCAGCGCGGTCAGCTGAGGAAACTGGCCATCAAGTATGGGGTTGGTGGTCTGCTGCTCTTCCTGCTCATCTTCATCATCTGGTTCCCTCTCTTGCTCTTCTCCTTG
GCCGACACCGTTCTCTACGTCCCCAACCCCCCAGTGGACTGCACTGTGCAGATCAAGATTGGAGGCTATCAACCCATGTTCCATATGAGCGCCCAGCAGAACTTCCTCAAGCCCATGAATGAATACAACTACAGT
CGCCTGCTCAACATGTACAGCGATAAACCGGAAGCCTTGTCGGTTCTTGAGCAGTACAAAGCGTCAGACATTACCATCGTGCAAATAAACGGGGAATCAACCGCATTATGGGGCATTAGCCCTCCCAGCAAAGA
CGCCCTCTTCCACGATCTCATAAATAAAAACACCAATTTGAAAATATACTTCTCCATATCATTTAAAAGAGATCCTAAGCCGGGCCTCAACACAGATACGGTCACTGCGCAATTTCCTCATCCTTTAAATCCCTCAA
ATAAAACGAGGGAGGAGTTCGCATTGATGATCAATGGCAATAGAACGAAGCCGGTGAATGTCGTTGGCCTGTTTCCACGCTTCATGCGAGTGCCTCCCAGTGAAGCCCCTCAGGAATTGTTAGCCCTCACCA

>*Piezo*_ISH6 Plasmid_ID:267
TCCCTCGCTCTGGAAACGAGGAGGTGGCGCAGAAAGCGTGAAGAAGCCCACGAGAATAATGCCGCACCTGCTGTTGCCACTTCTGCACGGTCTACACCAGGGGGCACCAGGAGGCCCAGAATGAGGCCGAGG
GGGTCGGTGACCTCTGAACGTCAGTCTCTTGTTGACGAACAGGAAGACCAAGCAAACTATGACTCCATGGCAACCGCGGGAGATTCCTCTGCGTCAGTCAGAAGAAGGGCGCAGAGGCAGTCGAGTGTCATG
AGTGAGAGTGAGGCCCTAGGACCCAAGCAGAGGACGGCATTCCAGTCCATTCTGGCCTACTTCATGAACCAGAGTTATGCCGCTGCCTTGATTGTCATGATGGTTTGGAGTATCACCTACCACAGCTGGCTGAC
GTTTGTCTACCTGCTGTCTGCGTGCGTAATTTGGATGACTCCCCAGAAGAGGCAGGTGGCGCTAGTGGGCTCCCCCATCATTGTGTTCTACGCAGAGAGCCTGCTCATCATCCAGTACATCTACGGGCTGGCCC
TGACGGACGAGGAGCTTCCCACTGAGACGTCCTCCGGGTATAACTTCAAGGAGATTGGTCTCATGCGCTACAAGTACCCTTGCATCCCCCTTGCCATTCAGCGTGCCTACACGGGCGTATTCTGGCTGACATTG
AGGCAGTACATGCGAGAACTCCACCTGAAACGTCAGAATCCCAGTCAAGGCGACCTTCCCCTGGAATCAGTCAGAGACCCCGCAATACACTCCTCCTCGAGACTCCACAGAATGCGCACCACAGACGTAGTGGA
GGGCACAGATGGGGAAACCGTCAAAGCTATAGGGACGTGGATCTGGCTGCTCATGTGCAAGTACTGGATCCTGGTGTGTGCCTCCCTGTTCCTGGTGATGGCCGTGCAGGAGGCCGTCATCTACAGAATCAT
CTACATGGTCCTCTTCCTTTACTTCGTGCTCACCTTCAAGTTGCTGTACAACTTCTGGCGCTTCAGCATCGTGGCCTTCTGGTGGTTTGTGATCGTCTACTCGATGGTCGTGCTGTGCCTCATCTACACTTACCAG
TTTGACAACTTCAAGGA

>*Piezo*_ISH8 Plasmid_ID:268
GAAGACAACACAGTTTCTCAGACCCCTCGGACATGTCACCCTCAGAAAGGGGAGCCCGGATGCAGAGGAGTGGAGCCATGAGGGATAGCATCCGGGAGGAGGGCCAGGACGCCGCCCCCCTCATCACGACCC
CCGAAGATGACTATGGCTCTGCTGCATCCACCTTGCCTGGGACCAGCAAGGCAGCCGTACCTGTCACTGAACAGCCCTCTCCCTCGGACAAGACTGACCCAGCCAAAGATGGGATTGCTGAGGACGACTCTGGC
TCCGACGTTGCAGACCTGGAAGAGGATGTGCAACCTGATGGACTGTGGCAGCGCATTGTGACCAGCATCAAGCTCAGTTGGGCGGTGGTCATCAGCATCTTGGATTATGTGATAGAATCTCTCAACAACTGGT
CGCGGGAATACCGAGATATGGCGAGGGTGCTGGAGAAGGAGAAGGAGCAGGAGAAGCAGAGGATCAAGGAAGAAGGCGCGGACTCCCTTAAGAAACCTGAATCGAGAGACAGCTTGAGGAGGGAGAGGG
CCGCGGACCCAAGAGACTTGACTCTTCAGTTGGCTACCACTACAAGCGACGCCCTCGATGGGCCTGCTGAGGAGGTGCAGGATCCAGAGAAGGAGTTCAGGGACAGCACTCCGCGCATCTTCCAGCTGTTCAT
TGCTCTCTTTGACGTGATGATCTCCCACTCGGACTTGCTCTGCTACTTCGCCATGATCCTGAACCACCTGGTGTCCGCCTCCCTCCTCTCCCTGGTCTACCCCATCCTGGTGTTCCTCTGGGCGCTACTCTCCGTCC
CTCGGCCCAGCAAGACCTTCTGGGTCACCATGATCACCTACACTGAGGCAGAGGTGGTGGTCAAGTACCTGTTCCAGTTTGGTTTCTTCCCTTGGAATGTGGAGGTGTCAGCCAGCAGTGCGGCTGACCCATTT
TACCCCCCAAGAATAATCGGCATAGAACGGAAGGGCTACTATGCG

ASIC8016

>*ASIC8016*_ISH1 Plasmid_ID:306

TCACTCTTTGCAATTACAACCAGTTTAGGAAATCTAAGGCCGAAGAGAAACCCGGAGTTATGGGTGTGTTACGCGATTTATTTTCCGTGGATGCCATAATAGGACGACAAAAGAGCAAATTGCGGATGTCGGC
ACAGTCATTAGTTTCGCAGGATATTAACGTCATCAATGAACTTGTCGCTTTAGCTCACACTTTAAAGGACATGGTACGCCATTGTCAGTGGTGTGGTAAAACTATTCCTTGTTCCGATTACTACTCACAGACACTA
ACGGAAGTGGGAGTTTGTTATACTTTCAATTCACACGATGTTCGTAAGAAAATTCCACCTCTAGAAATCCACGACGCGGGCACTGACTGTGGGCTGAAGCTCCGTCTCAATATCCAGGAATACGAGTATTACTA
TGGTGAAGGCAGCTCAGCGGGCATTCAGATGTTGATCCATGACCCTGAGGACGTGCCCCTTGTCAGACAGCATGGTTTCTCTGCTGCTCCAGGAACTTCAGTTATGGCAGCCATCAGACTAACCAAGGAAACCT
ACGCTCAGTCGCCTTATGGAGATTGCGTGGACGCGAAGCATCCTTCTTTTGAAAATCCTCTCCG

ENaC417306

>*ENaC417306_ISH1* Plasmid_ID:298

GACAGGATCGACAGGACTTAATCTCCCCTCAATACCACTTGGCCGGAAGAAGAGACAGATTCTTTCGGGGAATACTCTTTCGCCGAGGGTTCCTCCAGGGGAGTCAGGAACAAATAATCTGCTCAGCATGGAG
CTTAAAGGTGACCATGACTATGAGACACTATTCGCAAGCTCAAAGACATCTGACTTTAGTGACGTCGTCAGTATGCTGCGCCCGAGCATGCCGGACTTGGACAAATATGGGCACCAGTTTGACGAATTTGTCCT
AATGTGTACGTTTGATGGGGAAAACTGCACAAGCGATGATTTCATCCGTGTATACAACGAGGTCTATGGAAATTGCTTCACATTCAACAGACAGACGAACGGCTCAACTGTAAGAACCACAACTAAGTTTGGCT
CTCAGTACGGTCTAAAACTTTCCTTGAATATAGAAAAGGATGAATATTTGGGACTGTTTTCGCATCAGTATGGAGCTAGAGTTACAGTCCATTCGAACAACATAACCGCATTTCCTCAAGATAACTCCGTTTCCAT
TCCTGTTGGATACTACGCATCTGTGGCAGTAGACACTGAAGTGATGGACCCCAAAGAAAGACCCTATGAAACGAACTGCACCCATGGCAAATCGGCCGACTTGTTCTACGAAGGACTTTATTCAACCGATAACT
GCCTGAACTCTTGTCTGAGGAAGATGGTC

>*ENaC417306*_ISH2 Plasmid_ID:297

TGTGTTCCTGTGTGGAGACTGTCCTGGTCACTGGCAACGAAAGCAGATGCTCGATATCCAACACCACTGAGGTTGCCTGTCGACTGGAAGTTTACGCCAATTTTAGGGACAACGCTTACTCATGTCAGCATGAA
TGCAACGAACCTTGTTTAGAAGTTACCTACCGTACCTCGGTATCTTACGCAAATTGGCCGACGGATGGTTTCAACGCACTTGTTGGGAGGCAGCTGGTCGGTAAAATGGCAAGGAATGAAATTAAGGAGTACA
TTGAGAATAATATATTGCGAGTGAACATATTCTTCC

ENaC2547

>*ENaC2547*_ISH1 Plasmid_ID:260

GAAGTCTATGGGGAATACCACAACCTCAGTCTGTCTGAGCTGAGCGTCGACCCTGACAACTTCATTCTGAATTGCATGTACGAAAGGGTCCATTGTGACATGAGGAATGACATTGAAATGACATCTATATTCAA
CTGGCTGTATGGTGGATGCATAGTGTACATAATCAAACAAAACACAACAAGAACAGGTCCCAGATTTGGGCTAGAACTCACTCTAAACATTGAACAAAGTGATTACATGGACCTTACCACACAAGCTGGGGCAC
TGGTGCTGGTTCACCCTCCTGAAGAAATGCCTTTTCCTGAAGACGATGGAATTAGTATACCCCCTGGTCAAGCTGCCCTCATTGGAGTCAAAGTGAAAAGGACTGAAAGACTCGGGGGTAAATTTGGCGATTG
TGCTCATGCGAATAAACTGCCAAACTTCACCAACTACTTCCAAGCGGAGAATCCATGGACAAACTACAGTCTCAAGGCTTGCCAAAGATCATGCTTCCAGTGGAACGTTCGTCAAATGTGTCAATGTCTTGATGT
TCGATTTCCGCCATTTGACGAATACAAAAACTTTACTTCTTGCGTAAACAAGTATTTTTCGACGTTCAATTTTTCACAAGGTTTTCCAGAATCTGTGGGTTGTTTGGACGAAGTGACCAAAAAATTTAACGCACAT
CAACTTGGCTGCGACAAGCTATGTCATGTTCCATGCAAACAAACGTATTATGAGCCATTTGTGTCATATGCAGCATGGCCTTCAGATGCAGCTTTGAATGTCACCCTTGAGAACTTGAAGGGCAGCTCCGCAAG
TGCGGATGAAATCCTGGAAGTTGGTCATGACAATTTCCGCAAGAATTTTCTGCAACTCACAGTGTATTTTCAAGATCTCAATTTTGAATTCATAACTGAGAAAGAAGCAGTTGAGCTAACTCAGATGTTGTCTGA
TCTGGGAGGCAACACAGGCATGTATGTCGGAGCTTCGTTATTCACAATGATCGAGTTT

ENaC415688

>*ENaC415688*_ISH1 Plasmid_ID:263

TGCTGATCTTTGGTGCAGACAGTCCTGATGTTTTCCTTATCATCAACAAGGTTATAAAGGGCAACTCCCTCCATTGGAATTTCCGGATCCCCACAGTTACACTCCTTGAAGAGATGAGCTTGTAAACATGTTTTGT
AACAGGCGTGGTGAGTATAGCGAACAGGATATAATTCTGAATAAGCACTGTAGTCGTAAGCCTCCTTTCGTGAAACATACAAGCAGTCACTTGGGTATGGACTCTGCATTCTCTGAATTTTTTCAACAACAATG
GCTATATTTGACTGGGCATCAGGCGTAATAAGAACTCCATCCTCTTCTGGAAAAGGCATCACATCCTGGTCATGTATTGCTATGACAACGCCAGCAGATTCCGTCAATTCCGGAATATACTCGTCTTGGTCAATG
AAGAGTTCCAGCTTCAGTCCGTAAAATGGACCAGGCTGGTGACTGATCTGGACATGTTGATTTTTCCAATTTGAGTTAAACATGTAACAATTACCAAACTTCGGATTGAAAAAATGTGTGAAATCTTCAGAGTAA
CAACGCTTTTCATTATAAGAGCAGTCCAGCAGCATTTCCTCAATCTGATGACCCTGTGATGCTCTCGTAGACTGATTCAACTGACTCCAAGAATTCAAAAACTTACGCCGAAGTTCAGCTCTTCCTTCTGAGTTCTC
CCAGTGAGACATGTCTGGGTCTCCTCCCGCAGTATAAGTCAGAAGATCCAATGCTTTCAAACCTTTCATTAAACCAAGCTGAGATTTTCTCACCATATTATAATTACAGATGGTGATAGTAGGGAAGACCAAAGA
ATTTTCATAAGTCACATTAACATTGATGCCAACAGGGTAGCTAAAATACGTTTCGAACACCTCACATACTTGCTGGAGGGTCAAAGCTGCTCCCGTAAG

>*ENaC415688*_ISH2 Plasmid_ID:312

TGGAGCTACTAATGAACTCCAAAAGAATATTCCACAGTGATTTTGGTCCAGACCAGATATAAAGTCAAAATATTACGGGAAGGCATTAGCGGCCTGAATCCGCCACTGAACAGTTATATGTAAGCAATACTGCT
ATTATTGTCTTTCCTCCCTCTCATTTTCTAAATCATTCAAGATTCTACTTTCTGGACTTGATTTAGCATTTCGATTGTCGTCCTTTTTTCTTGATTTTTGACAAGTAACGAGGAATAAATCGAACAAGTATTCAATCA
ATTCAAATGCTGCAATGATGGAGATCCCCAGGTAGAAACCGATGGATCCCCCAAGGTCTGACAACAAACTGCTCACTGGATATGATGCATATTCCTCAATGGCCTTGAAATTAAATTCGCTGTAGTAAATCAATA
GTTTCACTAAATTAGTCCCTGGATTGGTTCCACTAGGGAATGGGACTTCAGAACAGATTTCATCTGAGGGTTCCTTGTTGCTAACGACTGTAATACAAATCCTCGTAATTTGCATTTGGCCATTGGAGGGTGCTG
ACGGTCTTCTCAAAAACAGTATCTATGCATGCAACAGAACAATTGCATTTGAGTTCACCATTCAGATA

EST clones

Gene name	Jékely Lab EST library ID
ASIC8016	B-50_I03 B-35_N23
ENaC2547	B-224_A17
PKD1-1	B27_E14
PKD1-2/LOV-1	B69_A09

Wildtype (WT) and predicted mutant sequences generated with CRISPR

NOMPC

NOMPC_WT nucleotide sequence (1st and 2nd exon boundaries marked with |)

AATGAACGTTACATTCCAACACTAAAGCACTTCACTATAAAAGGAAACCTCTACTGAACTGAAAAGATAACTTGCTGGAT
TTTCAAAAAGTTGTTTAAATAGTTATATGAACTTTTCTCTACACTAACTGTTGGACATATTATTACTGACATTGTGTGAT
AGTGGAATAACTTGTTTCCATTGTTTTTGGAGAATTCAGGGACATAGCCATGACTTTCACTCGGACATCATTTTAGAAGG

ATTTACAGTCTAGTTCTAAGGCTCAGAATATAAACTTTTATGACAATGAAGATTAAAACGGATTAGCGATGTCTTCCACT First predicted codon
GCAAAAGAACAAAATGAAACAGACAAGACTCCTCCAGATAAAGCTGTTGACGAGAAAAAGGAGGAAGATAAAAAAGCTGA
GGAAGCTGCAAAACCCACACCATCGGACAAAGAGTCTAAACCGGAGAACGTAGACAACAAAGCCTCAGCCAGCGGCCCAAC
CCCAGCCGCCTAGACTAGAAAAAGAGCTTCCATTGGAACCGAAGGAAATACCAAAGACAGAAACAGCAGCACAAGACACT
AATCAAAATAACAACAAAGATACAGGCACTTCCAAATCTACAGCACCGCCAGCAGCATTAGACGAATCTTCCAAAAAGAA
AGAACCAGAAAACAAAGATGCACCGACAGAGCTTCCAAAACAAGATGAAACCAAAGAAGGTGGACTCACCAACAGAAAAAG
GAACGTCTGACACCAAACAGCCAGAAAAGAAAGATGGAGACCCTAAGCCAGTAGAAAACAAGGAGCAACAACCTACTCCG
CCAGCGACAGAAAAGCCGAAATCTCCTCCGAAAGCTTCACCAGCTGAAGATAAGGATAAAAAGACTCCACCAACTGGTCTC
AGATAAATCAAAACAACCAGAAGCAGATAAATCAAAACCTCCAGCAGATAATGCAGTAAAATCAACAGAGGCTGAAAAAG
TTCAACAAGCAACTGCTCCAGAAAAGAAGGTTGAAGCAAAAGATGCAGAAAAGGATAAAAAGACTCCACCAACTGGTCTC
CAACCTACTCCGCCGGCTATACCACCAAGAAAAAATGTCAGCCCGAATCCAGGAAGGAAAATTCCTGTTCCATCTCACCC

GCAACCACCGAAGAAGACCCCTTTGAGTACCCAACCAGAAAGAAGAACCCCACCGACTCCTCCGACAAACAGAAAGACTC unstructured region
CTCCAACCTCCAGAAAAGAACCTCCGACACCTCCCGCTTCCAGGCGTCAGCAACCGCCGACGCCCCCAACCAGAGCGTAAT
CCGACGTCCTCCAATGCCGGTAAACGATCTTCCTCTAGTTCACGACAGTCTTCCGCAGAAAAGCAATCGAAACAAAATGG
CAAAGCTAATGAGGATAAAACTGACAACAAAACTGAACCTAAAGCGGGCAATAAGCCTGATAACAAAACGGTTAATAAGA
TTGATAACAAAACGGAAAACAAGACTGGAGTTGCAGCCACAGGACAAAACTCGAACAATAACACAAATGGTCCAAAACCG
ACAATTCCGGCGAACGACAAAAATGTAAAGACTGCAACTGACAAAAGGCCTCCAGTTTCTGCAACAGACAAATCGAATGA
TGCTGCTAATGCGAAGAATAATCCTCAAGATAAAACCAAGATTCCAAAATTGACAGAAGGAACAGCTCCAACAACAACTG
ATGCTAAGTCAACAAAAGGTGATCCCAAGAATGCCTCCAATGTTAAAGACTCCCATCAACCTAAACCTGATGCTAAACAT
GATAAAGATCCTAAACATGATGCTAAACAAAATGACGCTAAACATGATGACAACAAAGCTTTGCCGCACAAAGATGACAA
GAACAAAAAATCACACGATGCGACCCAAAAATCTTCCAAAGGCGAAAACGAGGGAAATATCAGGAGTGACGCCGAGTTAA
ACGGTGCCCCCTTAGACCCTGAAAATCCTGAAAAACAAGCTCCGCCATCCAGTCGTCTGCTTAATTTGGCTCAAAAGGA

GAATGGTCAATTCTGGAACAGTCACTTCGAACTCTGGACAAAGGAAATCCTCAAATATTTGTTGTGG | ACAAGGACACTGG
AGTTACACCCCTGATGCTGGCTGTGAGGGAGAACAGACTGGTCGTTGCTGAGAGGCTAATAGATTTGGGAGCACCTATCAA PAM+target seq (alt. ATG)

CCATCAGGCAGCAG | ACAATCGCACAGCCCTGCACTATGCTGCATCCTACTCAAGAGACGACATTGTCAAGTTGTTGCTG
AACAGAAAGGCTGATGCCACAATTATTGGCGGGCCAAAGGACCAGCTTGCTCTGCACATGGCCTGTGGTAGGGCTAGTGG
AGTTGTGCCGACTGTTCAGTTACTTTTGAGGTCAACAGGGAAAGAAGCCCGGCTTGCTGAAGACAAGGAAGGGTCAATTC
CACTGTTTTTGGCAGCTGAAGTTGGCAACACAGCAGTATGCCGGGAACTGCTGCAGCAAATGACAGAACAGCAGATTATG
GCACAAAAGAAGGATAATGGAGACTCAGTATTACACTTGGCCTGTCATCGGAAAGACATTGACATGACAAAATTGTTTAT
AGAAAGCGGGGCTGTGATAGACAGACAAAATACCGAAGGCCAGACTCCCCTGCACATAGCAGCCTGGGAGGGGGATGAAT
CGATCGTCAAGTACCTTCACCAAGTCAAGGCCAACCCCAACATTATAGATACGGAGGACAGGTCACCTTTACACATAGCA
TCAGAGAGAGGTCACACAATAATAGTGGAGATTTTAGTCGACAAATTCAGAGCAAATGTCGCGGCAAGGACGAAAGATGG
GAGCACTCTGATGCACATAGCATCTCAATATGGCCATCCTGACACAGCTCTGGCTTTCTTGAAGAAGGGAGTACCTATGC
TGATGCCTAATAAGTCTGGGGCTATCTGCCTCCATGCAGCTGCTATGAAAGGCCACACTTCTGTGATCAAGGCTCTTTTG
ACCAAAGGGGCCCCAGTGGATGCCATGACTAAAGATGGATACACTGCTCTGCATCTTTCTGTACAGCATTCCAAACCTTT
GGCAGTCCAGACATTGTTGGGGTATGGAGCAGCAGTGGAACTCAAGGGAGGACAGTCAAAAGAGACACCCTTGCACATCG
CAGCCAGGACGGAGGATGGAGAGAAGTGTGCTGAGATGCTTGTGAAAAGTGGGGCTGATGTGAACGCAACTAAAGAAAAT
GGAGAAACAGCGTTGCACATTGCTGGCCGATACGGTCACCTGCAGATGATCTGTGCTTTACTGGCAGAAGGCGCAGATCC
AACCTGGCAATCTAAGAGTGGAGAAACGCCTCTTCACATCGCAGTGCGCTACTGCCACTGGGAAGTAGCAGATGAATTAT

TGCGGTTTGTGAAGAAAGAAAGGTCACCGATTGACGCCACAATGTTGGTTAATGCACCAAACATTGAGGGAGAAACTGCT ANK repeats
GTCCATTATGCTGGAGAATTAACAAAGAAACAGGCACACCATGAGTTTGAGGACACTGATATGATCAGCTTGTTACTCCA
GAATGGGGGAGATACTAACACCCATACTAAATTGACCCAAGAAACACCGATCCACTACTGTGCCAGGGCAGGCAATGCAG
ACGTCATGCTGGAGATAGTGAAGCACCTTGGGCCCAACCAGGTGCAGGTGGCTGTCAACAAGCAGGCTAAGAATGGCTGG
TCACCTCTTCTTGTGGCAAGTGAACAAGGTCATTTGGACATCTTGAAAATTCTTCTTCAGCACCATGCAAGAGTGGATGT
CTTTGATGAGCATGGCAAAGCTGCATTACATTTAGCGGCAGAAAACGGTCACGAGGAAGTCGCGGACGTTCTTTTGCAAC
ACAAAGCTTTCGTCAATGCCAAGTCCAAACTTGGTTTGACTCCGTTGCATCTCGCTGCCCAATCAGGCTACAACAAACTG
GCAAGGCTACTCATAGAGAAGCATGGGGCCACCAATGATGCTTTGTCTTTGGCCAAGAAGACCCCACTCCACCTTGCAGC
CCAGACTGGTCAGATTGAAGTGTGTAGTACCTTGCTGAAGATGAAAGCAGATGCCAACGCCAAGGATGCTCACAGTCAAA
CGCCTCTCCATTTAGCAGCAGAAAACGACCATTCAGATGTGGTCAAACTTTTCCTGAAACACAGACCAGAACTGGTCTCA
ATGGCTAACACTGATGAGAATGGCATGACATGTGCCCACATAGCTGCCTCAAAAGGAAGTGTGGCTGTCATCAAAGAGTT
GATGAGGTCTAATAAGTCAGTGATAGCAACTGCAAGAACCAAGACAACAGATTCCACAACATTACATCTAGCTGCAGCTG
GGGGCCATGCTGTTGTCATCAAATACCTGCTGGAAGCTGGAGCCTCTGCTACTGATGAGAATGCTGAGGGAATGACGGCT
ATCCATCTGGCAGCCAAGAATGGCCATGTGAACGTCCTGGAAGCCCTCAAAGAAAATGTGTCATTCAGGGTGACTTCTAC
TAAGACTGGTCTAACAGCATTGCACGTAGCTGCCCAGCATGGTCAGATTGACTTCGTGAGAGAAATGCTGACCAAAGTTC
CAGCCACTGTCCAGAGCGAACCTCCCCATACAGATGGACCTATGAGGGAACTGGGAACTGCTGAGTTTGGCCTAACCCCC
CTCCATCTTGCAGCCCAGTCTGGACATGAAGGCCTAGTCAGATTACTGCTTAATTCTCCGGGAGTCCAGGCAGATGCTGC
CACTGCTGTTCACGGTGTGATTCCACTTCATTTGGCAGCCCAGAGTGGACACACAGCAGTGGTCAGCTTGCTTCTCAGCA
AATCCACCAAACAGCTTCACATACAAGACAAGAGAGGGCGCACTGGTCTTCACCTGGCGGCCTCTAACGGACATTTGGAC
ATGCTGTCTTTGCTGCTTGGTCAAGGGGCTGATATCAATGGATATGATAAAAATGGCTGGACTCCTCTGCATTACGCAGC
CAAAGCAGGGCACTTGGAAGCCGTGAAACTCCTCTGCGAAACCGGAGCCAACCCGACCAATGAAACAAAAGATGGAAAAG
TTCCTGTGTGTTTCGCCGCGGGAGCAAACCATGAAGAAGTTGTCTCATATCTTCTACAGACAGACCACAACACTTACAAC
TTGATGGAGGACAGAAAGTGGGTATTTGACCTCATGGCACTTGGCAAGATGAACGACAACAAATCGATCCAGGAATTTAT
TCTGGTATCTCCGGCCCCGGTTGACACTGCTGTCAAGTTGTCGCGTGCCTTCAGGAATCTCTCAACAAAGGAAAAGAAA
GAGCGAGGGATTTACTGGCCGCTGGTGACTTTTGTGAAGTAATGGCGACAGAGTTGAATGCCATTGCAGCCGGTATCAAC
GGTCCGGCTATATTACTCAAATCAGTGGACAATCGAGGAGTGCCTCTTCTGGATATTCTTATTGAATGCGAACAGAAAGA
AGTCGTTGCTGATGCCGTTGTCCAGCGGTATCTGTCAGAAGTGTGGGCGGGTTACCTGAGTCACTGGTCTTCTCCCAAAA
TGATGCTGCTCTTCTGTGTCTTACTGTTTGTGCCTCCTGTCTGGCTGGCTTTCTCTCTTCCTAACAAACACATTCGCTTC
AACAAAGTGCCTATGATCAAGTTCATGCTCTACTTAGTATCTCACATCTACCTTATCATGCTTTTCACTCTGACTATCGT
GTATCCGATCTATCCTCTCTGGGAATCTGGAAACTTGATTCCGCACTGGCACGAGTGGCTGCTGCTTTCCTGGTTTTCCG
GCTTGTTGGTCAACCAGCTGACGAATCCAGAAGATCGTCAAGGTTTGGGATGGATCAAGGTCATCGTGATCGCCATCAGT
AACTTCGGAGTGACCACTCATCTCATCGCCTTCTCATTCTCTGGCAACGATCGCCTGTACTGCTTGTATGCTCGTAACCA

ATTCTTCGGTCTTGCCCTTATGCTGTGTTTCGTCCAACTCCTGGACTTCTTGTCCTTTCATCAACTGTTCGGACCGTGGG TM region
CCATCATCATCCGTGATCTCATGAAGGATCTGACCCGTTTCCTCGTAATCTTAGTTATCTTCATGGCTGGCTTTACCTTG
GAACTAGCTACTATTTACCAGCCTGTGTTAGCTCCTTTGATACCAGAGCCTGACAAGGGAACAGGGGATGGTTCTGGAGG
AGCAGGAAAAGAGACTCCCATCAGTACTTTTGAACTTCTCTCTTCTTTGCGCTATTCGGTCTCGTTGACCCAGAGAACTTAC
CCCTTGTCAACAGGAGTCCCCTGTGGAGTGTTACCCTTGCTAAGGCTGTTTTTGGATGTTACTTGACACTGACAATTATC
GTGCTTATCAATCTCCTCATTGCTATGATGTCCGACACTTATCAAAGAATTCAGGCGCAGTCTGACACAGAATGGAAGTT
CGGAAGAGCGAAGTTGTTCAGGAATATGAACAAAACTTCTGCCACTCCATCTCCATTGAATCTCTTCACAAAACTTTATG
TGTATATAGTCATCCTCATTAAGCATAAAAG|CAAAGTATGTAGTCTCACAGATGTCCAGCAGTACATAGACGAAGAAGAG
GAGCTGGATCAGGACTCTGAGAATGCAGCAGCTGGGCAAGGGAATTGGAGGAATATGATGAGGAGAAACACCCAGGTTGC
CCCTGAAG|ATTTCATGCAAGCTGCCTTGCGGAGCAACGGTCCGACTCCAGTGGAAGAAGTCGTCGATTGGAAAGTTGTTG
CGGACAGATTTCTTTATCTTCGAGGTCTCGCTGACGGAAATAATATGGATAGCGACAAAGAAGATGATGATGATGAC

AAGTTTAAGAGCAATGAAAACGCTGCTGAACAGCAATGATTGA Stop codon
AACGGAATAATTGGGTTGCAATTTGTGTAAATTGTTATTATACATTATGTATCTAGCAGCAAAAACTATGCATTGTATCA
CAAAATTAGTGTATAACGATAAGGGTGTTATCCATAAAGAATTTTATCACAAAAAAATTTTGACACAGATGAATCATGTT
TTCACCTATAAATTTATTTATCTAAAATGATGCTGCATTTGGAAAAACTTGAAGTATTTCGCGAAGAGTGATTTTTACAA
TAGTACTTCATTTATCCTAGAATATTTATACTACCATGACAAATATGGCATTAGCCTACAAGCTTGTCGACGACTGCCTT
AAATGACATTTTTACCTGCTGATATAATATTATTGGTTCCATTCCGAAGTGAAGTATTGAAGAATAGTTTATATAGTATTT
TGGAATTAAATGCGATGCAAATAGAAATTGGTCTCATCAGGTACGTGTAATCATTTAAGTCCTAGTTCGAAGGAAGGTAC
CAAAGTTTAAATATGCAAATATTTGGAATTTCACACGCCGGTAACCAAGTGCTTTATAATCTTTCAGCTTTTTCATATAT
GAAACCTCTCAGCATTCACTGTCAGTCAAAAACTGGCCGCCATAGGTCCGTTATTAGCAGAGTAGGATCGGTTGGCTGAT
GCGGGCCATCCAACATGCTGATTGTGAAGTTGTATGTACATTCAATGTTTCTTCCCATAAATTGACATATCCATAATGGC
GACAAGTGCTCTCTTTACAAAATACCCACAATGAATGATTTTTATAAATAAAAAAGCAGCCTTATTGGGGCAACCATTAG
ATATTTAGGGATTTGAAGCTTTCTGCGACAACGAGATATTGTTTACTTCACGAAAATGCTGGAGACAATTATTTATTTG
AGCATGTGACATCCTAAAATGTTAATTTCTGAATCAAATAATAATCTATAAAGAAATATCCTCGCATGAAGTTATTTTTT
CAGGACTGACTGGAAAACATTCATTTATTTCACAAAAATGGGCAAGACATTTTTTATTAAAAAAAACTCCAAGCGTCCCT

GGAGATCTGGAGTTTGTCCCCTAAAAGGAAGGTTCGCTCTCTAACGAGAGATCATCCTGAGACTGAAAAGGAAGACAACT
TGTTTGTTGCAGGAACTTTTGAAAGCAGATTGCTTGTCGAAGGGATTACCGTCCACTGAAATACATAAAAGCAATATTAT
GAAGCTGCATGGCAGAACTTCTCAAATATTAACTTTTTTCGTAACACATACAATTCTTGAAGTATATGACTCCGTCCGTG
ATATTGCTATATTATAAGCTTAAATTAGTTGTGATATATTTACTGTTGATAATCCTTGTAAATTTATGATTATTATTACT
TTTTATTGACATTTATGTGTACAAAAACTGCAGATAGAACACCGTTATTGTTATTATTGTATAAAATAAACTGCTTCTCT
ACTGGAATATATATATACAATCATGTAAGTTACTGAATTTAATAGAAATTTCAATTAATAAAAA

NOMPC aminoacid sequences

- **NOMPC_WT** (colored according to nucleotide sequence)

MSSTAKEQNETDKTPPDKAVDEKKEEDKKAEEAAKPTPSDKESKPENVDNKASASAQPQPPRLEKELP
LEPKEIPKTETAAQDTNQNNNKDTGTSKSTAPPAALDESSKKKEPENKDAPTELPKQDETKKVDSPTE
KGTSDTKQPEKKDGDPKPVENKEQQPTPPATEKPKSPPKASPAEDKDKKSPPTPEDKSKQPEADKSKP
PADNAVKSTEAEKVQQATAPEKKVEAKDAEKDKKTPPTGLQPTPPAIPPRKNVSPNPGRKIPVPSHPQP
PKKTPLSTQPERRTPPTPPTNRKTPPTSRKEPPTPPASRRQQPPTPPTRRNPTSSNAGKRSSSSSRQSSAEKQS
KQNGKANEDKTDNKTEPKAGNKPDNKTVNKIDNKTENKTGVAATGQNSNNNTNGPKPTIPANDKN
VKTATDKRPPVSATDKSNDAANAKNNPQDKTKIPKLTEGTAPTTTDAKSTKGDPKNASNVKDSHQPK
PDAKHDKDPKHDAKQNDAKHDDNKALPHKDDKNKKSHDATQKSSKGENEGNIRSDAELNGAPLDP
ENPEKQAPPSSRLLNLAQKGEWSILEQSLRTLDKGNPQIFVVDKDTGVTPLMLAVRENRLVVAERLIDLG
APINHQAADNRTALHYAASYSRDDIVKLLLNRKADATIIGGPKDQLALHMACGRASGVVPTVQLLLRST
GKEARLAEDKEGSIPLFLAAEVGNTAVCRELLQQMTEQQIMAQKKDNGDSVLHLACHRKDIDMTKLFI
ESGAVIDRQNTEGQTPLHIAAWEGDESIVKYLHQVKANPNIIDTEDRSPLHIASERGHTIIVEILVDKFRA
NVAARTKDGSTLMHIASQYGHPDTALAFLKKGVPMLMPNKSGAICLHAAAMKGHTSVIKALLTKGAPV
DAMTKDGYTALHLSVQHSKPLAVQTLLGYGAAVELKGGQSKETPLHIAARTEDGEKCAEMLVKSGAD
VNATKENGETALHIAGRYGHLQMICALLAEGADPTWQSKSGETPLHIAVRYCHWEVADELLRFVKKER
SPIDATMLVNAPNIEGETAVHYAGELTKKQAHHEFEDTDMISLLLQNGGDTNTHTKLTQETPIHYCAR
AGNADVMLEIVKHLGPNQVQVAVNKQAKNGWSPLLVASEQGHLDILKILLQHHARVDVFDEHGKAA
LHLAAENGHEEVADVLLQHKAFVNAKSKLGLTPLHLAAQSGYNKLARLLIEKHGATNDALSLAKKTPL
HLAAQTGQIEVCSTLLKMKADANAKDAHSQTPLHLAAENDHSDVVKLFLKHRPELVSMANTDENGMT
CAHIAASKGSVAVIKELMRSNKSVIATARTKTTDSTTLHLAAAGGHAVVIKYLLEAGASATDENAEGMT
AIHLAAKNGHVNVLEALKENVSFRVTSTKTGLTALHVAAQHGQIDFVREMLTKVPATVQSEPPHTDGP
MRELGTAEFGLTPLHLAAQSGHEGLVRLLLNSPGVQADAATAVHGVIPLHLAAQSGHTAVVSLLLSKST
KQLHIQDKRGRTGLHLAASNGHLDMLSLLLGQGADINGYDKNGWTPLHYAAKAGHLEAVKLLCETG
ANPTNETKDGKVPVCFAAGANHEEVVSYLLQTDHNTYNLMEDRKWVFDLMALGKMNDNKSIQEFIL
VSPAPVDTAVKLSRAFRNLSTKEKERARDLLAAGDFCEVMATELNAIAAGINGPAILLKSVDNRGVPLLD
ILIECEQKEVVADAVVQRYLSEVWAGYLSHWSSPKMMLLFCVLLFVPPVWLAFSLPNKHIRFNKVPMIKF
MLYLVSHIYLIMLFTLTIVYPIYPLWESGNLIPHWHEWLLLSWFSGLLVNQLTNPEDRQGLGWIKVIVIAIS
NFGVTTHLIAFSFSGNDRLYCLYARNQFFGLALMLCFVQLLDFLSFHQLFGPWAIIIRDLMKDLTRFLVIL
VIFMAGFTLELATIYQPVLAPLIPEPDKGTGDGSGGAGKETPISTFELLFFALFGLVDPENLPLVNRSPLWS
VTLAKAVFGCYLTLTIIVLINLLIAMMSDTYQRIQAQSDTEWKFGRAKLFRNMNKTSATPSPLNLFTKLYV
YIVILIKHKSKVCSLTDVQQYIDEEEELDQDSENAAAGQGNWRNMMRRNTQVAPEDFMQAALRSNGPT
PVEEVVDWKVVADRFLYLRGLADGNNMDSDKEDDDDDDKFKSNENAAEQQ

- **NOMPC_Δ1_Version 1** (new residues and colors according to nucleotide sequence)

MSSTAKEQNETDKTPPDKAVDEKKEEDKKAEEAAKPTPSDKESKPENVDNKASASAQPQPPRLEKELP
LEPKEIPKTETAAQDTNQNNNKDTGTSKSTAPPAALDESSKKKEPENKDAPTELPKQDETKKVDSPTE
KGTSDTKQPEKKDGDPKPVENKEQQPTPPATEKPKSPPKASPAEDKDKKSPPTPEDKSKQPEADKSKP
PADNAVKSTEAEKVQQATAPEKKVEAKDAEKDKKTPPTGLQPTPPAIPPRKNVSPNPGRKIPVPSHPQP
PKKTPLSTQPERRTPPTPPTNRKTPPTSRKEPPTPPASRRQQPPTPPTRRNPTSSNAGKRSSSSSRQSSAEKQS
KQNGKANEDKTDNKTEPKAGNKPDNKTVNKIDNKTENKTGVAATGQNSNNNTNGPKPTIPANDKN
VKTATDKRPPVSATDKSNDAANAKNNPQDKTKIPKLTEGTAPTTTDAKSTKGDPKNASNVKDSHQPK
PDAKHDKDPKHDAKQNDAKHDDNKALPHKDDKNKKSHDATQKSSKGENEGNIRSDAELNGAPLDP
ENPEKQAPPSSRLLNLAQKGEWSILEQSLRTLDKGNPQIFVVDKDTGVTPLIWLSTOP

- **NOMPC_Δ1_Version 2** (new residues and colors according to nucleotide sequence)

MVNSGTVTSNSGQRKSSNICCGQGHWSYTPDLAVRENRLVVAERLIDLGAPINHQAADNRTALHYAASY
SRDDIVKLLLNRKADATIIGGPKDQLALHMACGRASGVVPTVQLLLRSTGKEARLAEDKEGSIPLFLAAE
VGNTAVCRELLQQMTEQQIMAQKKDNGDSVLHLACHRKDIDMTKLFIESGAVIDRQNTEGQTPLHIAA
WEGDESIVKYLHQVKANPNIIDTEDRSPLHIASERGHTIIVEILVDKFRANVAARTKDGSTLMHIASQYG
HPDTALAFLKKGVPMLMPNKSGAICLHAAAMKGHTSVIKALLTKGAPVDAMTKDGYTALHLSVQHSKP
LAVQTLLGYGAAVELKGGQSKETPLHIAARTEDGEKCAEMLVKSGADVNATKENGETALHIAGRYGH

LQMICALLAEGADPTWQSKSGETPLHIAVRYCHWEVADELLRFVKKERSPIDATMLVNAPNIEGETAVH
YAGELTKKQAHHEFEDTDMISLLLQNGGDTNTHTKLTQETPIHYCARAGNADVMLEIVKHLGPNQVQ
VAVNKQAKNGWSPLLVASEQGHLDILKILLQHHARVDVFDEHGKAALHLAAENGHEEVADVLLQHK
AFVNAKSKLGLTPLHLAAQSGYNKLARLLIEKHGATNDALSLAKKTPLHLAAQTGQIEVCSTLLKMKAD
ANAKDAHSQTPLHLAAENDHSDVVKLFLKHRPELVSMANTDENGMTCAHIAASKGSVAVIKELMRSNK
SVIATARTKTTDSTTLHLAAAGGHAVVIKYLLEAGASATDENAEGMTAIHLAAKNGHVNVLEALKENVS
FRVTSTKTGLTALHVAAQHGQIDFVREMLTKVPATVQSEPPHTDGPMRELGTAEFGLTPLHLAAQSGHE
GLVRLLLNSPGVQADAATAVHGVIPLHLAAQSGHTAVVSLLLSKSTKQLHIQDKRGRTGLHLAASNGHL
DMLSLLLGQGADINGYDKNGWTPLHYAAKAGHLEAVKLLCETGANPTNETKDGKVPVCFAAGANHE
EVVSYLLQTDHNTYNLMEDRKWVFDLMALGKMNDNKSIQEFILVSPAPVDTAVKLSRAFRNLSTKEKE
RARDLLAAGDFCEVMATELNAIAAGINGPAILLKSVDNRGVPLLDILIECEQKEVVADAVVQRYLSEVW
AGYLSHWSSPKMMLLFCVLLFVPPVWLAFSLPNKHIRFNKVPMIKFMLYLVSHIYLIMLFTLTIVYPIYPLW
ESGNLIPHWHEWLLLSWFSGLLVNQLTNPEDRQGLGWIKVIVIAISNFGVTTHLIAFSFSGNDRLYCLYAR
NQFFGLALMLCFVQLLDFLSFHQLFGPWAIIIRDLMKDLTRFLVILVIFMAGFTLELATIYQPVLAPLIPEPD
KGTGDGSGGAGKETPISTFELLFFALFGLVDPENLPLVNRSPLWSVTLAKAVFGCYLTLTIIVLINLLIAMM
SDTYQRIQAQSDTEWKFGRAKLFRNMNKTSATPSPLNLFTKLYVYIVILIKHKSKVCSLTDVQQYIDEEE
ELDQDSENAAAGQGNWRNMMRRNTQVAPEDFMQAALRSNGPTPVEEVVDWKVVADRFLYLRGLAD
GNNMDSDKEDDDDDDKFKSNENAAEQQ

- **NOMPC_Δ2_Version 1** (new residues)

MSSTAKEQNETDKTPPDKAVDEKKEEDKKAEEAAKPTPSDKESKPENVDNKASASAQPQPPRLEKELP
LEPKEIPKTETAAQDTNQNNNKDTGTSKSTAPPAALDESSKKKEPENKDAPTELPKQDETKKVDSPTE
KGTSDTKQPEKKDGDPKPVENKEQQPTPPATEKPKSPPKASPAEDKDKKSPPTPEDKSKQPEADKSKP
PADNAVKSTEAEKVQQATAPEKKVEAKDAEKDKKTPPTGLQPTPPAIPPRKNVSPNPGRKIPVPSHPQP
PKKTPLSTQPERRTPPTPPTNRKTPPTSRKEPPTPPASRRQQPPTPPTRRNPTSSNAGKRSSSSSRQSSAEKQS
KQNGKANEDKTDNKTEPKAGNKPDNKTVNKIDNKTENKTGVAATGQNSNNNTNGPKPTIPANDKN
VKTATDKRPPVSATDKSNDAANAKNNPQDKTKIPKLTEGTAPTTTDAKSTKGDPKNASNVKDSHQPK
PDAKHDKDPKHDAKQNDAKHDDNKALPHKDDKNKKSHDATQKSSKGENEGNIRSDAELNGAPLDP
ENPEKQAPPSSRLLNLAQKGEWSILEQSLRTLDKGNPQIFVVDKDTGVTPLMGCEGEQTGRCSTOP

- **NOMPC_Δ2_Version 2** (colors according to nucleotide sequence)

MACGRASGVVPTVQLLLRSTGKEARLAEDKEGSIPLFLAAEVGNTAVCRELLQQMTEQQIMAQKKDNG
DSVLHLACHRKDIDMTKLFIESGAVIDRQNTEGQTPLHIAAWEGDESIVKYLHQVKANPNIIDTEDRSPL
HIASERGHTIIVEILVDKFRANVAARTKDGSTLMHIASQYGHPDTALAFLKKGVPMLMPNKSGAICLHAA
AMKGHTSVIKALLTKGAPVDAMTKDGYTALHLSVQHSKPLAVQTLLGYGAAVELKGGQSKETPLHIAA
RTEDGEKCAEMLVKSGADVNATKENGETALHIAGRYGHLQMICALLAEGADPTWQSKSGETPLHIAVR
YCHWEVADELLRFVKKERSPIDATMLVNAPNIEGETAVHYAGELTKKQAHHEFEDTDMISLLLQNGGD
TNTHTKLTQETPIHYCARAGNADVMLEIVKHLGPNQVQVAVNKQAKNGWSPLLVASEQGHLDILKILL
QHHARVDVFDEHGKAALHLAAENGHEEVADVLLQHKAFVNAKSKLGLTPLHLAAQSGYNKLARLLIE
KHGATNDALSLAKKTPLHLAAQTGQIEVCSTLLKMKADANAKDAHSQTPLHLAAENDHSDVVKLFLK
HRPELVSMANTDENGMTCAHIAASKGSVAVIKELMRSNKSVIATARTKTTDSTTLHLAAAGGHAVVIKY
LLEAGASATDENAEGMTAIHLAAKNGHVNVLEALKENVSFRVTSTKTGLTALHVAAQHGQIDFVREML
TKVPATVQSEPPHTDGPMRELGTAEFGLTPLHLAAQSGHEGLVRLLLNSPGVQADAATAVHGVIPLHLA
AQSGHTAVVSLLLSKSTKQLHIQDKRGRTGLHLAASNGHLDMLSLLLGQGADINGYDKNGWTPLHYAA
KAGHLEAVKLLCETGANPTNETKDGKVPVCFAAGANHEEVVSYLLQTDHNTYNLMEDRKWVFDLMA
LGKMNDNKSIQEFILVSPAPVDTAVKLSRAFRNLSTKEKERARDLLAAGDFCEVMATELNAIAAGINGPA
ILLKSVDNRGVPLLDILIECEQKEVVADAVVQRYLSEVWAGYLSHWSSPKMMLLFCVLLFVPPVWLAFSLP
NKHIRFNKVPMIKFMLYLVSHIYLIMLFTLTIVYPIYPLWESGNLIPHWHEWLLLSWFSGLLVNQLTNPED
RQGLGWIKVIVIAISNFGVTTHLIAFSFSGNDRLYCLYARNQFFGLALMLCFVQLLDFLSFHQLFGPWAIII
RDLMKDLTRFLVILVIFMAGFTLELATIYQPVLAPLIPEPDKGTGDGSGGAGKETPISTFELLFFALFGLVD
PENLPLVNRSPLWSVTLAKAVFGCYLTLTIIVLINLLIAMMSDTYQRIQAQSDTEWKFGRAKLFRNMNKTS
ATPSPLNLFTKLYVYIVILIKHKSKVCSLTDVQQYIDEEEELDQDSENAAAGQGNWRNMMRRNTQVAPE
DFMQAALRSNGPTPVEEVVDWKVVADRFLYLRGLADGNNMDSDKEDDDDDDKFKSNENAAEQQ

- **NOMPC_i1_Version 1** (new residues and colors according to nucleotide sequence)

MSSTAKEQNETDKTPPDKAVDEKKEEDKKAEEAAKPTPSDKESKPENVDNKASASAQPQPPRLEKELP
LEPKEIPKTETAAQDTNQNNNKDTGTSKSTAPPAALDESSKKKEPENKDAPTELPKQDETKKVDSPTE

KGTSDTKQPEKKDGDPKPVENKEQQPTPPATEKPKSPPKASPAEDKDKKSPPTPEDKSKQPEADKSKP
PADNAVKSTEAEKVQQATAPEKKVEAKDAEKDKKTPPTGLQPTPPAIPPRKNVSPNPGRKIPVPSHPQP
PKKTPLSTQPERRTPPTPPTNRKTPPTSRKEPPTPPASRRQQPPTPPTRRNPTSSNAGKRSSSSSRQSSAEKQS
KQNGKANEDKTDNKTEPKAGNKPDNKTVNKIDNKTENKTGVAATGQNSNNNTNGPKPTIPANDKN
VKTATDKRPPVSATDKSNDAANAKNNPQDKTKIPKLTEGTAPTTTDAKSTKGDPKNASNVKDSHQPK
PDAKHDKDPKHDAKQNDAKHDDNKALPHKDDKNKKSHDATQKSSKGENEGNIRSDAELNGAPLDP
ENPEKQAPPSSRLLNLAQKGEWSILEQSLRTLDKGNPQIFVVDKDTGVTPLMAGCEGEQTGRCSTOP

- **NOMPC_i1_Version 2** (colors according to nucleotide sequence)

MACGRASGVVPTVQLLLRSTGKEARLAEDKEGSIPLFLAAEVGNTAVCRELLQQMTEQQIMAQKKDNG
DSVLHLACHRKDIDMTKLFIESGAVIDRQNTEGQTPLHIAAWEGDESIVKYLHQVKANPNIIDTEDRSPL
HIASERGHTIIVEILVDKFRANVAARTKDGSTLMHIASQYGHPDTALAFLKKGVPMLMPNKSGAICLHAA
AMKGHTSVIKALLTKGAPVDAMTKDGYTALHLSVQHSKPLAVQTLLGYGAAVELKGGQSKETPLHIAA
RTEDGEKCAEMLVKSGADVNATKENGETALHIAGRYGHLQMICALLAEGADPTWQSKSGETPLHIAVR
YCHWEVADELLRFVKKERSPIDATMLVNAPNIEGETAVHYAGELTKKQAHHEFEDTDMISLLLQNGGD
TNTHTKLTQETPIHYCARAGNADVMLEIVKHLGPNQVQVAVNKQAKNGWSPLLVASEQGHLDILKILL
QHHARVDVFDEHGKAALHLAAENGHEEVADVLLQHKAFVNAKSKLGLTPLHLAAQSGYNKLARLLIE
KHGATNDALSLAKKTPLHLAAQTGQIEVCSTLLKMKADANAKDAHSQTPLHLAAENDHSDVVKLFLK
HRPELVSMANTDENGMTCAHIAASKGSVAVIKELMRSNKSVIATARTKTTDSTTLHLAAAGGHAVVIKY
LLEAGASATDENAEGMTAIHLAAKNGHVNVLEALKENVSFRVTSTKTGLTALHVAAQHGQIDFVREML
TKVPATVQSEPPHTDGPMRELGTAEFGLTPLHLAAQSGHEGLVRLLLNSPGVQADAATAVHGVIPLHLA
AQSGHTAVVSLLLSKSTKQLHIQDKRGRTGLHLAASNGHLDMLSLLLGQGADINGYDKNGWTPLHYAA
KAGHLEAVKLLCETGANPTNETKDGKVPVCFAAGANHEEVVSYLLQTDHNTYNLMEDRKWVFDLMA
LGKMNDNKSIQEFILVSPAPVDTAVKLSRAFRNLSTKEKERARDLLAAGDFCEVMATELNAIAAGINGPA
ILLKSVDNRGVPLLDILIECEQKEVVADAVVQRYLSEVWAGYLSHWSSPKMMLLFCVLLFVPPVWLAFSLP
NKHIRFNKVPMIKFMLYLVSHIYLIMLFTLTIVYPIYPLWESGNLIPHWHEWLLLSWFSGLLVNQLTNPED
RQGLGWIKVIVIAISNFGVTTHLIAFSFSGNDRLYCLYARNQFFGLALMLCFVQLLDFLSFHQLFGPWAIII
RDLMKDLTRFLVILVIFMAGFTLELATIYQPVLAPLIPEPDKGTGDGSGGAGKETPISTFELLFFALFGLVD
PENLPLVNRSPLWSVTLAKAVFGCYLTLTIIVLINLLIAMMSDTYQRIQAQSDTEWKFGRAKLFRNMNKTS
ATPSPLNLFTKLYVYIVILIKHKSKVCSLTDVQQYIDEEEELDQDSENAAAGQGNWRNMMRRNTQVAPE
DFMQAALRSNGPTPVEEVVDWKVVADRFLYLRGLADGNNMDSDKEDDDDDDKFKSNENAAEQQ

PKD2-1

PKD2-1_WT nucleotide sequence (exon boundaries marked with |)

CCCGACTATGAAAGAAAATGGTGGCCCCATTATGATATGAGGAGCATGAATAATTCAGTAAAATTTGAACCGAGGCCAAT
GGAGCGTGACGCTTATCTTGACGTCAGACCTGTTGCCCGGCAGCCGCACCTTCTGAAACTGCCCAGCATCTCCGCTCCGC

ATCCCACACCTCAACGGCATCTTCAAGAGATTGACTTATTTGTAGAACGGACCACCAACCAACTTCTGGCTGATGTCAC Start codon
GCCCCGCAACGGCCCAACGGCCTATGTCGGCCATCTCGCGCAGGTCCAACAAGTCGGCATGGGAACCCGATGAGGGACGT
CCCGACTCGCGACAACAGCCTTACGATGATGACATGGCCCCCATGGAGAACGACCTCTACGACTCGTCTAACATGTCTGTG

GACAGTCGCCAGCCGGTCGTGCACGTGGCCGAGAACCAAAATGGCTGCTGGGACAAGTTCAAGAGAGGAGTGCGCT | CG

CTGTGGGCCACCCGTCAGACTGAGGACACCGGAGAGGACAGGGAGATGCACGTCAAGACCACCCTCCGAGAACTGGTCG target seq.+ PAM

TCTACTTGGTCTTCTTGGTCATCCTCTGTGTT | GTGACCTTTGGAATGACCAGTCCGACCATGTACTACTACACGCAGGTCA

TGAGCCAGCTGTTCCTGGACACCCCATCCGGAAATGGCGGAGTCACCTTCAGGAGTCTCACCACAATGGCCGACTTCTGGACT |
TTCTCTGAGCATGCCCTGTGCGATGGTCTCTACTGGGAGACCTGGTACAACCAGCGCAACGTGAGTGATAAAGAGCTCGGCTAC
ATCTTCTACGAGAACAAGTTGCTTGGTGTGCCGCGGGTGAGACAGCTCAAAGTTCGGAATGACTCTTGCACCGTCCACGATG
ACTTCAAGAAGGAGATCAAAGCTTGCTACGATGCCTACGCCCAGAGCATTGAAGACAAAGCTCCTTTCGGAAAAATGAATG

GAACTGC | CTGGACCTACAAAACAGAGAAGGAATTGGATGGCTCCGGCCACTGGGGTCTCATCTCATCCTACGGAGGTGGT
GGCTTTTACGAAGATCTTGACACCACCAAAGCCAAATCTGCCGCCATCATTGCCGACCTCAAACAGAACTTGTGGCTGGACAG

GGGAACCAGAGCCGTCTTCATTGACTTCACTGTGTACAACGCTAACATCAATCTCTTCTGCGTTATCAGG | CTTGTGGTTGAGT TM domain
TCCCGGCAACCGGTGGCGCCATCCCGCAGTGGGTGTTCAGGACCGTGAAACTGATCCGTTACGTGACTGTCGGGGACTACTT
CGTGATGGCTTGTGAGGGCATCTTCGTGCTGTTCATCCTGTACTACATGGTCGAGGAATCCATCGAAATCAAGAAACACAAAC
TGAACTACTTCAAGAGCTTCTGGAACATCCTGGACATCCTCGTCATCATAATATCCATCGTCTGCATAGCGTTCAACATATATCG
GACACTGGCGGTCGGGAAAAAGCTCGACCAGCTGCTTGCGAATCCGAACATGTTCCCCGACTTTGAGTTCCTCAGCTACTGGCA

GGTGCAGTTCAACTCTGCCATTGCCATCACCGTCTTCATGGCTTGGGTCAAGGT | GTTCAAGTACATCAGCTTCAACAAGACCA
TGACCCAGCTCTCCTCGACCCTTGCTGCCTGCGCCAAGGACTTGGCTGGTTTTGCCGTCATGTTCTTCATCATCTTCTTGGCCTTC

GCTCAGCTCGGATACTTGATCTTTGGAACCCAAGTCAAGGACTTCTCCAGCTTCGACGATGCCAT | CTTCACCCTGTTCCGAATCA
TCCTGGGAGACTTCAACTTCCAAGAACTGGAGGAGGCCAATCGTATCCTCGGACCCATCTACTTCTTGCTCTACGTCTTCTTTGTGT
TCTTCGTTCTGCTCAACATGTTCTTGGCTATCATCAACGACACCTACTCGGAGGTCAAGGCCGAGCTGGCCAACCAGAAGAGCGAGT

TCGAGATGACCGACTACTTCAAGAAGG | GCTACGACCGCATGCTTGACAAACTGGCCTTCAAGAGGGACAAGATCGTTGACATC
CAGAAGGCCATCAACTCAGCTGACCACAACCAGGACAAACAACTGGACTTTGACGAGTGGCGTCATGACCTCAAGAAACGTGGCTA
TGCTGAGGCGGAGATTGAGGCCGTATTTGCCAAGTACGATCTGGACGGAGACCGAGTGTTGGACGAGGGCGAACAGATGAGAATG
AACCGCGAACTGGAAGGACAAAAGGACCAACTCAACAAAGAGAGAGAGGAGCTGGAGGACCAGAGGGATGATGACCCTAACGGAGGCGGC
CCCATGGTCACCAATGGCGTCAGCTTCGATGAATTTAACGTGCTCGTGCGCAGAGTGGACCGCATGGAGCATTCAATTGG
CAGCATAGTCTCTAAGATAGACGCTGTCCTAGTCAAATTAGAAGCCATGGAGAGGGCGAAAGCAAAGCGCAGAGAAACCA

TGAGCAAGATTTTAGAAAGTATTACTGAG | TCTGATGGCGCAAGCGATGAGGCGAAGAGAGAACAGATGGAGCAGCTTGTC
CGACAAGAGTTGGAGAGGTGGGACACAGACGCCAACATGAAGGGCGCCTCTGGTGGCAAATCCAGTTCAAGACCCGGCTC

ATCAGCTCCCTCTTCTGGCAAT | TCGGAAGTA TAA GAAGATCGAGATATAACTGCTGAAACAATGATTTAAAGACAAGAAC Stop codon
GTTGAAAAAAAAAGTTGAAAAACTTCAAAAAA|GATTACTTTCTGTCTGCACTCAAAATTCGGACTATTGCACAAGCTTGC
CTGCAAAACGAAAATACCTAGTTATTAAAAACTAAAATACTGTTATATTGAGTCAAGAAGAAATTCTGTGATAAGTTCCT
GTGATATTATGAATAATCATATAACGATATATATATACACGATAATGTGTGATAAGAATTTGTGATATTTATTGACTTGC
CAGCTCAGAGATGAAAATTATAAAATCATATATAAGAATAAATTCGAATTTATCAGATCTACACTTCAGATTTCAACATT
AATTTTGTTGTGGTAAATGAAATAAAATTATTAAATTTTGAAGGAAAACTCTGCAGCTGGCATGCATAGAATTCTGTAAG

AAGCTGTTCACTTAATGATTGGAATAAAATTTCTGATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAA

PKD2-1 aminoacid sequences

- **PKD2-1_WT** (colored according to nucleotide sequence)

MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAEN
QNGCWDKFKRGVRSLWATRQTEDTGEDREMHVKTTLRELVVYLVFLVILCVVTFGMTSPTMYYYTQV
MSQLFLDTPSGNGGVTFRSLTTMADFWTFSEHALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPR
VRQLKVRNDSCTVHDDFKKEIKACYDAYAQSIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYG
GGGFYEDLDTTKAKSAAIIADLKQNLWLDRGTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVF
RTVKLIRYVTVGDYFVMACEGIFVLFILYYMVEESIEIKKHKLNYFKSFWNILDILVIIISIVCIAFNIYRTLA
VGKKLDQLLANPNMFPDFEFLSYWQVQFNSAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAG
FAVMFFIIFLAFAQLGYLIFGTQVKDFSSFDDAIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLL
NMFLAIINDTYSEVKAELANQKSEFEMTDYFKKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQL
DFDEWRHDLKKRGYAEAEIEAVFAKYDLDGDRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRD
DDPNGGGPMVTNGVSFDEFNVLVRRVDRMEHSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITE
SDGASDEAKREQMEQLVRQELERWDTDANMKGASGGKSSSRPGSSAPSSGNSEV

- **PKD2-1_Δ4_Version 1** (new residues)

SRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAENQ
NGCWDKFKRGVRSLWATRLRTPERTGRCTSRPPSENWSSTWSSWSSSVLSTOP

- **PKD2-1_Δ4_Version 2** (new residues and colored according to nucleotide sequence)

MHVKTTLRELVVYLVFLVILCVVTFGMTSPTMYYYTQVMSQLFLDTPSGNGGVTFRSLTTMADFWTFSE
HALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPRVRQLKVRNDSCTVHDDFKKEIKACYDAYAQ
SIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYGGGGFYEDLDTTKAKSAAIIADLKQNLWLDR
GTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVTVGDYFVMACEGIFVLFILYYM
VEESIEIKKHKLNYFKSFWNILDILVIIISIVCIAFNIYRTLAVGKKLDQLLANPNMFPDFEFLSYWQVQFNS
AIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLIFGTQVKDFSSFDD
AIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNMFLAIINDTYSEVKAELANQKSEFEMTDYF
KKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQLDFDEWRHDLKKRGYAEAEIEAVFAKYDLDG
DRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRDDDPNGGGPMVTNGVSFDEFNVLVRRVDRME
HSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITESDGASDEAKREQMEQLVRQELERWDTDAN
MKGASGGKSSSRPGSSAPSSGNSEV

- **PKD2-1_Δ137_Version 1** (**G** splice site in 2nd exon; new residues and colored according to nucleotide sequence)

MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAEN
QNGCWDKFKRGVRSLWATRQSDLWNDQSDHVLLHAGHEPAVPGHPIRKWRSHLQESHHNGRLLDFLS
TOP

- **PKD2-1_Δ137_Version 1** (**G** splice site in 3rd exon; new residues and colored according to nucleotide sequence)

MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAEN
QNGCWDKFKRGVRSLWATRQMTFGMTSPTMYYYTQVMSQLFLDTPSGNGGVTFRSLTTMADFWT
FSEHALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPRVRQLKVRNDSCTVHDDFKKEIKACY
DAYAQSIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYGGGGFYEDLDTTKAKSAAIIADLKQ
NLWLDRGTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVTVGDYFVMACEG
IFVLFILYYMVEESIEIKKHKLNYFKSFWNILDILVIIISIVCIAFNIYRTLAVGKKLDQLLANPNMFPD
FEFLSYWQVQFNSAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLG

YLIFGTQVKDFSSFDDAIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNMFLAIINDTYS
EVKAELANQKSEFEMTDYFKKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQLDFDEWRHDLKK
RGYAEAEIEAVFAKYDLDGDRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRDDDPNGGGPMVT
NGVSFDEFNVLVRRVDRMEHSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITESDGASDEAKRE
QMEQLVRQELERWDTDANMKGASGGKSSSRPGSSAPSSGNSEV

- **PKD2-1_Δ14_Version 1** (new residues)

MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAEN
QNGCWDKFKRGVRSLY**STOP**

- **PKD2-1_Δ14_Version 2** (new residues and colored according to nucleotide sequence)

MRDVPTRDNSLTMMTWPPWRTTSTTRLTCLWTVASRSCTWPRTKMAAGTSSREECARCTEDTGED**REM**
HVKTTLRELVVYLVFLVILCVVTFGMTSPTMYYYTQVMSQLFLDTPSGNGGVTFRSLTTMADFWTF
SEHALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPRVRQLKVRNDSCTVHDDFKKEIKACYD
AYAQSIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYGGGGFYEDLDTTKAKSAAIIADLKQNL
WLDRGTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVTVGDYFVMACEGIF
VLFILYYMVEESIEIKKHKLNYFKSFWNILDILVIISIVCIAFNIYRTLAVGKKLDQLLANPNMFPDFE
FLSYWQVQFNSAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLI
FGTQVKDFSSFDDAIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNMFLAIINDTYSEVK
AELANQKSEFEMTDYFKKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQLDFDEWRHDLKKRGY
AEAEIEAVFAKYDLDGDRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRDDDPNGGGPMVTNGV
SFDEFNVLVRRVDRMEHSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITESDGASDEAKREQME
QLVRQELERWDTDANMKGASGGKSSSRPGSSAPSSGNSEV

- **PKD2-1_Δ5_Version 1** (new residues)

MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAEN
QNGCWDKFKRGVRSLWATHSTOP

- **PKD2-1_Δ5_Version 2** (new residues and colored according to nucleotide sequence)

MRDVPTRDNSLTMMTWPPWRTTSTTRLTCLWTVASRSCTWPRTKMAAGTSSREECARCGPPTEDTGEDR
EMHVKTTLRELVVYLVFLVILCVVTFGMTSPTMYYYTQVMSQLFLDTPSGNGGVTFRSLTTMADFWTFSE
HALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPRVRQLKVRNDSCTVHDDFKKEIKACYDAYAQ
SIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYGGGGFYEDLDTTKAKSAAIIADLKQNLWLDR
GTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVTVGDYFVMACEGIFVLFILYYM
VEESIEIKKHKLNYFKSFWNILDILVIISIVCIAFNIYRTLAVGKKLDQLLANPNMFPDFEFLSYWQVQFNS
AIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLIFGTQVKDFSSFDD
AIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNMFLAIINDTYSEVKAELANQKSEFEMTDYF
KKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQLDFDEWRHDLKKRGYAEAEIEAVFAKYDLDG
DRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRDDDPNGGGPMVTNGVSFDEFNVLVRRVDRME
HSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITESDGASDEAKREQMEQLVRQELERWDTDAN
MKGASGGKSSSRPGSSAPSSGNSEV

PKD1-1

PKD1-1_WT nucleotide sequence (1st and 2nd exon boundaries marked with |)

CGAAAAAATCTTCAGAATTCGACGAGCCTCGATCAGGGCGATACAGACGGAGATGGTAGAGCCACTTACTTCCGCTCTTTCCTC
TCTTTGCACTTTCTTGTCTTTTGTTTCAACTTTCTTGGGCCGCGAGCGGTTCAGGACCAGTGGAGGTCCGCTCCTTTGCC
ATTCCGAATGGACAGTCACTAACCACTCCACAGAACCTTGGGGCCTTTCCCCATGACTGTGTGGGCAACATCCAGGTGTG
TGACAATGGCTTCACTGTTGGCTTCTGGTACAAGGCCCTTGCGGGAGATCCTGACCTTCCGGAGAGGTACGTTATCTCTT

CTGGAGGACAGAGTGCC**ATG**TCCGAGGGCTTCTACGTCCTCCAGCACTATGGAAAGGACTATGAAGTTGGAGTCGCAAAG Start codon
GGTAACAAACTCTGGACCGTTAGCCTGTTTGTGCATGAAGATGCCCTCGTCTACTTCGTCCTGACCTGGAGCGAACAGGG
CTTAGACGTTTACGTGGACGGCATACTCGCAGGAGCAGACGCAAAAGGAAAAGCACGGGTGTTCACCTCTCCAGCCTACG
ACCCTTTCCCCGAAATTGTGGTAGGGTATGCAAATGACAGACCTTACACCGGGAAGACCATCAATGAACACGTCGTCAAT

GTCACCCACTGGTCAGCCTATTACCATCAGGAGGAAGTGGAGCACCTTCAAG | GTGAAGCTCCAAAGGACCTTTATCACAA

GGGCTGTGTGTCAACTTTACCCCAGGGCGTCGATGATGGAGTCAACGTGAAAGATGTGGACACGAAAGGCGTCTACGACA target T1 seq.+PAM
GGGTGGCCAAATGCCGAAGGAAGTGCGCAGAGAAAGGGGCTTATGTGGCCTTCATGAAGGGAGGAGAAATGTGCTACTGC
TACATCGATGGCTTCACCATGGACTTCTTGGATGGTATGTCGACATCTACTTCCTGTTCCGATTACGATGTCTTCTGGGT
CAAGAGCCACGAGACTGAGCCAGATTACCCGATTGAACTGGAGGTGGAAGTCATCCCTTCAACTCCAAAACCATACATCA
GGCCGGAAGAAACAGTGACCTTCAAATCAACATTCACGACTGACGAGTATGTGACCTTCACTTACGACTTTGGCGATGGG

ACAGACAAGACCTCGAAGAATCCTGTCATCAGCCATTCATGGGAAAAGGAAGGGTCCTTCCTTGTGAAAATCACAGCCAC
AACTAGGCTCATGTCAGAGACCGTTTTGAAGACGATCAACATCACTTATGTAGAAGAAGGTGTTGCTCCTGAACAAGTGC
TTGTCAAGATGAGCCCCAAAGATTCATCTTTGGTTGAGTTTGACCTGACAACCATTGGCAAGTATGAGACTGAATGCACA
ATAGAGTTTGGCGATGGAGAATCTGAAGATGTTTCCGTCAAAGACTACATTGAGCAACTCTCGCTCTCCCATGACTACAA
GACGCCTGGCTTCTATGATGTTGTTTACAGCTGTGAGAATGAATATGGTGTGACAGCAGATCAAGGTCACGCCTTTGCTG

TCAAGC**ATG**AGATTGAATACCAGAATGAA**CCACGCTTCACTGATATCCGCCTT**TGAAGTTGTCGGAGCTGACTCCCATCAA target T5 seq.+PAM

GATGTCTTCGGTCTGAAGCTGAATGGTCTCCAAGTGGCAGCTAACACCACAAGTGATAATGTCATCATCCAAGCCAACAA
GTTCCTGTACAGCGGAGAACACATCGTAGAAGTCACAGCCCAAGATGATGAAGTCAGACGAAGAATCTTCAACATTGAAG
AGAAGATCGAGAAAGTAGCCATCTTGACCAACAACTCACAGGTGCTGCCTGGAGAAGAAGTTACCTTCACTTTCCAGATC
TTCAAAGGAGACCACATCCACGTTCTGATTGAATATGGAGATGGTGCATCTGAAAGAGTCTATATCCCACAGGGAGGTCA
GCCAGCTGCCATCACAAGGGTCCACAAGTATGAGTCATTGAATCACTACAAGGTCAAGGTTACTGCAGTTAACTACATCA
GCAAGGTCATCAAAGAACATGATATATCAATTGAAAGGCCCCTCATGTCAGCAAATATGCATGCCCAGAATGTCACCGTG
CTCCTCACTGAAACAATCATTACAATCAAGGTGGATGAGAACAGCCCTGCAACACCAGTGGATGTCATGTTTGACTATGG
AGATGGAAAGAGGGAAACAGTTCGATTGGGTTCCCACCGAAGTAAAGCAGAGCCAATGGTCCACAGTCACATATATTCAG
AATATGGCATTTACGTTATCCATGCTATCATCAGGAACAACATTAGTGAAATAGAAGTTGAAGCTCTGCACCAAGTTGGT
GAGAATATCACCATGTTGGACTTGTATGCAAACCGATACAGATTACTTGATGGTGAGGAAATCAACATTACCATGGACGC
TCCCTATGGTTATCCACTTGTGTACAATCTTGACATGGGAGATGGAACCACCCTCACAGTCAGAAGACCTCTTGGATACA
CCGGAGCTTTTGAATCAGATCCAATTGCCTCAACAACCGGCAAGACCTATTACCCACCAGATTTTGCTGTGACGGAGGTC
ACCAAAACAGAAAACACAACTTCAGTCATACTCGATCGGCAAAGGAGGGATGTTAACATCATGCCCACCACTGTTATTGT
CCCAGGAGAGGGGGAGGTCCTTGCAGTAATGAGCACTCCAGCTCCAATGGCAGCAGATAGTGACATGTCACAAAACAGAG
GGCAGAGGAATGAAAAGGTCACAGTGACATACCGCTATTCACGAAGTGGTACCTTCAACGTCATGGCTACAGTTGACAAT
CCATTTGGAGGAGAAAGCACATACTTATGTCCTGATATTGTTGTACTCCCAGCTAACCAGAAGCAACCAGATTGCTCAGG
GCTAACAGTTTCATTCCATAATGCAACATCAAAGTCATCACCTCTTGTCATTCAGAGGTCAAAGAGCCTCACAATCCCAA
CTGATGTTCAGCTGAATTGTCAAAAGGATGCTTCTCAATACCAGGTTGACTACACTTGGAAGGCTGCTAGACACGTCATG
GGTGACTTATGGAGGCCTGAATTAGAGGTCTGCCATTCTCAAGTACCTGAGGCAAACTTGGTGATCCCAACCAACACTCT
TTGGTACGGTCTATACAGGATCAAGGTCATCACCGGCATCAGACTACCACATCCTACAAGGAACCGAAGGTCAACGAACA
CAACAAAGGGAGCCACCAATTCCATCTTGAACTCCGATCCTTCCCTGATCACAGCTGAGAAAGATGTCTACCTCAAAGTG
GAGAAGACCCCATTGGTTGCTCAACTCCGCAACTCAAGCAAGAAACTGGATGTGTTTGCCAACACCGCAGTCACCATTTC
CGGTACGGGAAGTTTTGATCCGGATGTCGTCGGAAAACACAGCAACAAGACAGGCATGACTGGTGCAATCATTTGCTACC

CTCA | CTCAAGAGATGCTGAAATGAAAAAACTCTCATCCAAAGAACTATTGGCACTGACAACACTAATCCAGAACAACAGC
AGAAACACCGTCAGCTTGTATGACACTGACTGCTTCGTCAATGCGACCAATGTTCAGATGTTGGACTGGGATGTCACCTT
CTTCGGTGACTTATTGGTGAAAGACAAGGACCTTACTTTCAGGTTGTTTGTATGGAAGGATGACAGGGAAGCCACTACTT
CTCAAATTGTCAAAGTGCTTCAGGATCTTGATGGCGACAATCTTTTTGCCAACTTGGACAACCTGTTGAACAGTGGAGAC
ACTTCCGCTGCCCTCTATGTGCTTGGACACTTCACAGCTGGTTTGAACGGCGGAGGAGAGGAAGATGGGGAAAAGAATAA
CACAGATGTGAAAAATGAACGTTCTGAGCTAAGAGGTAAGCTCATCGGTGTCCTTGGCCGTGTGGCTGACCAAGTGGGCA
ACTTTGACCAATTGACCAAAACAACTGATGTTCTGTCCGGTGTGACTGAGGTCAGGGATGAGGTGACAGAAGACAGCAGG
GCTTCTGCAGCAGATGCCATGTCCAGTGGAGCAGCCAAAACAGTGAATCTGACAAATGATGCTCCGCTGGATCAGGTTGC
TGGTCTGGCTGGAGGAATGGTTGGAGTGGCAGGAAATGTTCTGCCATTTGCTGATGAAGAGGAAGAAACAGTCATAGAAG
CAATTCCTGCTGCTTCTGGAGTAACAACGATCAAGTACACAACCGTCGAAACAAAGACGCTCGCCACAACGATCAAGCCT
CCATCAAACTTCGAACCCACAACCCAGACTAGTGTGACCACGGCTTATGTGCCTACTACAGCTGCTCCAACTGTTGCAGC
TAGTGACAACGACACAATTCTTGTGAATGGCACCTGGGTGACAGCCACACCTGTAACAACTGTTGTCTCCACCATGGTTG
TGCCTTCAGTGACAACAGTCACAATCAAGAGCCCCAAACAAACTGTGACCAACACCAAAGTTACTTCTTACAAGACTCAA
CAGGTGACCACTACCACAGTTGCATACACGACCACAGCTCCTCCCACAGAAGCCACCACACAGAACATCATAGCAAACCTGAA
GACCACCATGCCTCCGGGTCTAGCAGCAGTTGTCGAGGAAGACGTTCCTCCATGTTGTTCCACCAAGGGCGGCCTCT
GTGCCTGGATTGAACCGCCTTGCGAGTTCTTGTCTGAGGATGACAGAAAGAAGCAAGTTTGGGACGACAACATGTGCAAC
GCTCTCCACCCGGATGACTTGAAAGAAATTGCATGGAGATATGGACTCTTCAAGGCTCAACAGAAAAGGATGCTCTTTGC
CAGGAGGAAACGTACAACACAGGGAGCCATGGGAGCTGTGGACAACGTAGGAACAGCTGTTCTGGGCAAAGCCGATAAAG
ACACCAATGTGTCGATCAGCACACCCAAAATGAAGTTGGAATTGGCTAAGGTTGAACCTGGACAAAATGGCACCATGAAT
GGCATTGCCTTCCCGGGATCTGCTCTTGGTGACACAGAAGGAGCTGGAGCTTCATCCATGGAATCTGACAACGCGTTCAC
CTACAGTGACTCTGCCAAATCAGTGAAGGGCAAGGTCAAGTCTTTGGCTTTCAAGGACAAGGATGGGAATGAGATCAAGA
TGAACGGGACTGCTGAGCCTTTTGTCATCTGGATGGATGGTCCAGGGGACAGCTTTCCAGATCCAAAACAAGCAACTTTT
GCTACGTCTACTCAGGTTCTGAAATTCAACTACCACACCATCAACACAACCTACAATGACTCATCCATCCACTTCATGAT
CAGGCCAAGATCTCCTCATGATGTGTATGAAGTCTACATCAAGTGGGAGAACTTCACACTGATGAACAGTTCAAAAGAAA
ATGGCAATGTAGGATTCCCCAATGAGACATGGTATGACCACAAAGGAATTGTTCCCAATGATGAATCTCCTCAAGGTGAT
GATCCTATTGTTAATGAGTTGCTCCGATACACCTACTTCCCTCCTCCTGAGGTTACTCGCCTCAACGGAACTTACAGAGT
TGGCGTCCGTCTGCTAGAAACCAAACTGAGTTTTGCAAATGGACCGTCCCAGAATAACTACACCTACTTCTACAATTCAT
TTGTGTCTGGATGTCGATACTGGGACGAAGAAGCAGACGCATGGTCTGCCGCTGGTTGTACTGTTGGCAACTTGACTTCA
ATCTACAGCACCCAATGTCTCTGCAATCACTTGACGTCGTTTGGAAGTGACATGGTTGTTCCGCCAAACACTATTGATTT
CTCCACTGTCTTCTCCTTGGAGAACTTCATCGAAGCCATCCCTGTCTTTGTGACTGTTGGTCTGGTTCTCCTTCTCTATC
TTGGAGTCATTATTTGGGCAAGGAGGGCGGACAAGAAGGATGTCATCAAGTGGGGAGCTCAACCTCTTGAAGACAACCTG
CCATCAGATCAGTACCACTACCAGATCACAGTGTACACAGGAGGCAACAAGAATGCTGAGACTACTTCAAGAGTAAGTCT
CATCATTTCTGGGGATGATGGCGACACCGGAGTCAGACGTCTTTGGGATGGCAAGCGAAAGGAACTTGAAAAGGCCAGCA
TCAACAACTACATTCTGTCAACAGAAGGATGCCTGGGACCCCTTTCTTTCCTTCGAATCTGGCATGACAACACTGGAAGT
GGGAAGAAGAGATCCTGGTTCTTGGACCAGGTCCAGGTGAAGGACCTCCAGACTGGTGAACTCTTTTGCTTCCTGTGTGA
CCGATGGCTTGCTGTTGAACAAGATGATGGATTAGTGGAAAGGATCATCCCTGTAGCAGGACTGAGAGAAATGACAGCCT
TCAAACACCTCTTCAACTCCTCGGTCAAGAAGAAGCTGGCAAACGACCATCTCTGGGTGTCAGTGGTAAAGCGTCCTACT
CGCAGCAACTTCACCAGAGTTCAGAGAGTGTCTGCGTGTGCCGCTTTGCTCTTCCTCACCATGATTACCAACTGCATGTT
CTTTAAGGCTGATGAAAACAAGGAGAATGTCCGGGCAATCAGCTTCGGTCCATTTGAATTCACTCTGCACCAACTGTTCA
TCTCTGTGGTGACGACCTGCATCGTGTTCCCTCCATCTCTGCTCATCATCACCATCTTCCGACGTGTCAAACCCAAGAGC
AACGGTGTGATGCAATCCAACCAGAGGTTCTCCCAGAAGTCCAGCAAGTTCAAGTGGAGAAGGTACAATCCTGGATCATC
TCTGTGGGACAACCAGCCAAGCAGGGCCCAGAGGATGAAGGAATCTCTGAGGAACATCTTGAGCCAACACCAGAAGACCA
AGTACCAAGATGACAATCTGGATGAACTGGGACTGCCCAAGAAGAAGAAGCCCTGGATGTTGCCTCACTGGTGCATCTAC
ATTGCCTGGACATTGTGCATCACTGCTATCCTGGCCTCAGCTTTCTTCACTGTGCTCTACAGTATGCAATGGGGAAGGAA
GAAGTCCCTGGACTGGCTCAGCACATTCGTCCTGTCCTTCTTCCAATCTGTCATTGTCGTCCAGCCAGTCAAGGTGCTTT
TGCTGGCAGCTTTCTTGTCATGGCGTTATGAGGAAGCCGGACTTGGACAATGAAGACGAAGTGTTTGACGAATGCCACAAT
GCTTTACCTGGACCTGATGAATCTATGGACAAGGCTAATGCTAGAAATATTGATGACATCCTGATCCAACGCAAGTCAGC
AGTGACCCACATGCAGCCTCCGGACATGGCCAGCCTGGAGGAAGCAAGGCAGCTAAGGATGAAGGAGATGGCAATGGAGG
CCATCTTGAAGGAGTGTATTGTCTACTTCTTCTTCATCCTAGTTGTTTTCTTCCTCTCCTACCAAGCAAGAGACAGTGAC
TCCTACCAATTTGCCAACAACATCAAGAACCAGTTTGTCCGCGGCAGCACTCCTCTTGATTCGATCTCAAAACCAGACTC
TGTGTGGACCTACCTGAGAAACACAGTCCTTCCAAACATGTACTCCAACAACTGGTACAACGGTAGAGAGTTAACAGATT
GGAGAGAATTGCTGCAGATAGATGACAGATGCACCATCCGAGTGGGGGTGGCAAGGCTGAGGCTGCAAAGGATCAAAGAG
GATTCTTGCGAGGTGCAAAAAGTCTTCAGGCCGATCTTCGACCACTGTCGAAACAACTACAATTGGATGGATGATGACAC
CAAAGACTATCTCCCTGGCTGGGTTGTTCCCACAGCAGAGAACAGACACAGGCAAGGCGTCATTGACTGTTATTCTGAAG
CTCTTGAGTTCGAAAATGAAACCAGTCTCAGAGATGAAGAGAGTCCTTGGATCTACAGGAACTCGATAGAATTAAAGAGC
GCCCCTATGCCGGTACACATGCAATGTACAAAGGTGGCGGCTACACGTTTACCTTCAAGAGGACCCCAGAAGCAACCAA
GAGACTTCTTGATGAGATGGAGCAGTACGAGTGGGTCAACCCCAATGCCAGAGCCATGTTCATCGAGTTCACTCTCTACA
ATGGCAACGTCAATCTCTTCGGCAGCGTTATTTTGCTTGTTGAGTGGATAGCCGGTGGTGGAGCCACTGTACGACATGAG
GTCAAGGTTTTCCGACTGAACAACTACGTTGGCAGTTTTGGAGTGATAGTCATCTTGTTTGAGGTGCTCTTCTTGGCCTT
CACCGTGTACTTCTTGGTCAAAGAAATCACCAAACTAAAGAAGGAGAGAATGAAGTACTTCAAGGAGTTCTGGAACATGT
TGGAACTCAGCACTCTAATCTTATCCTTGGTCGTCATTGCCATGTATGCACTCAAGGCTATATTCGGAAAGGTAGCTCTC
CATGTTCTGGAACAGAGCGAATCTGGGGAGTATGTCAACTTTGTGACCATTGGTGTTTGGGATGAGATGTTCACCTTCCT
GCTGGCATTTGTTGTCTACCTCACGACCGTCAAGTTCATCAAGATGCTGCGATTCAACAGGAGAATGGGCATGTTGGGAG
ACACGATCCGCATTGCCACCAAGGACCTGAAGAACTTCATCATCATGTTCTTCATCTACTTCTTTGCGTTCTGCATGTTT
GCTTATGTCATCTTCGGCACGCAGATGGGCACTTACGGCAACTTCATCAAGACCATGGAGTCCCTGTTCTCGTTCGCACT
CGGCAGCTTCGACTTCGATGCAATGCGCCAGGCCTCCAACATCGGGGTCTTGTTCTTCATCCTCTACGTGTTTGTGGTCT
TCATCGGTCTCATGGGCATGTTCCTCACCATCATTGCAGACGCATTCACTGCGGTCAAGGAAGACACGGCCCTCCGATCC

AACGATTACGAGAATGTTGACTTCATCCTCGGCAAAATCAAAGGCGTGTTCGGCTGG**TAG**GGGGACGCATGTGCGTAAAC Stop codon
TGCTTCTGCAGGCATGTGCGTCCGGAGCTAACCTCATGAACTGATATACTGTAAAGCAACCAGGACTTAGATTCTGTAGT
TCAGTACCATATTGTTCAATTCCATTTGAGTATTCTACTTAATCACTCTAAATAAACCAGTGCTCTGAGAGACAATGGGA
TTGTCACCAACTGTGATTGAAAATTCTGTATAAGATAAAATAATTGTCTGTGATGGAAAAGCACTTCTTGAAATTTGCA
CAAAATGATCTTTCGATGTAGAATTTCCACATTGTTACCGAAATGTATTAGTATAAACAATGCAATTAAGAAATAAAAAT
TTTGTATAGACCACCCGCAAATAAACATGGAAGTTTAAGACACTGTGTGTGATGTCACATTTCATGCCTCCTTGGCTAAT
TTCTGTATTCTCACTTAAACAATAAAAGCTCACTCAGCAAGAACTGCAGAACTAGAGTAAAAAAAA

PKD1-1 aminoacid sequences

- **PKD1-1_WT** (colored according to nucleotide sequence)

MSEGFYVLQHYGKDYEVGVAKGNKLWTVSLFVHEDALVYFVLTWSEQGLDVYVDGILAGADAKGKA
RVFTSPAYDPFPEIVVGYANDRPYTGKTINEHVVNVTHWSAYYHQEEVEHLQGEAPKDLYHKGCVSTL
PQGVDDGVNVKDVDTKGVYDRVAKCRRKCAEKGAYVAFMKGGEMCYCYIDGFTMDFLDGMSTSTSC
SDYDVFWVKSHETEPDYPIELEVEVIPSTPKPYIRPEETVTFKSTFTTDEYVTFTYDFGDGTDKTSKNPVIS

HSWEKEGSFLVKITATTRLMSETVLKTINITYVEEGVAPEQVLVKMSPKDSSLVEFDLTTIGKYETECTIE
FGDGESEDVSVKDYIEQLSLSHDYKTPGFYDVVYSCENEYGVTADQGHAFAVKHEIEYQNEPRFTDIRL
EVVGADSHQDVFGLKLNGLQVAANTTSDNVIIQANKFLYSGEHIVEVTAQDDEVRRRIFNIEEKIEKVA
ILTNNSQVLPGEEVTFTFQIFKGDHIHVLIEYGDGASERVYIPQGGQPAAITRVHKYESLNHYKVKVTAV
NYISKVIKEHDISIERPLMSANMHAQNVTVLLTETIITIKVDENSPATPVDVMFDYGDGKRETVRLGSHRS
KAEPMVHSHIYSEYGIYVIHAIIRNNISEIEVEALHQVGENITMLDLYANRYRLLDGEEINITMDAPYGYP
LVYNLDMGDGTTLTVRRPLGYTGAFESDPIASTTGKTYYPPDFAVTEVTKTENTTSVILDRQRRDVNIMP
TTVIVPGEGEVLAVMSTPAPMAADSDMSQNRGQRNEKVTVTYRYSRSGTFNVMATVDNPFGGESTYLC
PDIVVLPANQKQPDCSGLTVSFHNATSKSSPLVIQRSKSLTIPTDVQLNCQKDASQYQVDYTWKAARHV
MGDLWRPELEVCHSQVPEANLVIPTNTLWYGLYRIKVITGIRLPHPTRNRRSTNTTKGATNSILNSDPSLI
TAEKDVYLKVEKTPLVAQLRNSSKKLDVFANTAVTISGTGSFDPDVVGKHSNKTGMTGAIICYPHSRDA
EMKKLSSKELLALTTLIQNNSRNTVSLYDTDCFVNATNVQMLDWDVTFFGDLLVKDKDLTFRLFVWKD
DREATTSQIVKVLQDLDGDNLFANLDNLLNSGDTSAALYVLGHFTAGLNGGGEEDGEKNNTDVKNE
RSELRGKLIGVLGRVADQVGNFDQLTKTTDVLSGVTEVRDEVTEDSRASAADAMSSGAAKTVNLTNDA
PLDQVAGLAGGMVGVAGNVLPFADEEEETVIEAIPAASGVTTIKYTTVETKTLATTIKPPSNFEPTTQTS
VTTAYVPTTAAPTVAASDNDTILVNGTWVTATPVTTVVSTMVVPSVTTVTIKSPKQTVTNTKVTSYKTQQ
VTTTTVAYTTTAPPTEATTQNIIANLKTTMPPGLAAVVEEDVPPCCSTKGGLCAWIEPPCEFLSEDDRKK
QVWDDNMCNALHPDDLKEIAWRYGLFKAQQKRMLFARRKRTTQGAMGAVDNVGTAVLGKADKDT
NVSISTPKMKLELAKVEPGQNGTMNGIAFPGSALGDTEGAGASSMESDNAFTYSDSAKSVKGKVKSLAF
KDKDGNEIKMNGTAEPFVIWMDGPGDSFPDPKQATFATSTQVLKFNYHTINTTYNDSSIHFMIRPRSPH
DVYEVYIKWENFTLMNSSKENGNVGFPNETWYDHKGIVPNDESPQGDDPIVNELLRYTYFPPPEVTRL
NGTYRVGVRLLETKLSFANGPSQNNYTYFYNSFVSGCRYWDEEADAWSAAGCTVGNLTSIYSTQCLCNH
LTSFGSDMVVPPNTIDFSTVFSLENFIEAIPVFVTVGLVLLLYLGVIIWARRADKKDVIKWGAQPLEDNLPS
DQYHYQITVYTGGNKNAETTSRVSLIISGDDGDTGVRRLWDGKRKELEKASINNYILSTEGCLGPLSFLR
IWHDNTGSGKKRSWFLDQVQVKDLQTGELFCFLCDRWLAVEQDDGLVERIIPVAGLREMTAFKHLFNS
SVKKKLANDHLWVSVVKRPTRSNFTRVQRVSACAALLFLTMITNCMFFKADENKENVRAISFGPFEFTLH
QLFISVVTTCIVFPPSLLIITIFRRVKPKSNGVMQSNQRFSQKSSKFKWRRYNPGSSLWDNQPSRAQRMKES
LRNILSQHQKTKYQDDNLDELGLPKKKKPWMLPHWCIYIAWTLCITAILASAFFTVLYSMQWGRKKSLD
WLSTFVLSFFQSVIVVQPVKVLLLAAFLSCVMRKPDLDNEDEVFDECHNALPGPDESMDKANARNIDDIL
IQRKSAVTHMQPPDMASLEEARQLRMKEMAMEAILKECIVYFFFILVVFFLSYQARDSDSYQFANNIKNQ
FVRGSTPLDSISKPDSVWTYLRNTVLPNMYSNNWYNGRELTDWRELLQIDDRCTIRVGVARLRLQRIKED
SCEVQKVFRPIFDHCRNNYNWMDDDTKDYLPGWVVPTAENRHRQGVIDCYSEALEFENETSLRDEESP
WIYRNSIELKSAPYAGTHAMYKGGGYTFTFKRTPEATKRLLDEMEQYEWVNPNARAMFIEFTLYNGNV
NLFGSVILLVEWIAGGGATVRHEVKVFRLNNYVGSFGVIVILFEVLFLAFTVYFLVKEITKLKKERMKYF
KEFWNMLELSTLILSLVVIAMYALKAIFGKVALHVLEQSESGEYVNFVTIGVWDEMFTFLLAFVVYLTTV
KFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFCMFAYVIFGTQMGTYGNFIKTMESLFSFALG
SFDFDAMRQASNIGVLFFILYVFVVFIGLMGMFLTIIADAFTAVKEDTALRSNDYENVDFILGKIKGVFG
W

- **PKD1-1_T1_Δ1 Version 1** (new residues)

MSEGFYVLQHYGKDYEVGVAKGNKLWTVSLFVHEDALVYFVLTWSEQGLDVYVDGILAGADAKGKA
RVFTSPAYDPFPEIVVGYANDRPYTGKTINEHVVNVTHWSAYYHQEEVEHLQGEAPKDLYHKGCVSTL
PRASMMESTSTOP

- **PKD1-1_T1_Δ1 Version 2** (colored according to nucleotide sequence)

MKGGEMCYCYIDGFTMDFLDGMSTSTSCSDYDVFWVKSHETEPDYPIELEVEVIPSTPKPYIRPEETVTF
KSTFTTDEYVTFTYDFGDGTDKTSKNPVISHSWEKEGSFLVKITATTRLMSETVLKTINITYVEEGVAPE
QVLVKMSPKDSSLVEFDLTTIGKYETECTIEFGDGESEDVSVKDYIEQLSLSHDYKTPGFYDVVYSCENE
YGVTADQGHAFAVKHEIEYQNEPRFTDIRLEVVGADSHQDVFGLKLNGLQVAANTTSDNVIIQANKFL
YSGEHIVEVTAQDDEVRRRIFNIEEKIEKVAILTNNSQVLPGEEVTFTFQIFKGDHIHVLIEYGDGASERV
YIPQGGQPAAITRVHKYESLNHYKVKVTAVNYISKVIKEHDISIERPLMSANMHAQNVTVLLTETIITIKV
DENSPATPVDVMFDYGDGKRETVRLGSHRSKAEPMVHSHIYSEYGIYVIHAIIRNNISEIEVEALHQVGE
NITMLDLYANRYRLLDGEEINITMDAPYGYPLVYNLDMGDGTTLTVRRPLGYTGAFESDPIASTTGKTY
YPPDFAVTEVTKTENTTSVILDRQRRDVNIMPTTVIVPGEGEVLAVMSTPAPMAADSDMSQNRGQRNEK
VTVTYRYSRSGTFNVMATVDNPFGGESTYLCPDIVVLPANQKQPDCSGLTVSFHNATSKSSPLVIQRSKSL
TIPTDVQLNCQKDASQYQVDYTWKAARHVMGDLWRPELEVCHSQVPEANLVIPTNTLWYGLYRIKVIT
GIRLPHPTRNRRSTNTTKGATNSILNSDPSLITAEKDVYLKVEKTPLVAQLRNSSKKLDVFANTAVTISGT
GSFDPDVVGKHSNKTGMTGAIICYPHSRDAEMKKLSSKELLALTTLIQNNSRNTVSLYDTDCFVNATNV
QMLDWDVTFFGDLLVKDKDLTFRLFVWKDDREATTSQIVKVLQDLDGDNLFANLDNLLNSGDTSAAL

YVLGHFTAGLNGGGEEDGEKNNTDVKNERSELRGKLIGVLGRVADQVGNFDQLTKTTDVLSGVTEV
RDEVTEDSRASAADAMSSGAAKTVNLTNDAPLDQVAGLAGGMVGVAGNVLPFADEEEETVIEAIPAAS
GVTTIKYTTVETKTLATTIKPPSNFEPTTQTSVTTAYVPTTAAPTVAASDNDTILVNGTWVTATPVTTVVS
TMVVPSVTTVTIKSPKQTVTNTKVTSYKTQQVTTTTVAYTTTAPPTEATTQNIIANLKTTMPPGLAAVVE
EDVPPCCSTKGGLCAWIEPPCEFLSEDDRKKQVWDDNMCNALHPDDLKEIAWRYGLFKAQQKRMLFA
RRKRTTQGAMGAVDNVGTAVLGKADKDTNVSISTPKMKLELAKVEPGQNGTMNGIAFPGSALGDTEG
AGASSMESDNAFTYSDSAKSVKGKVKSLAFKDKDGNEIKMNGTAEPFVIWMDGPGDSFPDPKQATFAT
STQVLKFNYHTINTTYNDSSIHFMIRPRSPHDVYEVYIKWENFTLMNSSKENGNVGFPNETWYDHKGIV
PNDESPQGDDPIVNELLRYTYFPPPEVTRLNGTYRVGVRLLETKLSFANGPSQNNYTYFYNSFVSGCRYW
DEEADAWSAAGCTVGNLTSIYSTQCLCNHLTSFGSDMVVPPNTIDFSTVFSLENFIEAIPVFVTVGLVLLLY
LGVIIWARRADKKDVIKWGAQPLEDNLPSDQYHYQITVYTGGNKNAETTSRVSLIISGDDGDTGVRRL
WDGKRKELEKASINNYILSTEGCLGPLSFLRIWHDNTGSGKKRSWFLDQVQVKDLQTGELFCFLCDRW
LAVEQDDGLVERIIPVAGLREMTAFKHLFNSSVKKKLANDHLWVSVVKRPTRSNFTRVQRVSACAALLFL
TMITNCMFFKADENKENVRAISFGPFEFTLHQLFISVVTTCIVFPPSLLIITIFRRVKPKSNGVMQSNQRFSQ
KSSKFKWRRYNPGSSLWDNQPSRAQRMKESLRNILSQHQKTKYQDDNLDELGLPKKKKPWMLPHWCI
YIAWTLCITAILASAFFTVLYSMQWGRKKSLDWLSTFVLSFFQSVIVVQPVKVLLLAAFLSCVMRKPDLDN
EDEVFDECHNALPGPDESMDKANARNIDDILIQRKSAVTHMQPPDMASLEEARQLRMKEMAMEAILKE
CIVYFFFILVVFFLSYQARDSDSYQFANNIKNQFVRGSTPLDSISKPDSVWTYLRNTVLPNMYSNNWYNGR
ELTDWRELLQIDDRCTIRVGVARLRLQRIKEDSCEVQKVFRPIFDHCRNNYNWMDDDTKDYLPGWVVP
TAENRHRQGVIDCYSEALEFENETSLRDEESPWIYRNSIELKSAPYAGTHAMYKGGGYTFTFKRTPEATK
RLLDEMEQYEWVNPNARAMFIEFTLYNGNVNLFGSVILLVEWIAGGGATVRHEVKVFRLNNYVGSFGV
IVILFEVLFLAFTVYFLVKEITKLKKERMKYFKEFWNMLELSTLILSLVVIAMYALKAIFGKVALHVLEQSE
SGEYVNFVTIGVWDEMFTFLLAFVVYLTTVKFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFC
MFAYVIFGTQMGTYGNFIKTMESLFSFALGSFDFDAMRQASNIGVLFFILYVFVVFIGLMGMFLTIIADAF
TAVKEDTALRSNDYENVDFILGKIKGVFGW

- **PKD1-1_T5_i1 Version 1** (new residues)

MSEGFYVLQHYGKDYEVGVAKGNKLWTVSLFVHEDALVYFVLTWSEQGLDVYVDGILAGADAKGKA
RVFTSPAYDPFPEIVVGYANDRPYTGKTINEHVVNVTHWSAYYHQEEVEHLQGEAPKDLYHKGCVSTL
PQGVDDGVNVKDVDTKGVYDRVAKCRRKCAEKGAYVAFMKGGEMCYCYIDGFTMDFLDGMSTSTSC
SDYDVFWVKSHETEPDYPIELEVEVIPSTPKPYIRPEETVTFKSTFTTDEYVTFTYDFGDGTDKTSKNPVIS
HSWEKEGSFLVKITATTRLMSETVLKTINITYVEEGVAPEQVLVKMSPKDSSLVEFDLTTIGKYETECTIE
FGDGESEDVSVKDYIEQLSLSHDYKTPGFYDVVYSCENEYGVTADQGHAFAVKHEIEYQNEPRFH

- **PKD1-1_T5_i1 Version 2** (new residues and colored according to nucleotide sequence)

MRLNTRMNHAFTDIRLEVVGADSHQDVFGLKLNGLQVAANTTSDNVIIQANKFLYSGEHIVEVTAQDD
EVRRRIFNIEEKIEKVAILTNNSQVLPGEEVTFTFQIFKGDHIHVLIEYGDGASERVYIPQGGQPAAITRV
HKYESLNHYKVKVTAVNYISKVIKEHDISIERPLMSANMHAQNVTVLLTETIITIKVDENSPATPVDVMF
DYGDGKRETVRLGSHRSKAEPMVHSHIYSEYGIYVIHAIIRNNISEIEVEALHQVGENITMLDLYANRYRL
LDGEEINITMDAPYGYPLVYNLDMGDGTTLTVRRPLGYTGAFESDPIASTTGKTYYPPDFAVTEVTKTE
NTTSVILDRQRRDVNIMPTTVIVPGEGEVLAVMSTPAPMAADSDMSQNRGQRNEKVTVTYRYSRSGTFN
VMATVDNPFGGESTYLCPDIVVLPANQKQPDCSGLTVSFHNATSKSSPLVIQRSKSLTIPTDVQLNCQKD
ASQYQVDYTWKAARHVMGDLWRPELEVCHSQVPEANLVIPTNTLWYGLYRIKVITGIRLPHPTRNRRST
NTTKGATNSILNSDPSLITAEKDVYLKVEKTPLVAQLRNSSKKLDVFANTAVTISGTGSFDPDVVGKHSN
KTGMTGAIICYPHSRDAEMKKLSSKELLALTTLIQNNSRNTVSLYDTDCFVNATNVQMLDWDVTFFGDL
LVKDKDLTFRLFVWKDDREATTSQIVKVLQDLDGDNLFANLDNLLNSGDTSAALYVLGHFTAGLNGG
GEEDGEKNNTDVKNERSELRGKLIGVLGRVADQVGNFDQLTKTTDVLSGVTEVRDEVTEDSRASAAD
AMSSGAAKTVNLTNDAPLDQVAGLAGGMVGVAGNVLPFADEEEETVIEAIPAASGVTTIKYTTVETKTL
ATTIKPPSNFEPTTQTSVTTAYVPTTAAPTVAASDNDTILVNGTWVTATPVTTVVSTMVVPSVTTVTIKSP
KQTVTNTKVTSYKTQQVTTTTVAYTTTAPPTEATTQNIIANLKTTMPPGLAAVVEEDVPPCCSTKGGLC
AWIEPPCEFLSEDDRKKQVWDDNMCNALHPDDLKEIAWRYGLFKAQQKRMLFARRKRTTQGAMGAV
DNVGTAVLGKADKDTNVSISTPKMKLELAKVEPGQNGTMNGIAFPGSALGDTEGAGASSMESDNAFT
YSDSAKSVKGKVKSLAFKDKDGNEIKMNGTAEPFVIWMDGPGDSFPDPKQATFATSTQVLKFNYHTIN
TTYNDSSIHFMIRPRSPHDVYEVYIKWENFTLMNSSKENGNVGFPNETWYDHKGIVPNDESPQGDDPIV
NELLRYTYFPPPEVTRLNGTYRVGVRLLETKLSFANGPSQNNYTYFYNSFVSGCRYWDEEADAWSAAGC
TVGNLTSIYSTQCLCNHLTSFGSDMVVPPNTIDFSTVFSLENFIEAIPVFVTVGLVLLLYLGVIIWARRADK
KDVIKWGAQPLEDNLPSDQYHYQITVYTGGNKNAETTSRVSLIISGDDGDTGVRRLWDGKRKELEKAS
INNYILSTEGCLGPLSFLRIWHDNTGSGKKRSWFLDQVQVKDLQTGELFCFLCDRWLAVEQDDGLVERI
IPVAGLREMTAFKHLFNSSVKKKLANDHLWVSVVKRPTRSNFTRVQRVSACAALLFLTMITNCMFFKAD

ENKENVRAISFGPFEFTLHQLFISVVTTCIVFPPSLLIITIFRRVKPKSNGVMQSNQRFSQKSSKFKWRRYNP
GSSLWDNQPSRAQRMKESLRNILSQHQKTKYQDDNLDELGLPKKKKPWMLPHWCIYIAWTLCITAILAS
AFFTVLYSMQWGRKKSLDWLSTFVLSFFQSVIVVQPVKVLLLAAFLSCVMRKPDLDNEDEVFDECHNAL
PGPDESMDKANARNIDDILIQRKSAVTHMQPPDMASLEEARQLRMKEMAMEAILKECIVYFFFILVVFFL
SYQARDSDSYQFANNIKNQFVRGSTPLDSISKPDSVWTYLRNTVLPNMYSNNWYNGRELTDWRELLQID
DRCTIRVGVARLRLQRIKEDSCEVQKVFRPIFDHCRNNYNWMDDDTKDYLPGWVVPTAENRHRQGVI
DCYSEALEFENETSLRDEESPWIYRNSIELKSAPYAGTHAMYKGGGYTFTFKRTPEATKRLLDEMEQYE
WVNPNARAMFIEFTLYNGNVNLFGSVILLVEWIAGGGATVRHEVKVFRLNNYVGSFGVIVILFEVLFLAF
TVYFLVKEITKLKKERMKYFKEFWNMLELSTLILSLVVIAMYALKAIFGKVALHVLEQSESGEYVNFVTIG
VWDEMFTFLLAFVVYLTTVKFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFCMFAYVIFGTQ
MGTYGNFIKTMESLFSFALGSFDFDAMRQASNIGVLFFILYVFVVFIGLMGMFLTIIADAFTAVKEDTALR
SNDYENVDFILGKIKGVFGW

Sequences and data used for phylogenetic analyses

Sequences and accession IDs used for TRPP2/PKD1 phylogenies

Table 0-7 Accession numbers of sequences used in phylogenetic analysis

Species	Accession ID	Source Database
Homo sapiens	5669804	GenBank
Strongylocentrotus purpuratus	48094207	GenBank
Chlamydomonas reinhardtii	148832347	GenBank
Schmidtea mediterranea	847028803	GenBank
Schmidtea mediterranea	847028815	GenBank
Schmidtea mediterranea	847028821	GenBank
Dugesia japonica	675439102	GenBank
Tribolium castaneum	270002831	GenBank
Crassostrea gigas	405953847	GenBank
Crassostrea gigas	405968482	GenBank
Crassostrea gigas	405970250	GenBank
Capitella teleta	443713927	GenBank
Capitella teleta	443717936	GenBank
Capitella teleta	443732271	GenBank
Ophiophagus hannah	565304042	GenBank
Ramazzottius varieornatus	1101394602	GenBank
Corticium candelabrum	685230872	GenBank
Stegodyphus mimosarum	675368552	GenBank
Stegodyphus mimosarum	675370064	GenBank
Toxocara canis	734563164	GenBank
Octopus bimaculoides	918287730	GenBank
Octopus bimaculoides	918292252	GenBank
Octopus bimaculoides	918302224	GenBank
Exaiptasia pallida	999981301	GenBank
Exaiptasia pallida	999981318	GenBank
Exaiptasia pallida	999987444	GenBank
Homo sapiens	4505835	GenBank
Danio rerio	50539686	GenBank
Homo sapiens	205360954	GenBank
Mus musculus	115583681	GenBank
Strongylocentrotus purpuratus	148539638	GenBank
Strongylocentrotus purpuratus	148539556	GenBank
Homo sapiens	359465612	GenBank
Homo sapiens	5174633	GenBank
Mus musculus	115583675	GenBank
Homo sapiens	116006951	GenBank

Caenorhabditis elegans	392891400	GenBank
Caenorhabditis elegans	71999568	GenBank
Drosophila melanogaster	116008215	GenBank
Homo sapiens	19923084	GenBank
Homo sapiens	31559825	GenBank
Strongylocentrotus purpuratus	47551289	GenBank
Hypsibius dujardini	1174757226	GenBank
Mizuhopecten yessoensis	1205898338	GenBank
Mizuhopecten yessoensis	1205910066	GenBank
Macrostomum lignano	1236532622	GenBank
Pristionchus pacificus	1253289448	GenBank
Nematostella vectensis	156394387	GenBank
Nematostella vectensis	156395095	GenBank
Nematostella vectensis	156401013	GenBank
Nematostella vectensis	156406861	GenBank
Trichoplax adhaerens	196000368	GenBank
Ciona intestinalis	198420844	GenBank
Hydra vulgaris	828228329	GenBank
Branchiostoma floridae	260785879	GenBank
Branchiostoma floridae	260798008	GenBank
Branchiostoma floridae	260801585	GenBank
Branchiostoma floridae	260802732	GenBank
Branchiostoma floridae	260823633	GenBank
Branchiostoma floridae	260824333	GenBank
Branchiostoma floridae	260827108	GenBank
Saccoglossus kowalevskii	291224264	GenBank
Saccoglossus kowalevskii	291234667	GenBank
Anolis carolinensis	1033394660	GenBank
Dictyostelium fasciculatum	470247587	GenBank
Aplysia californica	871204620	GenBank
Aplysia californica	871233656	GenBank
Aplysia californica	871254812	GenBank
Guillardia theta CCMP2712	551655682	GenBank
Guillardia theta CCMP2712	551658247	GenBank
Latimeria chalumnae	556956196	GenBank
Latimeria chalumnae	557013709	GenBank
Saccoglossus kowalevskii	585644120	GenBank
Saccoglossus kowalevskii	585651646	GenBank
Saccoglossus kowalevskii	585652993	GenBank
Saccoglossus kowalevskii	585678603	GenBank
Saccoglossus kowalevskii	585688456	GenBank
Saccoglossus kowalevskii	585696579	GenBank

Chrysemys picta bellii	641798600	GenBank
Aphanomyces invadans	673034662	GenBank
Helobdella robusta	675861632	GenBank
Helobdella robusta	675875990	GenBank
Lottia gigantea	676439278	GenBank
Lottia gigantea	676474191	GenBank
Lottia gigantea	676479791	GenBank
Opisthorchis viverrini	684373764	GenBank
Phytophthora sojae	695457291	GenBank
Crassostrea gigas	762156788	GenBank
Crassostrea gigas	762157119	GenBank
Crassostrea gigas	1139740091	GenBank
Crassostrea gigas	762107655	GenBank
Crassostrea gigas	762108557	GenBank
Strongylocentrotus purpuratus	780039266	GenBank
Strongylocentrotus purpuratus	780160480	GenBank
Hydra vulgaris	828192580	GenBank
Hydra vulgaris	828232853	GenBank
Aplysia californica	871215788	GenBank
Aplysia californica	871242653	GenBank
Aplysia californica	871245685	GenBank
Aplysia californica	871263454	GenBank
Biomphalaria glabrata	908506947	GenBank
Biomphalaria glabrata	908475875	GenBank
Biomphalaria glabrata	908481541	GenBank
Biomphalaria glabrata	908400046	GenBank
Lingula anatina	919098890	GenBank
Lingula anatina	919100788	GenBank
Lingula anatina	918996724	GenBank
Lingula anatina	919018042	GenBank
Lingula anatina	919023136	GenBank
Lingula anatina	919052864	GenBank
Lingula anatina	919066838	GenBank
Lingula anatina	919074653	GenBank
Lingula anatina	919076662	GenBank
Lingula anatina	919076678	GenBank
Lingula anatina	919082289	GenBank
Thecamonas trahens ATCC 50062	923132353	GenBank
Thecamonas trahens ATCC 50062	923138555	GenBank
Limulus polyphemus	926626589	GenBank

Latimeria chalumnae	942132194	GenBank
Latimeria chalumnae	942189627	GenBank
Priapulus caudatus	957833931	GenBank
Octopus bimaculoides	961088331	GenBank
Acropora digitifera	1005492071	GenBank
Acropora digitifera	1005450057	GenBank
Acropora digitifera	1005458745	GenBank
Acropora digitifera	1005475249	GenBank
Acropora digitifera	1005480473	GenBank
Acropora digitifera	1005483664	GenBank
Parasteatoda tepidariorum	1009565777	GenBank
Mus musculus	1039738876	GenBank
Danio rerio	1207182030	GenBank
Hyalella azteca	1067095036	GenBank
Hyalella azteca	1067079595	GenBank
Galendromus occidentalis	1078802969	GenBank
Ciona intestinalis	1087812391	GenBank
Ciona intestinalis	1087817330	GenBank
Ciona intestinalis	1087817335	GenBank
Branchiostoma belcheri	1126169400	GenBank
Branchiostoma belcheri	1126215987	GenBank
Crassostrea gigas	1139813591	GenBank
Crassostrea gigas	1139824403	GenBank
Crassostrea gigas	1139824617	GenBank
Rhincodon typus	1160139095	GenBank
Rhincodon typus	1160100767	GenBank
Rhincodon typus	1160116546	GenBank
Exaiptasia pallida	1191059136	GenBank
Danio rerio	1207165955	GenBank
Danio rerio	1207107604	GenBank
Mizuhopecten yessoensis	1207917521	GenBank
Acanthaster planci	1229127695	GenBank
Acanthaster planci	1229145338	GenBank
Acanthaster planci	1229159547	GenBank
Acanthaster planci	1229186594	GenBank
Acanthaster planci	1229191925	GenBank
Limulus polyphemus	1238854278	GenBank
Limulus polyphemus	1238871302	GenBank
Crassostrea virginica	1242857867	GenBank
Dictyostelium discoideum AX4	66823157	GenBank
Danio rerio	528498254	GenBank
Planaria torva	v3_32544_1_1	PlanMine

Planaria torva	v3_35851_1_1_	PlanMine
Clytia hemisphaerica	unigene014425	Compagen
Bugula neritina	contig_206993	Compagen
Petromyzon marinus	ENSPMAP00000001872	Ensembl
Petromyzon marinus	ENSPMAP00000007053	Ensembl
Platynereis dumerilii	Contig9160	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig14805	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig14759	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig15290	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig12681	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig9951	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig9950	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig15048	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig688	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig1457	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig14675	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	Contig8524	Transcriptome assembly (Jekely Lab)
Platynereis dumerilii	lcl|comp408112_c0_seq1	Transcriptome assembly (Jekely Lab)

Platynereis dumerilii PKD2 and PKD1 aminoacid sequences from transcriptome assembly

>**PKD21** (Contig8524)
MSRPATAQRPMSAISRRSNKSAWEPDEGRPDSRQQPYDDDMAPMENDLYDSSNMSVDSRQPVVHVAENQNGCWDKFKRGVRSLWATRQTEDTGEDREMHVKTTLRELVVYLVFLVILCVVTFGMTSPTMYYYTQVM
SQLFLDTPSGNGGVTFRSLTTMADFWTFSEHALCDGLYWETWYNQRNVSDKELGYIFYENKLLGVPRVRQLKVRNDSCTVHDDFKKEIKACYDAYAQSIEDKAPFGKMNGTAWTYKTEKELDGSGHWGLISSYGGG
GFYEDLDTTKAKSAAIIADLKQNLWLDRGTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVTVGDYFVMACEGIFVLFILYYMVEESIEIKKHKLNYFKSFWNILDILVIIISIVCIAFNIYRTLAVG
KKLDQLLANPNMFPDFEFLSYWQVQFNSAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLIFGTQVKDFSSFDDAIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNM
FLAIINDTYSEVKAELANQKSEFEMTDYFKKGYDRMLDKLAFKRDKIVDIQKAINSADHNQDKQLDFDEWRHDLKKRGYAEAEIEAVFAKYDLDGDRVLDEGEQMRMNRELEGQKDQLNKEREELEDQRDDDP
NGGGPMVTNGVSFDEFNVLVRRVDRMEHSIGSIVSKIDAVLVKLEAMERAKAKRRETMSKILESITESDGASDEAKREQMEQLVRQELERWDTDANMKGASGGKSSSRPGSSAPSSGNSEV

>PKD2-2 Artificial fusion (|comp408112_c0_seq1)
MTEPSSSDSRSRHPGWEGGGVDNPAMDNSTSPSRQSGYLDDEIGPAENDLYESSVQVHSEAAPPVDVPVHPSVFDKFSRGVRSLWATRQTENTDKDRELHIKTTLRELVIYIVFLIVLCIITFGMTSPTTYYYTKVMSELF
LDTSFQGRLTTFRSMTTMHDFWRFSESSLVNGLYWETWYNGKNVSEDELGYIFYENKVLGVPRIRQLKVRNDSCTIHDDFRDTITACFDAYDKGIEDRNPFGKMNGTAWEYHSESELGGSSQRGRLTSYGGGGYIQDL
AGTKAESLNIILDLRENLWLDRGTRAVFVDFTVYNANINLFCVVKLMVEFPATGGAIPSWNFRTVKLIRYVTVQDFFIMACECIFCLFILYYTIEEVIEIKKHKLKYLKSFWNLLDILVIGISLVCIGFSIYRTIAVRDTLDQLL
QKPGEFADFEFLSFWQVQFNSALAVTVFFAWVKVFKYISFNKTMTQLSRTLGSCAKDLAGFAVMFFIIFLAFAQLGYLIFGTEVKDFSTFGNAVFTLLRIILGDFNFHELEEAERILGPVYFILYVFFVFFVLLNMFLAIIND
TYSAVKEDLANQKSEFEMRIQQNEGETQLQERQDHGIQKAIQAADINQDRMLDFEEWRRDMKTRGYGEAEIEAVFAKYDQDADRVLDENEQSKINADLEGQRVAVNKEYEDLEERGNRPGFDSGDESDDEDSGT
KSSRTGRVINGVSYDEFTVLSRRVDRMEHSIGSIVSKIDAVLVKLEAMERAKSKRRETMSKLLDSITEADGSSDESKREQMEKLVREELERWDSESSISQPPGGARGMSPGGSSSGNSGGVRHAPDPLPQSRTTVTTLKGPT
AVAMHNSGFLLSSIALSQTPVLSLDFCPKDVYRIRDNSVSAQRNAFVKAMNGASFGERRHTTLAALMRMRTHSPFFSSRFMKKHVFG

>PKD1-1 (Contig15048)
MSEGFYVLQHYGKDYEVGVAKGNKLWTVSLFVHEDALVYFVLTWSEQGLDVYVDGILAGADAKGKARVFTSPAYDPFPEIVVGYANDRPYTGKTINEHVVNVTHWSAYYHQEEVEHLQGEAPKDLYHKGCVSTLP
QGVDDGVNVKDVDTKGVYDRVAKCRRKCAEKGAYVAFMKGGEMCYCYIDGFTMDFLDGMSTSTSCSDYDVFWVKSHETEPDYPIELEVEVIPSTPKPYIRPEETVTFKSTFTTDEYVTFTYDFGDGTDKTSKNPVISH
SWEKEGSFLVKITATTRLMSETVLKTINITYVEEGVAPEQVLVKMSPKDSSLVEFDLTTIGKYETECTIEFGDGESEDVSVKDYIEQLSLSHDYKTPGFYDVVYSCENEYGVTADQGHAFAVKHEIEYQNEPRFTDIRLEV
VGADSHQDVFGLKLNGLQVAANTTSDNVIIQANKFLYSGEHIVEVTAQDDEVRRRIFNIEEKIEKVAILTNNSQVLPGEEVTFTFQIFKGDHIHVLIEYGDGASERVYIPQGGQPAAITRVHKYESLNHYKVKVTAVNYI
SKVIKEHDISIERPLMSANMHAQNVTVLLTETITITKVDENSPATPVDVMFDYGDGKRETVRLGSHRSKAEPMVHSHIYSEYGIYVIHAIIRNNISEIEVEALHQVGENITMLDLYANRYRLLDGEEINITMDAPYGYPLVY
NLDMGDGTTLTVRRPLGYTGAFESDPIASTTGKTYYPPDFAVTEVTKTENTTSVILDRQRRDVNIMPTTVIVPGEGEVLAVMSTPAPMAADSDMSQNRGQRNEKVTVTYRYSRSGTFNVMATVDNPFGGESTYLCPDIV
VLPANQKQPDCSGLTVSFHNATSKSSPLVIQRSKSLTIPTDVQLNCQKDASQYQVDYTWKAARHVMGDLWRPELEVCHSQVPEANLVIPTNTLWYGLYRIKVITGIRLPHPTRNRRSTNTTKGATNSILNSDPSLITAEKD
VYLKVEKTPLVAQLRNSSKKLDVFANTAVTISGTGSFDPDVVGKHSNKTGMTGAIICYPHSRDAEMKKLSSKELLALTTLIQNNSRNTVSLYDTDCFVNATNVQMLDWDVTFFGDLLVKDKDLTFRLFVWKDDREATT
SQIVKVLQDLDGDNLFANLDNLLNSGDTSAALYVLGHFTAGLNGGGEEDGEKNNTDVKNERSELRGKLIGVLGRVADQVGNFDQLTKTTDVLSGVTEVRDEVTEDSRASAADAMSSGAAKTVNLTNDAPLDQVAG
LAGGMVGVAGNVLPFADEEEETVIEAIPAASGVTTIKYTTVETKTLATTIKPPSNFEPTTQTSVITAYVPTTAAPTVAASDNDTILVNGTWVTATPVTTVVSTMVVPSVTTVTIKSPKQTVTNTKVTISYKTQQVTTTTVAY
TTTAPPTEATTQNIIANLKTTMPPGLAAVVEEDVPPCCSTKGGLCAWIEPPCEFLSEDDRKKQVWDDNMCNALHPDDLKEIAWRYGLFKAQQKRMLFARRKRITQGAMGAVDNVGTAVLGKADKDTNVSISTPKMK
LELAKVEPGQNGTMNGIAFPGSALGDTEGAGASSMESDNAFTYSDSAKSVKGKVKSLAFKDKDGNEIKMNGTAEPFVIWMDGPGDSFPDPKQATFATSTQVLKFNYHTINTTYNDSSIHFMIRRPRSPHDVYEVYIKWE
NFTLMNSSKENGNVGFPNETWYDHKGIVPNDESPQGDDPIVNELLRYTYFPPPEVTRLNGTYRVGVRLLETKLSFANGPSQNNYTYFYNSFVSGCRYWDEEADAWSAAGCTVGNLTSIYSTQCLCNHLTSFGSDMVVP
PNTIDFSTVFSLENFIEAIPVFVTVGLVLLLYLGVIIWARRADKKDVIKWGAQPLEDNLPSDQYHYQITVYTGGNKNAETTSRVSLIISGDDGDTGVRRLWDGKRKELEKASINNYILSTEGCLGPLSFLRIWHDNTGSGK
KRSWFLDQVQVKDLQTGELFCFLCDRWLAVEQDDGLVERIIPVAGLREMTAFKHLFNSSVKKKLANDHLWVSVVKRPTRSNFTRVQRVSACAALLFLTMITNCMFFKADENKENVRAISFGPFEFTLHQLFISVVTTCIV
FPPSLLIITIFRRVKPKSNGVMQSNQRFSQKSSKFKWRRYNPGSSLWDNQPSRAQRMKESLRNILSQHQKTKYQDDNLDELGLPKKKKPWMLPHWCIYIAWTLCITAILASAFFTVLYSMQWGRKKSLDWLSTFVLSFFQS
VIVVQPVKVLLLAAFLSCVMRKPDLDNEDEVFDECHNALPGPDESMDKANARNIDDILIQRKSAVTHMQPPDMASLEEARQLRMKEMAMEAILKECIVYFFFILVVFFLSYQARDSDSYQFANNIKNQFVRGSTPLDSIS
KPDSVWTYLRNTVLPNMYSNNWYNGRELTDWRELLQIDDRCTIRVGVARLRLQRIKEDSCEVQKVFRPIFDHCRNNYNWMDDDTKDYLPGWVVPTAENRHRQGVIDCYSEALEFENETSLRDEESPWIYRNSIELKSA
PYAGTHAMYKGGGYTFTFKRTPEATKRLLDEMEQYEWVNPNARAMFIEFTLYNGNVNLFGSVILLVEWIAGGGATVRHEVKVFRLNNYVGSFGVIVILFEVLFLAFTVYFLVKEITKLKKERMKYFKEFWNMLELSTLI

LSLVVIAMYALKAIFGKVALHVLEQSESGEYVNFVTIGVWDEMFTFLLAFVVYLTTVKFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFCMFAYVIFGTQMGTYGNFIKTMESLFSFALGSFDFDAMRQASNIG
VLFFILYVFVVFIGLMGMFLTIIADAFTAVKEDTALRSNDYENVDFILGKIKGVFGW

>PKD1L1_Cterm (Contig1457)
MMKKTKYRGHCTAQKQANSPNLSELSHNLAGQNVKSEVIRDLNQQPTYKPPAAYAHSNHTVEVFISILDSLGARTRVCSLPITVRPRSFSALNKTEEQILYDIVKGDESPLQIYVSSGDNRNVRLQTQLVAHRLNRLNSSG
EDPEQRNMRRQIREAMLQALHDLPMRDEYEVEQQTTAAVASSTAVTNELEVASLVFASDILEKAVRQVETLASRGKQLDTTNIAFLTIDGAANILAAVTTPDVAEQLEKDLTTPSPLPFVIEYEAPTQVEEVVQKSRNAVH
AVLKSEMQQWVAGENVFTKDTGNLLTEAGCFRASVAYTRQSQLTNFVLPSNMEQVLNTNVKVKQGNSSSNMEEDCYQAHMLTFRTNPYIYSKNRQRQVQTQVASLTLFTCDGQELSVDGLPADDLVNMHIPHQDLN
SQSMKSEMLTLEKHHMNVHQFNVSGANLDQSLHIHVELEPAKDKHVFPISILVSVDGPPNPENYFIRKNFGRELENLKVYVEANKLEAGKQYYIGLVDVEYNAGRPREGEVRQRQYRLWLWWGDCLYWKEADNEW
SPDGCHVSPAANFSHIQCNCNHLTSFGAHFSLAPNDLSFTDVEALIDLHENPVTFVFVLSIFVLYFILLAVSIKLDRHDSMKGSIVHLVDNRNVDKQTYMVTLETGFRKGAGTTAKVSLVLHGEQGMSEMRELVSESGQLL
FERNSRDTFFIVTLKESLGKVWKVQLWHNNAGQQPSWYLSRVTVKDLTAGHTYYFSCEKWFAVEEDDGKIEREIMLQDGGLGFGKVLWAKGTQYFADFHLWLSVFSRPSYSYRFTRAQRLTCALSLLLSYMAVNAAWYQ
FREKEYRGEFGLIDLSWRSVFVGIITALLVLPVNFHVLIFRRSKVSKVRLEEFPEKKHFANKRDSLSSLDSTVRLLQPVVTQSIFDQSMINWQNLQEWAQRQWAKRYQQSVNSSLQNVKPPNPQQVQMADTVPPGATDA
NVAGNVAGNVASNDQATPDIAAGQTDPVDLAAASPEVGDSAAASNRSSASLGPPSQRRRRGQEDQPRAEDATGQSREPLLRNRSTHSEPSGHSSNNSFSIATSTEGSRHQPPQLLLIQGGCISYLRSEKKKERCFLPSWCIY
IAWILSVVIILGCATTTILYGFRFGRARSIMWLQSLFFSFIQCVFTITHPFIILIISVVTAIRYHDDLSVFDHHNDGVLETFQEVVNNDGSPPPYTEKEEEFDLATGIAQRQRSRYLRFARPPQEKHLIEARKTTLKQRKTGAVFSE
TLSFSILLFLLLVVAFGKDSQTPFRLNSTLETLLTKTGSLPFKDIQTVDQWWQWSHNDLLDLVYADGLGKDMVVGNSQLVGQVSIRQKRVKPSLCGLPTAYTQFSTCHGGYSKSASEASWFAVDNFTMPYKEKDQTGAS
CVVGRFAYYDSAGHQVFLPRDKTSAAAMLEDLYNYEWLGHQTRAVIVEFVLFHAPSNLFSSVHLVLEIPPLGRAFPYTDITSAYVYKYISHFDNFILACELLLLIFTMILLHSEVESAIRMRKEYIKSPWNCVQVTVLCLLMA
YIACYLYRYVLVGNVVSALRKSYNSEFVDTSFIAYWDHLLNSFLGFIIFLLLVMSLRLLRNQPLFSKYGAIYKKTVHEIITFGCLFLVWVCAYTSLGVALYSAISLDFSRWDRGLGTVVGLMTRHYDLQELDKMDPHGSRLLV
ISFLLLALSSLTGYIVAVLIYHLKRHHSSGLPPLSTSQALRYLWNRLVMFAQLEQPRYVHEDDGGLPPEFTMAEIEYQVEEMLFKMNALAGTSGLPDKPSQYFTDSDCTYPSPGEDAVSSGSEPRFDDGRFERRINKIEDNL
VSYEPHLAQLLQLDAIGARILTAEHEKQIRSQLELEIFRRLQLQRPDVRPHHTNRLPSSGSNSSGKVQQQQASNSSSGCQASAEASPSPASPPGHSHKSFFLPTANTAGQSPPPSLGFTPKTNASLPCCVAAAT

>PC1 (Contig14675)
MMVTLNVSIAGSDVPVGSDVLKLNFEPPTVSQANIECLSCRESNIYTISRTKPIKLLGSCTNCGTDAHFEWSVVDEDMKEIHFTAENSTTGLHGANFAIKKNVLNSNKGHHFRLKVWSQRQSKPGYARLYLSPNHPPEGG
TCTLTIGDDPMQPVEVMHTQVVKALTDMVYIECSGWNDPDDYSTLTYHVFVQRAGLVGLITGEWYPLYRGTQSSGSFYLSQWELDQHVYVVVEVIDAMGGVAEGFGNPIKLHVVAPRADELKNITLDTLKILRPHND
PKLLRQYCLASAIQMNILYDDPAYQDDRATLQKEILKTSNDTAHTANSLDDAQQFIYILQHLTERFELSNEDNSDTSAMILELTAITGALSRNANPSTDGDLFLPEQLIETVGNMIQVDYINTNINTITMSQELVSQYTIE
YFNVAHLVMVLYSLTQISGERPFEFNTVSLNSMGTRSYLKDLPKTQKFSNFTISLPNTDNIGLQTREVFQSVMTLQFNPFNWGYNNPITSQLCSVSMYNTNGIPINVKDLNALDGMFEITLPASVGSEVPDFDDSFNVTLQ
PGQTLTSALNLTGDGAAFINFRICMFSSVVMPNKKSFSATNIGQHGLHPPKGSVVAYLSEVSEPLYDQYKKKLFIETDTESSTFFISERVEKENSLHYLTLTNSDLEKSITVTMTLFYSQCLFYNDTTDKWEDNGCEVTDAS
TATSTTCRCNHLTLFGSSVQHAPINIHFIELERLNIKDNPVVIILISVIAGVYILALLFGKYQDNKTLKLTSVIPLCGKDGVFKYEINITTGRYPSAGTSAHVGIIIHGLEGRSSARHLTKTGAFKRNGDDTFMVAMDHCLGH
LTKVKVWHDNSGLSPSWYVHRIIIRDLQSGQRYLFHANKWLSLETKEPVIEYEIPVTVDGELVKFETYFKESTAVGFADSHMWLSLINRPPRSRFTNVQRLTCCVTLLYCFMMTNVIWYGVLSAQPSELDYALGHIAIDWQ
DLVIGALAGLLAFPPIYVLTVIFKRSRIKAKVRDFDEKSETAQTIEIDAHCDVSSAAGSSFISRSSRGPPSDRQLNRSVSDVVRRSKPGIVSRLRRTGALTKLFGVDGLDDMTDTSSAVSGFTNNNNTIKDDLAKQIPGANRPL
YSKDGIMEWPQQPLKWWNNSPSTGKRISTNSNISDDSGILEGSSESSRPPDRCHPRIVLRRPTAQTVCLDASIDEKEEESYVSPVPPRAQKHKQTKAEDIPLKTFTSDKNVQTETRSQGSNTSKSGLMVLPRRRRRSLYRTP
QPAPEPILEHQVSEIAETTLFNPRSETNVEIQTAMTEIQRNMGERSPEKRRCIARKNHSPNTSRGHKRDASDASNDSRGSQRSNKSEQSGQVNEAFELDETQPKGHRRQESGVSFTSDISSASEMNANPLHKTKSSRISKR
WDNIVIPEDMWYSDGTRSGRYDSDESGSWTLPPVYHSNHLNNMPNYYPQEFISDQGRSTSMTLPHYFIYVAYLLCFALSVFSIIVTLLYGNSFGRHLAQKWLIAFFISLIGSIFIIEPLKVIIVALVMSTFKPSLYEYENIGDPYA
TVQMNETLKDVSIRPVGGFALMQAKEETKKVQRMRLILRQMTAYCLFFWIMVVLCYTRHLDTSYQLNQAVQSAYGWTPTVNEDALNQKSWYENVFFDVLQKLYRPNNDNFQMYMVGIPRLRQIRSKIGTCIIAPVNA
PYTELPNGDTITGCRTKNEDKASYLEVWLVPEGAANYSWIHTDAEKALGQTYRGKTAEYHGGGYIINLASSKSATANVIHLLKMQNWIDRSTIAVFLEATFYYPPKNTYTSLILAAEFPLSGGAYYSSSVKTLELKPWISG
VVDSMTVFLGLLAATIMYLLIHIGLQIKKQKIEYLKEFWNCYEFVMFLLAVLSIATVVLKIEEDTEKTAQIAAKQREFISLAKNLELDTRWSIINAFIMFMLTIWIVKQIRFIKSWSLHGNTLRVSAPALLGISFILTLVILAYAQ
LGYLWYGPIIGAFRSFGSSFCNLVVLIRGSFTELDGMLEYNEVFTQLYVMSFVSITLVGIMGLVIGILNSTYRDLKKQANYYSNMDLKDYEMIEFTIRRFKRWLGLTKEKPAFRRVKFQGLPSASTASSMKSGSGTVTGSIGS
SGSSGMSEVMDILEAESQAYLDGLLERALPTWNNMLHQLNLVEEVDQREKATERHVYEKLLKVRIRGMKEGNYVEDDAIPGTSGEGACAFTNEAKVQYHPPRGDKNTPKALPHTPPPSRRSRRYSEQLSTKSKSNKIY
PMRAHSDMEYKSKSSQRSRALSDPSVQAKPAW

>PKD9160
MKEGHSVEQMQLPSVDSMLGNSDGEGNSTDTPYSIECKMSASSVNTHPNDISTMITNTQVMNFDIVDELGIRMDVTKSDDWLEIMLLHTNGKMNETFVNVTEDQKVQPIHQFQVLNASALFVGVRDIPIPKADAATTS
SPDDQTTAAEEATTAGGSSRRKKRSVGSGDNSTTTVVLVTTAAATAQNVVTANAQNNVAITSNTPAPAPVQPAPTPANQTAPNNTTNANQGEAKTPWCLKDPELMFTIYINEGTESPTPEKHLLKCTLPLKDAEIQAYI
DRYNKNKKWKIKDFDRCTCFFRQNQMNLTEPRPFSMSVVPMNITSNETTTDGLQYAVKVWTPSCQYFDKNESSYKTDGVEVGPMSTENATQCLTEHLTSFGGGFVVPMNSINLKESAFTKLNENPVVFAVMVSLMCL
YLVMLIYVRKKDKTDEVHAGATPLLDNNPMDKYLYELVFFTGSRKNAGTSAKVHCIISGENGDTEPRLIFDDKRAVLKRSGVDVFLMSVPQPLGPLTHLRIWHDNTGEYPSWYFERLQVTDLQTEDVYFYVCDRWLA
VEEDDGMIDRVMPVAGMEELTHFNHLFWIKAKRDLTDGHIWFSVFQRPKRSRFTRVQRLTCCLTLLFSSMMASILFYKSDENIQGKQVLRLGPLEITPQQIYVGFISSMIALPANIIIALIFRYAGDIPKCKKEKEQPPVQH
GSNVLVVDNHEHRKIAKPEDDDRKMHDELQRQSDYLDTGDIDQRLLVTPNHRPATASTSSSQGESYEGQGQDITALALPMPPPKKKKFQLHWGFKIFAYVFSFCIVAVTFWVTVEVAGVFGKKKSLQWLLAFVVSVV
ESIFLSQPIKVLVLAAFYALVIRKPDTDSDEAGTTDLRKNEEWLHKHFTEEDMNNPETMKLVASYKSALPEPPDEEYLQEARQMRFKEIKMNKIIWEIIFYFLFLWLLMMVAYGNRDPQTYNLYKHYQKLLVKASYDKV
NPRGTWDLQEVRSRAEFWDYVSQVLLPGLYDNKWYDGTYHSQLQGFMADGTGYLVGQPRLRQVRTVPEECHRKVKHLEKNCEKGYSIFTQDEGEYSFGWKKVNYSDPANNKPETSWIYRSWLDLSGYPYWGKYN
SYSGGGYTADLGTTLTEAQQKYEELRSNRWVDSNTRAIFIEFLVYNPSVNLFGISMVAMEAQESGAMYLYFKFHVMRLDRYVGPFMFVVMAGEILTLVFIIYFLVKEIKKLRKFGKAYFKDSWNIFETLILSLAIVTIVMYF
FRLVHSARQSKKLREEPNRFHNLQYAALWDESFTYLMAFMCSLSTLKFLKLLRFNRRMAMLSSTLRYAAGPIIQFLIMFLILFLAFVLASYIVFGMDLEDFSTFARSMATLFSAMLNKFNFYNLKEADRIFGPTILFLFTLMITF
ILINVFTILICEAFSIVKDDLSQQSNEYEIVDFVVDRFKMVTGLGEPKTKMPRRRRKVKEYNYDDYVEATDEDIQKEQERIDKMMAKLEFLIFHKDHKPGKPVSDPKGKFEDWRERKKKGILYR

>PKD12 (Contig14805)
MRTPILSTAVLVLLYNILSSVNGQNCPAPGCYEDIPRHFTYNPGNYQPTNMTTDICKGVCGVWDMKYAALFSGSFCFCANTLPASVTPSAATNCNTACSGDSASLTCGGLRHVNVFCVTPPAQDVKVTMSDVVEVLTETT
VNITTLKQSPGYVYKLDSAEDNGFVPLTSSVIQKSFSTAGLHKVSASVEDSGGLSAVAYGSRTISAQVPVEHVDLQCPEVWVATEELSCNSTIISGTLMTADMDYKDGKPHDNFPVKDAGWFTAGTGLVPMTGPATANK
VYIRFHSDFMNSGTLVAFEYYASAAGTANFVILRPTCTTNTYCHDSNTCVASCTSLSQATCTTGQVFCGYKTTCTANTLPTGSCAAAPNRHASLPANVEYEVVRVFTHAITAPGYGSHKLDPNTNPLEVQPGDVVGYYSG
TAVLGVSTAGAKFDYTVNTVTTTAAATKHAVTDLTLGAHRLLLRAIASKPSLVYYYHTFDNSGEYNVVLNVTNDLLPGKLVSKTKKIFIEVGINTTVISNPPYGVTGESTTIEVEPVSGCLNLTFEWNFGEGTFLNISRLGRF
FNYTWKARGEYNITVRVVNRLQSKNNDSTIIIEDRIRNMGLYGTHAATGKPYQITFNMTSGSDYACTWNWGDGTTFTTNSSNLLASGSVQTHNFPAEGWYNVTVFCRNNVSSGWAHKRVLSIDPPINISLIKTGAIQNT
MFKVDWKIDKGTNTTWTLLLQNVSYPIVWNEATKKGETTNFLNGLPVDRYLVGLHVINPVADEWFFKNFTVDKPITGQWFAPSNLIEILTGQPATFTADMTGGSSVIIDINYGDGGANGLFDTTQGGNVLIDWPGM
KNFTRTFNQPGSFNVTATFKNNFDTYLYYCAVDVYNRVENLTLMTNSPIPYLNNEALVLWYYSTPTYPPTKAFIELSYGNGDPDIKGVSFKPPGTKTGHTYKAIGRYNLHSNVSNKISWGLFNTSIHILRPVVSLELAVVP
QHSIKNTPVNFKIAMESGDDVTLYCNPGDGSPTVTKPRTQTSPYKWDICTHTYTATGVFLFVAIAENALGNTTKEYRIHVQNPVKSTFDLTSNFPQEFIGGSATVTYTLTYPPGDILPTDPSIILDYGDTPDVYPTAVPLNI
PGSRTLTFTKTYTTPGDYTTNFTILNLASMANFTVASGVYERISGVSCVAKQRKLIPQGGADSPGYGTNQLTFPLHNPLTAHFTVAAGTISKYIVKAATTPPGITEESKTQPVAVKLTQAGDYTLTCEAWNPLNSASATVN
VKMVEGIQGISISNRVPIKPGVETTFALSVASLGVDSCMIVDYGEPNGVKEYYGNTAGCASRLPSGFSVKGVIQSTMSLKHTYKNKGTYDCTVNGYTDVLQTTNSSKVSVTVSDKTCLSPLVTMKPANPRFYEPFSVPRNE
YVKVIASNSIDCQETLENVKEWRVFSINEATGADISEVDLKSVTTRTTANLVLNPRFLAYGTYRFGYKMYMKAGAEQFPVEESIYIKIAKTPLMASMIDGGMSIIKRGFNRLIKLSPRKYSFDPGEDPSGPQGLNFECWCRR
VDETFKTYRNGSLISTDASAFMTSAQLAADTAPDQGGCAKRGPGVLTLADHGEFSFNSSNMRVNESYEFQLRAWKDTREVWTKLIINLVDGDPPEIVQRCNGENSALCRPFPGGQFINPNNRMALKAFCVANCGNGSP
LTHQWYVERRRINGAWVEINITDHVQGSTSSQIALPTTLFMSDYAKYGEYMDYRVKVQVVVPGNTVVGESFTVLDMNIPPKPGTCTLDKTAYTIGEKAQFTFSSWVDGEGVEEATVEAAVNGETATLTFLNVPRTELV
LPLGGSVVLKCSVKDKYGAVTVLNMTSASVTMITAQQTATITNLKNIAKGVDVSLTKAGEEGDSVSILAKVAQVMTQVAAVSSTKLVDLKAAWAAEEARLQAALKNTSLSEAQKRDLIVQRNKAEQAYEELRDIVTQC
QEKSILAMEDLETSTPEGLKEQAAVLNTMDNPDHMTLIARKSLMKKGAAMAEGLKLNALSLPREDVEVIARGALTSIMTALHSTHSVTNNPTLTELETSRGSANYAYDSDWESDQDEFAADDYKGLMKWTVGVDAGI
QSKSNSKSGVQGLQTVALASLQTMVNGEEIINMTIGDGPAASQFTAQKQTLSKLINPIEQQSVGFQFPSNLLPRPKLPDGSFEDPILGLAALTSARNPLPTLGQGERVTTHTIDLTLTDENGTLLSFSNLSPPLDLFFPVDAS
KSTVYTQQPKLPALREGRPRRYFWYNSIDCLPYEGVRLRLTPLTTNTSNITYLVMIDNNQFPNCTLEGVCNTIDWLVLPNENTTNTSDPYMALFDNKKMGNGKYKAIIAVRELQGSELQAYKNDSTKIPKFTIPQDKVF
TTPYRTEFTCMGAFFFDSTNVEKGWHGEGLTVGSKTVPSKVHLQSTHLSLFGAGFAVQPNKIDWAFVFSNMDFFKNPTLYIVEIITCIIFIMLIWARYKDKQDKMKLGVCPLPDNEKEDKYLYEILVYTGLRTNGGTES
KVNFIISGEHDETDVRTFEDPERKIFKRGAVDAFLMATPHPLGHLNYLRVWHDNKGKGKYGSWYLHYVIVSDLQTRHRYYFICNRWLAAEEDDGSVDRLLAVAGEEEMTEFGHLFAHSTKKKLDDDHLWFSVVARP
PQSRFTRCQRVCCCLLLLYTTMLADAVFYETAASNSGKKVFSVGPFALTPQQFYVGIISNLIVFPINFLVVFLFRKSRPRTKRPSRIAEALRRMQEQKKQQKLSNSDLKAHSSMNFNEEDIPKDPDANPISTISRPMSAQSRRS
VVSDAFDPEQPPEEPYRPPTPLEEEEKKKKFSFPWWCVIIAWILLVLSVAVSVCFVTFYGIMFGDDKCKKWLTSMLVAFFTSVFFTQPIQVLVVAALLSVICKNPSLSEDDDGDDDEEDLELGNNEAYLHEMPREGYARP
RKIAYKPPDPEALEKARQQRMNEKQMYAVMREIIFYFIFLWMLLVISYSFRDPNAFLIKESFVKNFIRSGEDSEYNFKNLKTMKQFFDWSENVLAPELRAGKWYNGRWLPGQRGFINDRNNRIMGFATMRQLRVKKAK
CNLVKPMDKILRECNVAYAFYHEDTETRGVGWEPLYSNSTYNNSAYEYVHRSAKSLDSFPFWAVHHVYGGGGYVRELRGSTNTLKQHIRELHDGGWFDHYTRAVFIEFTVYNAQVNIFTICTLVAEFLPTGSLFTSYRF
EPVNLLGYSMDTAGFEIICQIIYMLYILFFIISEARELYRKRSAYFTEWWNVEVMMIIFLSLSGAVIFFYRLVMASKLTKKFEESGGNAYMKFQYVGYWNELLLYMIGWLVFLATIKFLRLLRFNRRMSMLASTLRNCAKSLA
NFALMFLVVFFAYVQFFYLLFFITISEFSTFLVASETCLQMILGRFNFFEMKGANPTLGPLFFLAYVISVALILINMFVTILNEAFADVRDDLTKQMNDYEIVEFVVKRLKSWTGVGAVVQGGKSQKEQDEYAIQMATLDA
MAKAAEGNDYNFNPNYKRTEDMEDFPDKVDRLVRYISNMYFPQKEWLEAIAIKKMEEKRRSHDRGPDYGRF

>PKD14759
MMTWYVVQIDPPGHTEPVQVQYMDDNFNHPVRGLPYGTYKLNFTIQMKNVEGTNNSALAYLKIVKTPLTVELTGNYSESVALGTPYALDALQESFDHDVEPWNKSAMVFRWTCERVNGTFNESELNMDEIFKDQT
MLNIVKYPDGCYGEPAGVMSHTAGLYPIPTDYLVPGSCYRYTLTIVKDTRSDTDTKFVCVPIYKAPEMQMRCRENCKNRKGTASRLVVAGDCYADCGTKFGGLRYEWTLFYAENDDSVVYTEYEYLSAMTFAGVHEK
SIFIEPNRLIPGKIHEMKLETYTMPNSLKSTVNFTFLTNQPAYGGYCGVEPREGIASNTTFVVRCDRWKEWRPDRPDAPLRYDFRVKLPSTDRDFPFAFGYKSFTEPTTFPVGDPTCFDNYQDGFQLDIVVYIIDTVNDPK
RYQTHVRVIPKPKPALEVMSAIMEASSQQGEGSPIDELLNNGSSQAAISMVYAMISELNVPNKVYYDGAADDETATNITNDKIKARTNLIKKVSQGGFTIEEPEALSQQNEVMRTVLIQKEELEMDALEVGTNILGNLTG

TFANFRQNGSVENPTNMASDLIGSTMTCMAASDYFYAQANKKANATTEEYMTTESPLNATTAAYNDTISAQEEMKRKAAAVMKQGIDNCYNILNIVSEDMEPGDPDRIMSKPEITAVVQTMDDKQMKKANLVNSTQ
GGAEMPPGEILFGRDAGCVEYKMIFLGINPFIWGNQTETGVESSVLDMQVQKCVVETKAGGSRKKRSVGPASTAIKPASPFVISLAPKNLPEPQVIAIPRAMGLVYHKIPISNPARPIQINVYPREHEILEVFAHINGNPSPS
NYSWNFTFPASQYVDIVNATYDENFTLPEAAFTVTLSALQLEMALNGTLDQYKRSVINLNNLDDEGNGTGNGTSLNETTIQPTTTMAPTTTTTAPTTTRQIDLTTESFFDNSTTPFSNETNEAGNRTLIVGVRFGRCYPG
TDCEGGRESGVHYRVRIAEIKIQKWDPATEQWQEHVAEVEEDSTPDRIRFKSNLFGSLGAGMFSGPNTIDFSNLFDNIGAKLLENLPVVMTVIGFFLLYVPLLIYCRKKDKQDAAKWTTLPLVANVASASYRYEIHVYTG
IKRHSQTKSKVHFVLSGSEADTGVRQMDDGIRENFCRGDINLFLLNTQQDLGELMSMRIWHDGSGEGTDCGWFLDKVIVIDLQRNHWWVFHSNKWLDNTKGDRQTERVLKVSTFGSSSNTGVLLRNNLNRKFFDD
HVWLSVAFRNTPSVFTRAQRLSVCWAILYLTMVTNAMWYKGPDDITTGSSGGLFKIGPITITMQELTNSIMASLIVIPPILAITLFFTKSVPIPRKIKTFVDTEDKSYYTRRRKPGDDKAKEQYKIITPPTPAGERRKYPLPHW
CIHLGWLLIFLAISVSALFTIFYALEWGGVKSRAWLIAFFLSFFQSVVIVQPVKVLLISALLSFLFKKDTDVSEYEDSYGQQQTWFDIFTNPALIDTEINTKEIEEAKKKKEMDKKINNMMREIFMQAFFLMMLLIIAYGNQD
STVYRQNTDLRNRMQQYLKIDKIPEIHAWLKTNFTDEVYYYKNYLGKQVSAYDQKFIANTIGMRLGPPRLRQLRVKNETCEIPVPADIATKRCTHHYSETRTSILEKKNHSIGWMPYSDKGTPGGLTNLEKAFVYNTSS
QLRGLPFWGFLSTYDGGGYVANLGLTKPEAIKVIDGLNSSRWIDQYTRAVFVEFNVFNMQSGLFTQVIISFELPPTGAIHFWTDIRSVKLYRYNGGSGVIAIVTEILCTLVFIAMITYLIMDIVTQKLSWFKSFWNWTILASS
TFFIIALSIYAYRSVWTVKSVEYMMRHWGEVANFQIVTTADNIFLTFLGLATFLVMLQFLHILRYNKTIGVLNATVAQASDLIGMGFLVFVIFAAFLSLGYLGFGSRYQKYGSIKNTAIALCSNLLGNFDYQAIRLAVGIPG
QVFLITYLCTMIYLMINMFITLLCNYMEAVKGDESLIPEDHLVVDHFFENLRGLAMGDEEDEENDQRNSNQPVSDQGISNILFGSTKQADYSSGIPKKPRKILYASD

>PKD15290
MTIWLLKLSLTSIAGEGFFKITVHAKKAPECPSGSYRNGDLCVAPDYSKRAMLVNYFETACEHMEWTQRGYTYMVKEAQHMSHAYDVMSLSLSYTPPFYAGLREKNSLWVFDDGSNVTSNTPFNYFTKPAVANSHCGK
EYLETILTSSDCKLRGFCQTRAAIDPDDKCIPNWYHTYREKLPDNRRTHMIDGLSVGQCLNFCLYTLRNECLAVVYYPGDQECHLFNTTSLESELTYDVNARLHEWHCHDRSDVNVFRDLEGQFQYTIVVSSNYSSAYN
AYEAPYLPRNTWGWSSIAGDVDGKHYITVNFDREMSLIGMYALGGQRRTASYTKSMYVWYWNRYGWQGIVKTGSVDSLIDVGVNRYYGKFIRFNRRIETKTVRCLTSRNSGYSLLKFEYFGHIVKDPAQASDFPLGMA
TRQIGGYQIKASSFVSGAKTYPPSAFRYFSTTLGSIQYSWQPSTQLNQWVQIDLMGSCLITRALVSAGLETKCTKKYDMMTSMDGRNWTTVLTGASGNVFRFNIYRDILTNLTTPVVARYVRVIPTENHIAHDLKMELYGL
RMCSALIQKGLIPKLVYSKPRSFCWDNGNLTFAYKNEFNCSAWCIETSSTQTLVIDLKEIYTLTQVVLVITQQGNVSITANITFGFKTQDYVQDPKTHWLKNGTKSTYVLHHTRKARFVKVQLNLQQHEKFCVFADVYG
CSLNDCPASYMQHKKSCFRVYDMLVTRDEAQQVCEDDALSWYGKGQLAKLNEMSKVDAVHGLLDGEYLEAWVGATYNSGAWKWSDNDALSWASWAVQYSTSWTATFYSVNQRGMINVIKTLKKRFICELETQGK
NPGVVWLYQNEAVLNETGEEVPGDNEVSDCLSACQSGDLSSCASILRDSTLLNKGCSHMTWSRYSSNKGISQSPLKMTKYWELHNGGLCSRTIVGSIYAHTITASSNKDASTTPDKVAYGANPITFWQAASSDSSPYIQVE
FEQVMTAARILVAGNGKKLFYPQYVKSFSVNTSLDNIKWVAQMDGASQTYVFTGNYDAVTPYSVLLVSIPLFKYIKIFPLTWNIAPSMRIQIFGCHFVEPLTCNNAIRPTISVPGSPDFKPPCHDSWTVIMRHVNPSFSFRQS
WAGYKNGFGNTISDFWIGNILLAHLMKNRLFDLRLEIWTTTEDQFFSSEYSKVKVQNELQKFAISIQSHEKGLLEGVLLTNGTPFSTFDADNGGNGCSQSHAGGGFWFNKTSCSSGSLTGDLQTKTYWISARGSKVYAK
KVIMKTVPTKAQLDNLAALKSASRSSGDSTYPASLATDLQILQDRASGKCSLGVAGTDPWWRVDLEKPQQIFSLRIITVSDCCADNLHDFDIYIGSTSANGNFTDNGYKLCKKYTGTVSSGDVVTVNCTSVLTGRYVAIQI
PGASKVLSLCEVEVFGVKKEPKPIYTACERNSQCTQINTECRPKRGSMAEIYSCECRTYTRDVDAVCYPEADMEPNIQSVTVPQDEIVRNHTTEGVVNLKISHNEGAVLNNVSSGSNFDIDVFLTSEEPDVLAGNFTKSSA
SVSLDSKPKFMNGMVYATTIDVILNVTIHIDFVECFNLSHLCVNLTRGSTGFYFETNSSNNIHCISMKQYTQCRPEINLALSTISIPGSTEFVRGVTSDYPVSLTLSLTASSYGIVEPLPGDKNYVVHFLYTSTDITSYQQFLSAG
PANETKIDITNSRKSISQNSQVSVTVNVTSSPSDSECEQLKYICLAVVNANGASFNDSNKDNNVKCFDMATKKQCFPALSLNITSFSKVTGPKIVRGYNATIGVTFVIRNTDSTHNVVPVASPRENFAFELSLANSLASRRRK
RAATDVGPFTVVMFGNTSTGLGVNESKTLSGTVQFLIPRDKCSSLHYICGGLSAANGSSYTISGTSPISCFNFDPVKNCQGVPVTNPSLSSNASIIPTNSTIRTQVQWSAGTDIDFKIDFGDGHSFSWNWQSAGHIHNFSSFSV
NIAHTYVKYGTYALKLNISNEINSTVIIVNIVVEPILEYFTIFKDVWKPKPVPSIVNIEAGIISASLYNGPIRIFCDSQFDDNGSANSSRVNETQTVTLTPANITHTYTFDKPDANLTLNCSNHVSKILFAKRIILREPITNLTVTP
SVKILKTGSSVEFVFTMTSGSHAVFDINYGDVTANNIAHPNIRSYQQPHNLTHVYTTAGNYTVSAKAYNIFYNDTYSLADVIVVQNEVKNIQLVVPSVLEWSASPMSVIVRMNGSAPFPTDPYCHWMMNGSTVSHHYQA
ELKMGTDALTNVSIARSYIGRNVTFDARCYNLVSSVSLQKITGVFEKIVDLRIQGPNAVITTMPVTMKFTVVNGSNVHFEINDGAGNVYNQAHPSLLASASPFYNLTYGVIGNYTITVKATNPLGSETQTHNIVVQNAM
TDFTFLANDSVLYTPGEVDYKVIISSSQAPLKNIHCRWTYNNKSEEYIYYDTLNAGDTFMRQKTFDRSYLGLLNTSINCSNLVSKVVKKVQTEIILDAVILGSFETNGSVFTTNRTFFTVNITRFGVAPKSCFHLDLGDGITQ
IMFGLPVCATNATTLGIPFQQIAFNTMELEYDYEYQTVNNFEAKMFAFNHVSNHTLKTIAVVRAWTCFKPNITELDNFTDVSNPLVKMKSEQFSITPSLELDCMKTADTVDTWIVSLVSNVNAIIYTQPTGRAFNYTVRL
LQYGVYHLNYTVFMKDNPDTDRSESVYLKIIPTPLDVNIGGNNTQTLEFTKPFTLKALNFSYDPDVEPNDKSGLVFKWNCTQYNSTFQTDSVDIASLDNQMQEVFKHRQFCFGRQTADVDSTSGVFNPDTSLMLPRTW
YHIGLRMTKDTRVDTTNTWLYVKDLSAPVIQIGCIENCAAKKAFMKQIVLGTTCSANCQSRLGGLKFKWSLVANYENGTIETINNVIDWTEGGITGQNITFKKETLRAGTNHTATVVVFTPGSKQSQVSLKFLVNSPPY
QGNCSVSPFQGVTSVTNFVATCEGWKAFQPPGGVAQLRYDFRISYPGSVRTFPVGFGSSTKTEEVKFPVGAYDKSFQLNVVVRILDTTDDYFQFIFGVKVDPRPETISSVLEKVKLSAELELKNLSASGNYRKGAAMSFAL
ASELNVPDRYYVADPNSTVSFQQLKSAKSALRTSIITDIATYPINEPEELLQNTEVMRSAMTVRDELEADALRVGTSRLEGLTDSFRNFKSSGKVENVTQMSSDLVGLSVKCMGAGDFLMKQADTPKVPDNETLENSTAL
NSTMLEEEKRILKEKVKQVMRSVVNVVHNISGSLAEDMKPDDADKILAFDELDLLVSMATVDKLENGTLKETKQGSLKFGDPKLLFGSAVPRCLESQMFFTLINPFIYGNGSVNNVTSSVLDIQMRNCDNKLPSSSRRR
KRAATSNELLPKIPYDVSLVPTKLPDPIIVFNKSHYDKFIYHELNTSKPLKPIQIYIKHNLSFPLELFVGINDKPTDQISFSNFTVNETETTFFLSVAALSTFINDSMNGTVYVAVRSTACTSAPYCNGDEGTLTYSLYLYEIECQ
SWDETKEEWNSMSCKVEKETTPEKIVFKSNLFGSLGTGLLQAPNKIDFGTVFNNLGAKIVENMPVVATVLGVIVLYVVGLVVCRKFDKMDAMKWQCLPLIGSVVGGEYQYEVHVFTGFKTKSETTSKVYIEILGAHGT
TGTRELTDIENRKNFRRGECNAFKMSTTNDLGKLYAINVWHDGSGGDWYLSKVIVVDCHRNEWYQFGCNKWLSEERGDKMIKRTLDVSVLESHVNTGVLLRNNMSRKFFDDHLWLSVAYRKSPSRFTRAQRLSCCM
ALLYLTMIGNAMWYKAPDEEQGTVQALHLGPITISVSMLVASVYSSLVVVPPIMLITFLFTKANDKPKDNKVENDTEDFNELYANYYYGKKEVKKQKSEEKEVKKAFTLPYWCIFFGWLLVFLSVAASAFFTILYSFQW
GGVKSTGWLVAFLLSFFESVIIHQPMKVMLIAMIISFMFRRDVDVNEYEENYGEMQNWYDIYNDKEISLLRRKLPKDNKDENVLKLRRLDNKVGIMIKEIISNFLVLFVMMCVAYVSYDDNMYRQNANLKSALVENVDP
EVEHFLNVWPWVEKVLIPSLYYVEHYNGGKVSEYDKQFISNGMAIRFGPPRLRQLRVIPENCRIRPNIQTLSKRCIPRVTKTHGDEKNYLTNWTIPVNGSQDDFSIPKNVSETFRYQSASSLKILSYWATLEVLPGGGYVA
NLGNNKQDAEKLVEQLKAMQWIDHRTRALFAEFNIYNANTELFTQVILAFEFPSNGGVHPFHSFSTIQLYRYTGPMGLFSLIMEIISLFVFVGLLISSIIKMATKRAEYLKNLWNWVHMAALFLYAIAVVVYAVKMKWTI
QVIEYMMNNKEETVSFQHVTLGDQIFVCLFGLVIFLLTLAFLNILRFNWTIAVLAQVLVSSAKDLTLMALNILIMVMAFASMGNVLFGPYMSSYKDLPDTMTTLSTNFLGAFDYQEISESTGNAGRVFLIVYLVTMMFIVV
NLFITLLCTFVDAVRSDTSIIPKDHEVVDHFLDMVQGFVLPEKDEENEQDGDKKK

>PKD12681
MWQLFLLACIKYGLATVPDLLIRMEDDQAKSVEGPSTTVISGPISASDVVDGMFGKAIKFNGGPHLEVDVSAVSCFKTITTCSNGISGSMWIQPQTKTYQTSEYYYPDEEYMNNGGEQDLTPGFYFRYDGVLGKYEVGI
RRGTQLFKVSFKLQEMYWHEVFWAHTSGYLRVYVNGKRIQEITGASAAAVSNAANAATGKIIIGKGSKAILDEVAFWQTSVTYPTDNVGCLPGFTHSNFDCFLILRANQSQSWNASQALCEKYGGHLAIIKTETHKTDV
AALITGISATSSYVGLNDWDLDGEFIWYDGLFGPTNLQPEAPRVGNCTVFTTGKKFESLKCDTNHSAVCSASKSSTVLKDATDVVTTVQAPTNCTISFESSYASNRIKAESPFKFKASSSIGPVESMIVNMGDGTTVEAMG
DSFTYIYSTPGDRNISYLVANSGGSYSCSQNVTVDDVSSDFEIFDVHVTDAKPNENVTVAWSNRVNPYANGTCLLMELQYGNGAVDIMNIAIGQSVTGSMEYAYTTAGEYNLSATIDYGSGYVQSTRTVNIASSLVVTNA
PTYVNARYGYDFDIAYTANGTGISASLYKGTTETFGCVSTFNSSTNSGKLTVPFSQFNCAGSESLKLVVSDGIHHETKFVTVNLETPISNLFLTSEQQNYVFNDTSNITLTYGISAGSNVEFYIVFDDTNEVFISPCGDVSGK
NYTIQHTFMTAGVYTVFVLASNSFYEVTASRIIRLEYTMSPINLTWLQNDTDIATYGIPPYHNATLIVTPENPTFLPQGFRMMMTYGSPSKVKDEPIFNSSFTPTNSFNLMFATPGDRPVHVRLYSLSHSNATATVKVRLYDA
MTFIPSKFIAAVNEEITFNVSNITGVQVTYSFEAEGQSPQEENRTVLLNPTYAHTFKTSFNKTGIMTVTAKVSSRSYNLTKSVDILIEPHIMSLDMEVKNKSIPLPGATVNMQFKQNSSAVFSALANCTITYGDNRNRTETK
MTLNVSDWHPFRHTYATTGTFTIFVECSNNVSSVNKTTVTVNPLNLTVNPRSIIYNNTGAGLMNFQIIDNMITVYITPGLSAVFMYGSPSKVKDEPIFNSSFTPTNSFNLMFATPGDRPVHVRLYSLSHSNATATVKVRLYDA
LHFNVSDHIVNTATETVSVMVWGWVAPYKYTVDWADGSPQSSCSSSSGSCPNPATFTRTYTTPGVYYIQVFGSDNTVGTEKLIHNRSITVQHPVGPNVLQLALVKNKLILPNETAGFTVSVTATNTPTNATCTMNFGDT
GPTDSSSQTFEAGRTKSYARFYKNHLNLNGNHKNVTANVSCVNLVSSWSDTVPLQLFEAITVFKVTPEMKIAEIGYPVQFVLSVATGSDLQFNISFGDGNYQELTGSGYSVNHRHPYMMVGFDFMVNARVWNPMQTV
QLPEFKMIVTQTMFVPIENYTITAPEINYIDYITGKAIVPIKVVAGEALVAPDNVSIVLDYGDLNKTIGPLEMMLTYPYAYYLPGKYTVKLTVYLNNTEETRTVETTLNVTYACLSPVAAFPHIVRNASSPYNLFLGLPVVIS
ASLTVYRCRYGSHAHIWKFYGITNQSYRNTFTTTGTVLQVVKETLPPGEFRVELIISVDGKSSNASMYLRLNRSPLEAKIKGGSYLEVNKNQQVTIDATDSYDPDKADPKNIAHLAFAWACKQLTAAEENSDPAGITMDP
TCITPFSATNKNTLGFNSNLLTANKQYLFRLTMTSSDSRQSTVNPLNLTVNPRSIIYNNTGAGLMNFQIIDNMITVYITPGLSAVFMYGSKDNGFQRVKTPLKLEGICKSCTASDPVPNYVWSVMEDLPTPVVVPNQAGQTLVIERNKFAPGKKYIFKLEAHWSTNTA
AKGMATLNSVGNHLPRRGNCTVNRPKGSIFETFTFDCHGWLDEENAQNPGSASDKNHSLTYQISVDGKKIAFANNALQLPLKDYLLPSTSEETSTKKVVATVCDIYGGCSTSETSIEISNPVADIVNKIAASSPTDSKNTL
TATEEQSIIGLIENYVGRAEEDDAFPLKTLMFQQGLTAFASSFSKPNVEAEPTETTATSPEEEAANIAKTTNSLTGEMAAVAGSSGSTGASTGGSTNTGSSASGVSEAEIGSYGQVLLNIGRKDSATETQVDKTVKAAIQAS
LSRVDTLTQAVFANNGSRQDESITKQAIIMASSILESSAANPRSMSDESMGLGTSVMSNSAKALFDLPSQSAESMNETAKAVLNGLGSLSNSVSNNFDAGSGEGPEDNAVGEVDFAGVGMTEESIAALEQQPEASEPRVSPQ
EVEYFAMIRNRRNAVKSDDAEKRLQEHAAKILEAAATVTKAVSKSAGIGEGPWTYGNKEMNVSAVKINQADSSSTVASVGNQNFSFPRGFGSNSTTEAATISAVAFKRNIYGFLNGSADINGGMTHLAVTERGAQAPV
PAVLFSLQQPSTARPKQIAPAESKGDPSKMSYLRIYVPNPQNTLSVTLRQDDSLSASVSKDCDIYLQVNTFPQLTKFAAKKTVSMVKGVNTIVNFSPSERPTGVRYYLGLSCKVKFNTARRKRSVNGTASAVNNTLVKNALV
GLGIDLTNISASNPDLNQTAVVDGTTYTVGQLITLYNAANGNFAVQSAATGCRTNDDNSWNSNSCKVTATSTPDNIQCHCEKPEKFNFAASFVVAPNMIDFNMVFSNWDPSNVGVLLACLVVVLVYIPLAVCCRRLDK
RDIVEWGVTKLLDNKQGHSHYYLMSIYTGMRGGSATRSKIGFSINGEKGETGFRLLYDGLRRELKRGSVKRFLLAVPKDLGKLNYVHIWHDNSGGGSFASWFLDRIVVQDLQDDTGFVFVCDEWLAVDKSDGEIERIL
PAAGKDNLMSPKALFMDNVKLTATDDHLWLSLLTRPSPSNFTRLQRLSCVMTLLFMAMIANAMFYGTDGSEERPVGINVGPFKFTLGQIFTSLMSILIITPVNLLVINLFKRAKPRSKTFHKNLQDSLKQKNQNKHQELE
AIEEADVPPKRRFLLCCGKSKEKETFVEATDSEVLDQMLVTYINENGLSFPWWTTIIGWVLVVLIVLASSFFCILYSQQWGKEKANEWIVAFFISAFQDMFVLDPVKVCMIAGFLALVFRKPALQMESVDQTQLKADEEY
VSQPSDGENAADVKNKYIGTAGLTVLPPLAQDTITRLRRKKLKERRLYELIQEVIVYVLFLLVIWTISYSNRDPEAFKIYDIVAKNIIDNDGGFRQVRTPSDYLQWIKDQLIPWFAYDNYDGNRFHWRQRRFVQDLDNY
RLGPARLRQLRMPHVPAVDFYGTPYNTYGAYSPLKADSADHCIGWSSTPCNPDEKTYNFSSEAWKYTKATEVSGVPMWGQFNIYSGGGYVANLDVNRDVSKLILKELWDNLWLDWKTRMVFLEFTLYNANANMFCS
LTLMSEFPEFGGITQTVRIYTFRVYVHQGNTGMFVIACEILFICYLIYLIVRTVKRMYHKGCKSFFGNGWHVIDFLIIGSFAAIAMYFIRAALTDKILRNFRVDATEYVNFQIVAMWDEYFNIVLALLVFVYTMKLIGLLNV
NKHMGSLLSVLRVGKEELISYTVMFIVVFLAFSQLGWLLFGTDSKYYATMMRACEELLSTMLGKINLQTNKGIEFESTGAQTYYYMLYILVSTTILANTLICIIVDAINSVNAEGFDISKHLQLIDFIIYRVKEMFGFDAHFDF
DNFDLAEDEDCPEKGPFDGPRPVADEQAELWYRTLSLQINEIMSHLWRTVQNIEGQDYEKQDYKRLLISSEYSTMKSPEVPPT

>PKD9951-9950_Artificial_fusion
MPLIVNLIGGHAKTVGYGHSLIMNASKSKDPDMPHLGTTGINFSWTCTIVNNPQQPCPYQLSQLPNSDTLQLPTTGLVLHNSYSFTLNMSKDTRHRVFTQVVSVAEATKNDLLISPILNSEGILNPDLPFKIKAECSTCSPG
VLPNYDWEVWEEDGNNNLTKLPNIPTTTGLNGEYLAISGNALKSGKKYRIRLKSNGSYVEFPVQVNARPYGGNCSVNTKLGIAYETNFIISCAGWLDEGYREDTNPLVDNQTLVVLEYEIFYELVGDKLGIFLFVENG
AGSDFQSKPHQFRQGYFSNDYKVNVIVTVKDVRSTRTRYVIPVQVVPTTKATTDIISSSLFQDQRDRGDLRSAIRTGLEFLTMMSDKVTNVISGYGEGNSVEEGSCTAGVKFTPGSNETRIMQNVLDDVIEFLGNYSSQIE
SPSDIVAAAQLLLAFVPNMDIVTFETKYKTMDAVLALANRAVLGINSSVTDRERMTVAKGLLTGAYAFLESVGDTHKPVDPDLGFLEEDIAPWVKENRTIEARAARRDHERKRHLAISIHKLDDVRPRVTRTMTSGLVP
GERPTIVTTEFTTYIVERVKPRGLSKTNYSALLTSYVLGDMSDQDGVNQAETVDLEILAYQHQLLKFANSSGKSKMQQSSLTMYDSDGALKPISGLKNPIKISTKTEAPPANPTAMTSSMMVDGLHFRRVDLMEDVSNLK
VLVTINQRTNMTYYVKGHGFPNESHYDVTGVLNLDDAQTNDTIWIDMQPFSPNENSSTVLYFAFDVNDTKPGLSFDFQALALDCIYWNTRNLEWENDGCKVSPATSTTEAQCLCSHATTFSKNWFVPPNKIDFKTVFA

KFKNIGNNLAVLLTVIILFVLYVILSIIARQIDKDDILRWAALPLTDNTLHDSYHYEVTIYTGLRKGAGTRSKVFFVVTGEHNQTEVRALEDGKRKCIERGSILRYILSVPGPLGPLEYVTVWHDNSGEGGDASWYLDSAVI
KDLQTDEKSVFLCGRWLAVDEDDGMIERTIPAATEEDLRAFGHLIFISTTKHDMFDAHIWLSIVSRPTRSHFTRVQRLACGASLLYLTMIANAMFYRADQAMAMEGREKTDLFYIAPSELCSEPSTSRDDMPILGSPKHTAK
ALAPAYVIDGIPYYPTMPLEDETPSTEERSDRKKPFMLPHWVIYIAWTLSLLSIFGSAFFVVLYSLQWGKDKSEEWMTTFFLSFFQSVMVVQPLKVLFIAFLITLIFRRPRDDEDANYETRPKRDEELLPNRPNITHYKDLWT
YAAHRKPYQVDPGHAEMLDRMRARRATQLRAYAITKEIIVYCIYLLLISTMAYDNQDVNANRAYEHFNDMFVYNSMDRVHSMAEFWNWTETIFIPRMHADYWYNGQPVHWRDTYLTYDYNSFRVGAPRFRQIRTASS
AGCKVHPSLKSFIDKCYDSSSSSNDKHFAPGWKTLNTSVNRPTPWTHQSAFSLMTLPYVGRVGSYSGGGGYVAQLGRWPQQSKNFIADLKKHSWTDRLTQAVFVEFVLYNPNTNLFTAARLAFETPGSSGVITQHLQTRVL
RLYPYVGAKLLFRLLLDTAVLIFTIVYMVAQIKLMKKMKCQYLKTFWNIWECINISAFIVTVGMYALHFISVKLTVNEVHENPEQYVSFDAVALWGDLQNAFQSIMLFTAMIKLLRLIRFSQLVHQLGALMKMVSAPLTSF
FMVFFLLMFMAFTHTAYLTFGKQIRNYSTFISSFESQIEMLLMKFDFAGMRSAHGIVGPMFFILFVFCVVILLLNMMLSIINDTFLASSTSIPENKDREAAIFLVNQIKEVLMEFPTLTATQKGSKRWKKKNVEISTQTLATSET
EELINKTSPVDDFPRLTPKIPLKAKPKEAFIVPGLKTA

>PKD688
MKKQCLPRKCLPVGWVISVLAILTSGFFVILYSMDWGAEKANAWLVSYTTSFFQSVLIVDPIKVIAIAAFTALLVWRAKLSEAQHSKDGLDEEEDDNDKDESELTPEEIEKQRKRAVLTQAAKDMVKSLVAHALLLWVLV
SLSIMCSSSSGFQINRNIHNVFVNSFKSMATSKGGFFDWLHNEFIPNYFVEKDYAGNDLRWEDKLYTTDFQSHRVGPARLRQVTIKPRERCQQGGLSWKNDRSCYGTYSLADESEDTFGIALENITNKNALLTSASWSYNT
SLELDSLPIWGHLSVYGGGGFAASLDVNRDISIDILNELNEYRWLDRQTRAVFLEFSLFNANVGMLTYITLLAEMPNSGAAITFTKIQPFKVYSAGLSSIHLVICAAIFLIYVIYYIVSFFIGICKHKRKHFTKVMTWFNLVFIIT
AMGMIVFFALRTIALNITMDKLKEDKTKFVSFAYVAALDASFRYCVAFTLSLAIIKLIFLLKLYRRISQLMLTLRKSIKPLASVLVALSIMVLSYAFYAYLYLSPNLAKYHTIISSIESMVGMALGSFDVMEMFQVNPYFTWFFFT
TYIILMNFIMLSILITVIMDSHYEFKDHQPECHEHEVIDYFVKSVFRGLQKTAVTDASL

PKD phylogenies: alignment and trees

PKD2/PKD1 phylogeny

Sequence alignment (FASTA format)

>270002831
AMAEDF--
TRHFREAVLYIIFVIAVTICTVGFWHYTQSLMLENFYGVPRIRQVKVTESIPFSGYGGGGFYLDLLKENLWTTRGTRAIFIDFSIYNANLNLFLIFEFPPTGGIIPVTFVFECTVYAIAGFFLAEEYFLQFWTYGNLEYLAEMRI
FYNNFAASLLFFSYIKLFKYLNFNKTMGQLNNTLKRCAFDILGFSIMFFIIFFAFALLGYLLFGFGVAMFTLLRTILGLAPIYFLAYIFFVFFLLNMFLAII
>1253289448
S--
IKKSDDKKLKDLISHFIFLVIVCIIAFSVWEYMNNQLMAGLYGKPRIRMLKVLENSDLWGYGGGGYVQELLKANRWIDRGTRLIVIDFAVYNGNLNLFLMFELPATGGVIPFIGGCEAAFCVLVIILVIKEYLNKFWNSSP
MDDVVNAENMYNNIAALLLFFAWIKVFKYISFNKTMNQLSATLSRSAKDIGGFAVMFFVFFFAYAQFGYLVFGFYDAVFALFRTILGLGPIFFITYVFFVFFLLNMFLAII
>1174757226
ETEDTV-
TNKELRELFVYILFLSTTLVMTFGLFKYLKGTMLDALFGLPRIRQLRVTGMASYWGYSGGGYILDLLEARRWLDQATRAVIVDLTVYNANLNYFLCVEMPPTGVLVPAVLAFEVIFCLFVAYYIGEEYFRSFWNYVDFESL
SFYQTILNNALAITVFCAWIKVFKYLSFNKTMMQLSSTLGRAAKDVASFAIMFFIVFFAFAQLGFFLFGVGTASFTLMRIILGFGPLFFFMYIFFVFFLLNMFLAII
>847028821
HTENCD-
EDKNVRELVVYMLFLVLLFIITFSFWKFLMGPLLNGLYGVPRIRQIRVLQSGSFQGYGGGGFYVDLLFNNRWLDRNTRAVLVDFTIYNANINLFLVGEFPPVGGVLPFVLACEILFVLYLLYYIIEEYFCSLWNHPNFEYLS
FWQLQYDNLSAFLLFLAITKIFKYISFNKMMTQLISTLNYCLKDLVGFSLMFFIVFLAFAQAGYLIFGFYICIFTLFRIILNLGPLYFILYVFFVFFLFNMFLAIL
>675439102
KQKIVK-
GNRDLRELILYIIFLAILMIIAFGFWNFATGALLNGLYGLPRMRQVRMLGSYPFQGYSGGGGFYIDLLYTNLWIDRGTRAVIIDFTVYNANINLFLVAEFPAVGGVIPFILGCELAYLIFLIYYTIEEYFHSVWNHPNFDNLSYN
QTTFNYLLAVTVFLAWVKIFKYISFNKTMTQLSTTLGRCSKDLAGFAVMFFIIFFAFAQLGYLLFGLHYSIFTLYRIILGLGPLYFILYVFFVFFLFNMFLAII
>847028815
HTEDTK-
VDRQLRELVIYIFFLAVLVIISFGFWKFLTGPLLNGLYGLPRLRQVRVLDTTEYQGYGGGGFYVNLLFNNLWIDRGTRAVIIDFTVYNANINLFLVGEFPPTGGLLVFVLACELLFVVFIVYYIVEEYFFSIWNHPNFESLTF
WQLQFDNLIAITLFLAWVKVFKYISFNKTMTQLSATLSRCSKDLAGFAIMFFIIFIAFAQLGYLIFGLQYCIYTLFRIILGFGPLYFVFYVFFVFFLFNMFLAII
>Ptor32544
HTEDTI-
HDREIRELLVYIIFLIVLMIISFGYWKFLTGPVLSGLYGLPRLRQVRVLKGRSASGYGGGGYYIDLLFNNLWIDRGTRAVILDFTVYNANINLFLVAEFPPTGGMIPCIMACEMVFIMFILYYSIEEYFTSVWNYPNFDFLSYW
QLQFNYATAFTLFIAWIKIFKYISFNKTMTQLSSTLSACSKDLAGFAVMFFIVFLAFAQLGYLIFGFEISIYTLFRIILGLGPLYFIVYVFFVFFLFNMFIAII
>1160116546
AAHAKFFHRRNLREFLLYLVFLLDLSILTFGFWEYVEGPLLDSLYGHPQIRQLRVEFELWHWGYGNGGYIFELLKSKLWLDSGSRAVFIDFTVYNANVNLFLLVEFPATGGAIPFLAACEIIFCFFICNYLIQEYFKNFWNY
QNLYFLAYWQNRYNVIIAFTVLIAWIKIFKYINFNMTMTQLSTTLSRCAKGIVGFAIMFFIVFFGYAQLGYLIFGFGNCIFTQFRIILGLGPVYFITFVFCLFFLLNVFLAII
>1067079595
QMT--
KSDDRELRELVVYLLFLVILCIMTFGVWDFYENVMLDGLYGLPRVRQVRVLNGFAHWGYSGGGSYKNLLRSNLWIERGTRAVFLDFTVYNANINLFLVFELPATGGVIPFIMACEFIFCFFILYYLVEEYFKTVWNYANF
EVLGYAAVQYNSVVAVCVFFAWILKLFKYISFNKTMTQLSSTLSRCAKDLAGFAVMFFIVFFAYAQLGYLLFGFISSTFTLLRTILGLGPIFFLTYVFFVFFLLNMFLAII
>926626589
EMA--
TTEDKELRELVIYIFFVVVLCILTFGFWRFVDNVMLDSLYGSPRLRQLRVLNGSSHWGYSGAGSYIDLLRQNLWISRATRAVFLDFTVYNANINLFIVFEFPATGGMIPFIMGCEILFSVFIMYYIIEEYFKSIWNFADFEFLGF
WQTQFNNAVALAVFFAWIKIFKYISFNKTMTQLSSTLSRCSKDITGFAVMFFIVFVAYSQLGYLLFGFLKAIFTLLRVILGLGPLYFLSYVFFVFFLMNMFLAII
>675370064
QMKG-
KEDDKELRELFTYIVFLVLLSIMTFGFWKFAKDVLVENLYGSPRIRQLRVMNGSNHWAYSGAGSYRDLLRQNLWISRATRAVFVDFTVYNANLNLFLVFEFPATGGMIPFIMACEVCIVLFILYYIVEEYFYDVWNFPDFE
FLGFWQMQFNNAVAVDIFIAWIKVFKYISFNKTMTQLSSTLSRCSKDIGGFAVMFFIVFFAFAQLGYLLFGFGTSVFTLLRLILGLGPIYFLSYIFFVFFLMNMFLAII
>241171514
QLGGGAKADRELRELLIYIVFLVILCVLTFGFWSFVEDVMINSLYGSPRIRQLRVLDGADHWGYSGAGYYADLLKDNLWIGRATRAAFLDFTVYNANVNLFLVFEFPATGGMIPFVLACEVVFVIFILYYIVEEYFKSIWNF
PDFGKLGFYQTQFNNAVALAVFFAWIKVFKYISFNKTMTQLSSTLSRCSKDVAGFSIMFFIVFFAFAQLGYLLFGFVNAVFTLLRLILGLGPIYFLSYVFFVFFLLNMFLAII
>556956196
T-----SSSKELQELLIYLIFLIDLCILAFGFWKFAEGPLLEGLYGVPRIRQLKVY-
TAPHWGYSSGGYYLDLLKSNHWFDHRTRAVFIDFTTYNANVNLFLTVEFPPTGGAVPFVASCEIIFCLFTFAFLGQEYFKNTWNYADFYFLAYWQTQYNNMIAVIVFIAWIKIFKFINFNKTMTQLASTLSRCAKDIIGFAI
MFFIIFFSYAQLGYLVFGFENCIFTQFRIILGLGPIYFTIFVFFVFFLLNMFLAII
>5669804
L-----HYRKELQELLLYFIFLINLCILTFGFWKFMEGPLLEGLYGVPRVRQLKVN-
SPWHWGYRNGGYIFTLLRLNSWITRGTRVIFIDFSLYNANVNLFLVAEFPATGGILTFIASCEITFCIFLFVFTTQEYFKSIWNYSDFYFLACWHIYYNNIIAITIFFAWIKIFKFISFNKTMSQLSSTLSRCVKDIVGFAIMFFIIFF
AYAQLGFLVFGFQNSIFAQFRIVLGLGPIYFITFIFFVFFLLNMFLAII
>1033394660
L-----THSRELQELLIYIIFLTDLCILTFGYWEFLEGPLLDGLFGVPQIRQLKIL-
TPWHWGYGSGGFMFTLLKENSWITRGTRAVFIDFSMYNANVNLFXXVEFPATGGAIPFLASCEVIFCLFIVAFIIQEYFRDAWNYPDFYFLGFWQTRYNNMISVNVFFAWIKIFKFISFNKTMTQLSSTLSRCAKDIIGFAIM
FFIIFFAYAQLGYLVFGFQNCIFTQFRIVLGLGPIYFTIFVFFVFFLLNMFLAII
>685230872

HTLKDEQSDRELRELITYIVFVIIICIVTYGFWKYAEGPLLDGLYGRPRLRLLRVLKSRDYWGYDGGGFTHLLLKDNMWLDRASRVVFLDFTVYNANINLFLIVEFPATGGAISFVGFCEVVFCLFIIYYIVEEYFKNVWNY
ASFENLGFWQVQFNHMSAVLIFFAWFKLFKYISFNKTMNQLSSTLSKCAKDVLGFCVMFFIVFFAFAQLGYLVFGFGDSVFTLFRIILGLGPIFFVTYVFFVFFLLNMFLAII
>675875990
DTLNEN-
EDSNLRELVVYILFLIVLCVMTFGLWTFMQNNLLNSLYGVPRIRQVKVMKSSSYSGYGGGGFAQDLLFDNLWIDRGTRAVIIDFTLYNGNINLFLVAELPATGGLFPFVLACEIFFILFLIYYTVEEYFKSFWNYANFVFLSY
CQDMANSALAIMVFFSWIKIFKYISFNKTMTQLSSTLSKCSKDLGGFAVMFFIFFLAFAQLGYLLFGLGTCIFTLFRFILGLGPIYFVCFVFFMFFLLNMFLAII
>828192580
YTDDIK--
DRELRELAIYLIFIIILSIISFGFFNYCDKVLIEGLYGLPRLRQVRVLEGRVFSGYGGGGGYVSLLLKHNLWLDRGTRAVFIDFTVYNANINLFLVFEFPPTGGCVTFVMGCEVIFVIFILYYLIEEYFFDVWSYANFDFLSDWQC
RFNDFVAVGVFIAWIKIFKYIKFNKTMSQLQSTLARCAKDIAVFAVMFMIVFLAYAQLGNLLFGHIDCIYTLFRIILGLGPIYFVTFVFFVFFLFNMFLAII
>50539686
LLGESN-
SNRELREMITYILFLLTLCIITYGFWKFTEGPFLNGMYGVPRLRQLRVLGESSYWGYGGGGYYQDLLKNNLWLDRGTRAVFLDFSIYNGNVNLFLLVEFPATGGAVPFVGMCEVSFCLFVLYYLVEEYFKSLWNYPNFEP
LARLQVHFNNLAAIIVFLSWVKLFKFINFNKTMNQLSTTMSRCAKDLMGFAIMFFIVFLAYAQLAYLVFGFQACIFTQFRIILGLGPIYFTTFVFFFIFMLLNMFLAII
>1160100767
LTVDHA-
VSRELREMVTYSIFLVILCVLTYGFWKFAEGPLLDGLYGVPRIRQLKVLKETSYWGYSGAGYYLDLLKENLWLDRGTRVVFITFSVYNANINLFLVVEFPATGGALSFLAACELAFCIFVFYYIIEEYFKSVWNYPNLESLAY
WQIQFNNMAAVVVFFVWIKIFKFISINKTMTQLSSTMSRCAKDIIGFAIMFFIVFLAYAQLAYLVFGFHTCIFTQFRIILGLGPVYFVTFVFFIFFLLNMFLAII
>4505835
LMEESS-
TNRELRELVTYLLFLIVLCILTYGFWKFTEGSLLDGLYGVPRIRQLRVLNGSSHWGYSGAGYYLDLLKKNVWLDRGTRATFIDFSVYNANINLFLLVEFPATGGVIPFLAACEIIFCFFIFYYVVEEYFRSFWNFPNFEHLAY
WQIQFNNIAAVTVFFVWIKLFKFINFNRTMSQLSTTMSRCAKDLFGFAIMFFIIFLAYAQLAYLVFGFQECIFTQFRIILGLGPIYFTTFVFFMFFLLNMFLAII
>557013709
LMEGGT-
TSRELREMITYIIFLVILCLLTYGFWKFTEGPYLNGLYGVPRIRQVKVLNGSTHWGYSGGGYYLDLLKDRLWLDRGTRAVFFDFSVYNANINLFLLVEFPATGGMVTFLAACEVAFCFFVLYYIVEEYFRSFWNFPNFEPL
AYWQIQFNNMAAVIVFFAWVKLFKFVNFNKTMSQLSSTMSRCAKDVIGFAIMFFIIFLAYAQLAYLVFGFQDCIFTQFRIILGLGPIYFTTFVFFIFFLLNMFLAII
>1242857867
HTLESD--
DRKLKELVVYFLFLVLLSVVAFGTWKVLKGSLPSALHGLPRIRQVKVLDGSSSEGYSGGGFVQNLLFTESWLTRGTRAVFVDFTVYNANVNLFLTVEFPPTGGAVSFIMACEFFFALFIVYYVIEEYFKSVWNFQDFDFISL
WQIRFDNAIAITVFLAWVKLFKYVSFNKTMSQLSSTLARCAKDLAGFAVMYFIVFLAFTQLGHLLFGFTSSFFALFNIILGLGPTFFMLFIFFIFFLVNMFLAII
>Chem014425
LTESNKEDDRELRELIIYVAFIVILCILTFGFWTFADDGLMNGLYGRPRLRQVRVLNGRGFRGYGGGGGFTKLLLKANLWLDRGTRAVFIDFTVYNANINLFLLLEYPATGGCIPFVFACECIFVLFLLYYAIEEYFKDFWN
YANFDTLGYWQETFNDFIAVAVFVAWIKVFKYISFNKTMTQLQSTLARCAKDIAGFAVMFFIIFLAYAQWGYLIFGYLNSVFTLFRIILGLGPLFFISYVFFVFFLINMFLAII
>156401013
QTEDT-
DSDRELRELIIYLVFITILCILTFGFWMYTTETLSDGLYGRPRFRQLRVLQGRSYWGYSGGGFTQLLLKQNLWLDRGTRAVFIDFTVYNANINLFLVFEYPPTGGCIPFVMACEGIFILFIIYYTIEEYFKSFWNYANFDSLGF
WQTQFNNMVAVAVFFAWIKVFKYISFNKTMTQLSSTLSKCAKDVAGFAIMFFIIFFAYAQLGYLIFGFEDSIFTLFRIILGLGPLFFMTYVFFVFFLLNMFLAII
>Pmar01872
LTEDIT--
DRDLRELFVYIAFLINLIILTFGFWKVVQGPVLNGLYGVPRIRQLKVMNGSSHWGYSGGGYYLDLLFDKQWLDRGTRAILIDFSSYNGNINLFLLVEFPPTGGAVPFILACEGIFCCFILYYIVEEYFKSVWNYGNFQFLAF
WQTQYNNMVAVNVFFCWIKIFKYVSFNKTMTQLSSTLARCAKDVLGFAVMFFIIFFAYAQLGYLLFGFSNCIFTQFRIILGLGPLYFTTYVFVVFFLLNMFLAII
>Pmar7053
LTEVNS--
TRELRELIIYVVFLIDLCLLTFGFWKVVQGPVLDGLYGVPRVRQLKVLSGSSHWGYSGAGYYRDLLKDYLWLDHGTRAVFIDFTTYNGNINLFLVVEFPLTGGALPFIMACEVIFCCFILYYVVEEYFKSVWNYADFEYLA
FWQTQFNNMLAINVFFAWIKIFKYVSFNKTLTQLSSTLARCAKDVLGFAVMFFIIFFAYAQTGYLVFGFVNCIFTQFRIILGLGPIYFTTYIFFVFFLLNMFLAII
>675861632
QTEETA-HNQDLRELVIYCVFLAILCIITFGVCIYAEKVLLPGLYGVPRLRQVRV-
LGSNVRAYGGGGYYADLLKTNLWVDRGTRAVFIDFTIYNANINLFLIAEFPPTGGHIPFVFACEIIFCIFILYYIGEEYLKEFWNYIDFDYISYWQSVFNSALAIMVFFAWIKIFKYISFNKTMTQLSSTLGACAKDLGGFSIMF
FIVFLAFAQLGYLLFGYETAIFTLFRIILGLGPTYFILYVFFVFFLLNMFLAII
>Bner206993
SLMDSKTADSELRELIIYGIFLTILLVLTFGFWRFSEGPLINGLYGVPRLRQIKVLESGNVKGYSGAGYYQNLLKQNLWIDRGTRVVFFDFTVYNANINLFLLVEFPATGGVIPIVLGCEGIFVLFLLYYIVEEYLYDIWNFG
DFEFLSYWSVQFNYIIAGAIFLAWIKVFKYVSFNKTMSQLTTTLGKAKKDIVGFLVMFMIIFFAFAQLGYLLFGLINAMLTLFRIILGLGPIYFIFFVFFVFFLLNMFLAII
>528498254
LRKDTK-
ENPELKELLIYIVFLVDLCLLTYGVWTYMEGPLLDNLYGVPRVRQLKVIKGSSHWGYSGAGFYQDLLRENLWLDRGTRVLFIDFTAYNANINLFLVVEFPATGGAIPFIIGCEIVFCVFILYYIVEEHFTSIWNYADFEFLAF
WQTQYNNMNAVNLFFAWIKIFKYISFNKTMTQLTSTLARCALDIFGFAIMFFIVFFAYAQLGYLLFGFNKCIFTQFRIILGLGPIYFFSYVFFVFFLLNMFLAII
>359465612
LTENTA-
ENRELRELLVYIVFLVDICLLTYGFWDFAQGPLLDSLYGVPRLRQLKVLGGFSHWGYSGGGYYLDLLQEGLWLDRGTRVVFIDFSVYNANINLFLVVEFPATGGAIPFIVGCEVIFCVFIFYYVVEEYLSSIWNYADFEFLA
FWQTQYNNMNAVNLFFAWIKIFKYISFNKTMTQLSSTLARCAKDILGFAVMFFIVFFAYAQLGYLLFGFIKCIFTQFRIILGLGPAYFVTYVFFVFFLLNMFLAII
>908475875
YTSDVK-
EDQEFRELVIYVLFLIFVCIIAFGFWKFAKGPLLNSLYGVPRLRQLRVLKGFMYWGYSGGGYIKDLLIDKLWIQRGARAVFISFTVYNANINLFLLVEFPATGGAIPFILGCEVVFVIFVLYYIIEEYFLDFWNYAPFENIGFW
EVRFNNAVAITVFIAWVKVFKYISFNKTMTQLSTTLSRCAKDLLGFFIMFFIVFLAFTQLGFLLFGFMNSFFTLFRIILGLGPIFFMLFVFFVFFLINMFLAII
>260802732
QTEEFH-
GDKELRELFTYIIFIVILCILTFGWWKFTRGPLLDGLYGVPRIRQIKVIDGSGHRGYAGNGYVQDLLFNNLWTDRGTRVVFIDFTVYNANINLFLTAEFPPTGGLIPFIMACEGLYVLFIIYYIIEEYFKSFSNYADFEFLAWA
QGIYNDMVAITVFFAWIKVFKYISFNKTMTQLSSTLARCAKDIAGFLVMFFIIFFAYAQLGYLVFGFPDAIFTLFRIILGLGPMFFITYILLVFFLLNMFLAII
>918287730
DTEETK-
GDRELRELIIYIIFLAVLCVITFGVWKFAKGPLMDSLYGLPRMRQLRVLDGTSHWGYGTGGFVQDLLFDNLWIDRGTRVLFIDFTVYNANINLFLVVEFPATGGALSFVMACEAIFVLFILYYIIEEYFKCFWNYSDFQFLA
FCQTHFNNAIAVTVFMAWIKIFKYISFNKTMTQLSSTLGKCAKDLAGFAVMFFIIFLAFTQLGYLIFGFESAFFTLFRIILGLGPIFFILYVFFVFFLLNMFLAII
>405953847
QTEETK-
GNRELRELIIYLVFLIILCVVTFGFWKFARGPLIKGLHGLPRLRQLRVLDGSSHEAYSGAGYVQNLLFNKLWITRGTRVLFVDFTVYNANINLFLIVEFPATGGAIPFIMACELIFALFIAYYIVEEYFKSLWNYADFEFLSYW
QTRFDNAIAIAVFFAWVKIFKYISFNKTMTQLSSTLGRCAKDLLGFAVMFFIIFLAFTQLGYLLFGFGNSFFTLFRIILGLGPIFFMLFVFFVFFLINMFLAII
>871233656
QTEETK-
TNRELRELIIYIVFLVVLCVITFGFWRYAKGPLVNGLYGVPRLRQLKVLHGSSHWGYSGAGYYKDLLFDELWIRRGTRAVFVDFTVYNANINLFLLVEFPATGGAIPFILACEIIFCLFILYYMVEEYFKSFWNYADFERLSY
WETRFSNAIAIAVFMAWIKIFKYISFNKTMTQLSSTLGRCAKDLAGFFVMFLIIFLAFTQLGYLLFGFENSFFTLFRIILGLGPIFFMLFVFFVFFLINMFLAII
>1236532622
DTEDVS-
QNAELRELVMYCLFLVVICILTFGWWGFAQGPLISGLYGVPRMRQVRVLDGRGYSGYNGGGYYIDLLFHGLWIDRATRAVFIDFTVYNANINLFLIAEFPATGGGFTFVLACECIFMLFICYYIVEEYFSSFWNYPDFDFLS
FWQEQFNQAIAITVFFAWVKIFKYISFNKTMTQLSSTLGACAKDLLGFAVMFFIVFFAFAQLGYLIFGFQQAIFTLFRIILGLGPIYFILYVFFVFFLINMFLAII

>47551289
QTEDTS-
LDKELRELAIYLVFLVVICIVTFGFWRFAETPLMDGLYGVPRIRQLKVLDGSGHTGYGGGGGFYADLLKENLWLDRGTRALFVDFTVYNANINYFLTVEYPPTGGAIPFILACEGIFCLYIMYYMGEEYFKSSWNYPDFEY
MGYWQDLFNNIIAVNVFLAWIKTFKYISFNKTMTQLSSTLSRCAKDVGGFTVMFAIIFMAYAQLGYLVFGIGSCIYTLFRIVLGMGPIFFLTYVFVVFFLLNMFLAII
>291234667
QTEDTK-
TDRELRELIIYILFIIVLCIVTFGWWRFAEGPLVNGLYGVPRIRELKVLDGSSHWGYGGGGYYLDLLKEYLWLDRGTRVVFIDFTVYNANINLFLVTEFPATGGAIPFIMACEGIFCLFIAYYIVEEYFKSVWNYADFEYLSY
WQSMYNNMLAVLVFFAWIKVDNLFYAK----------F-------IFTGFTLFRIILGMGPIFFCTYVFFVFFLLNMFLAII
>198420844
QTEDTQ-
TDREIRELVIYVFFIATLCILTFGFWRFSEGPLMDGLYGVPRIRQLKVLDGSSHWGYSGGGYYQDLLKENLWLDRGTRVVLVDFTVYNANINLFLVVEFPATGGAVPFLMGCEVLYMLFILYYIVEEYFKSVWNYPDFEF
LGFWQMQYNNMVAIVVFFAWIKIFKYISFNRTMTQLQSTLSRCSKDIAGFAVMFFIIFLAYAQLGYLVFGFQESIFTQFRIILGLGPIFFLTYVFFVFFLLNMFLAII
>919076678
ETVDAS-
GNKDLKELIIYLLFLVVLCVTTFGYWQFARGPLMDGLYGVPRIRQLKVLDGSGHYGYEGGGYVQKLLKQNLWLTRGTRAVFTDFTVYNANINLFLLVEYPATGGALSFVMACECIFILFIIYYIIEEYFKHFWNFPDFEFL
SFWQMQFNSAVAVAVFFAWVKVFKYISFNKTMTQLSSTLSRCAGDLAGFLVMFCIMFFAFAQLGVLVFGFWNAVFTLFRIILGLGPLYFFCYVFFVFLLLNMFLAII
>PdumPKD22
QTENTD-
KDRELRELVIYIVFLIVLCIITFGFWRFSESSLVNGLYGVPRIRQLKVLGGSSQRGYGGGGYIQDLLRENLWLDRGTRAVFVDFTVYNANINLFLMVEFPATGGAIPFIMACECIFCLFILYYTIEEYLKSFWNFADFEFLSFW
QVQFNSALAVTVFFAWVKVFKYISFNKTMTQLSRTLGSCAKDLAGFAVMFFIIFLAFAQLGYLIFGFGNAVFTLLRIILGLGPVYFILYVFFVFFLLNMFLAII
>PdumPKD21
QTEDTG-
EDRELRELVVYLVFLVILCVVTFGFWTFSEHALCDGLYGVPRVRQLKVLDGSGHWGYGGGGFYEDLLKQNLWLDRGTRAVFIDFTVYNANINLFLVVEFPATGGAIPFVMACEGIFVLFILYYMVEEYFKSFWNFPDFE
FLSYWQVQFNSAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLIFGFDDAIFTLFRIILGLGPIYFLLYVFFVFFLLNMFLAII
>919076662
HTEDTK-
TDRELRELVIYLLFLTDLTILTFGFWRFAEGPLMDGLYGVPRLRQLKVLDGWGHEGYGGGGFVQNLLKQNLWLDRGTRVVFVDFTVYNANINLFLVVEFPATGGAIPFVLACEGIFLLFIIYYIIEEYFKSVWNFADFEF
LSWAQVLFNQGIAFAVFFAWVKIFKYISFNKTMTQLSSTIANCAKDLAGFGVMFFIILFAFAQLGYLIFGFPDTILTLFRIILGLGPLYFLLYVFFVFFLLNMFLAII
>Pdum9951
LDRMRARRATQTKEIIVYCIYLLLISTMAYDFWNWTETIFIPRMHGAPRFRQIRTLMTLPYVGYSGGGYVAQLLKKHSWTDRLTQAVFVEFVLYNPNTNLFLAFETPGSSGVITFRLLLDTAVLIFTIVYMVAQYLKTFWNY
VSFDAVALWGDLQNAFQSIMLFTAMIKLLRLIRFSQLVHQLGALMKMVSAPLTSFFMVFFLMFMAFTHTAYLTFGFISSFESQIEMLLMVGPMFFILFVFCVVILLNMMLSII
>676479791
LEKVREYRKKQFQEIIVYSFYVITIAILCYEFWEWCDERLLPNLYGLPQLRQARVVNGFPIAGYGGGGGYTAVLLKSQRWLDRQTRAVFVEVNLYNPSVNMFMMLEFPPTGSVSTFHLLDLVFVAFTLFYLIQQYFQEFWN
FVNFQRIVLWDEIFCFTQAFLLFFSMLKLLRLVRFNSKVTQLAGTLRRSVWPVFNFCIVFTIIFVAFAISGYIVFGPLTTMQTLFSMLLMIGPLYVFIYILMVVFLINMFMSIL
>1087817335
LREARLARE KHIHEVIVYLLYLILLLMAANTFWSWCNNSLIPGLHGPARIRQLRVLNGHPYTGYDGGGYSVTLLKDHGWVDRPTSAIFVEFTVYNANVNLFLLLEGSANGAFFPLLVITYIIFVLVLLYKIFNIYFHHFWNF
VNFQYLILWNDIFRFLMSFILFFATIKFTRLLRFNRRIRLLSMTLSYAAKPVVLFSVMFSIYFLGFLFLANLWFEIPLAAGEMFSMLLNLGPLFFSAFTGVCTILLNMFIVIL
>148539556
LQEARKERLRDIRDAFLYALFLMVVLYLAYIFWNYIEQTLIPGLYGPARLRQLRVLYGVPYQSYGGGGFVAEFLRDHGWYDKYTRAIFIEFTVYNANVNLFLLVETPTFGGGVPVAVAFQGIFAVYLLYFLGHFELKNFW
MYVNFNHIAKWDQIFGYLLATLTFLSILRFMRALRFSARVTHVAFALQRASKDMLYFLFMWCIIVIAFALWGYAVFSFQAVLESLLGMMIGIGPIFFTLYIFAAYFLANMLITII
>PdumPKD11
RQ----
LRMKELKECIVYFFFILVVFFLSYQVWTYLRNTVLPNMYGVARLRLQRILKSAPYAGYKGGGGYTFTFMEQYEWVNPNARAMFIEFTLYNGNVNLFLLVEWIAGGGATVIVILFEVLFLAFTVYFLVKEYFKEFWNYVNFV
TIGVWDEMFTFLLAFVVYLTTVKFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFCMFAYVIFGFIKTMESLFSFALGIGVLFFILYVFVVFILMGMFLTII
>871242653
RDEASLSRSHGRSGLGL-AFSVLCAKFLVRNFWVWLEGSFLPSVYGVVRLRQMRI----------
ARHQDWLDLNTKAIFVEFTIYNPNANLFALVEFLTTGSTTQIVLIMEVFFGLTTLYFFYQCYFKDFWSFVNFNSIALYDEMYGYVMAIVVFLATIQFLKLLQFNKKMGMLGSTVKLASKDLKVFSITFFLYFFAFTATAFLL
FGVITAAESMFAFALGWGPIFFFSYIGVVYILMSIFLTII
>919018042
RE----
QRMKEIKEILIYFFFLLILFFLSYQYWAWLNSVLLPGLYGVGRLRQMRVLKNAPYMGYKGGGYVQDLLKDDDWFDQYTRAIFVEFTLFNPNVNLFILKEFLAAGGSTTVVIIFELMFGAFTIYFLVREYFREFWNFVNFQS
LALWDETFGYMVGVVVFLATIKSLKLLRFNRRMGMLGDTIRIATQDLKVFSITFFLYFFAFCQAAYLLFGFIGSVEALFAFSLGLGPLFFFLFVAVVYILMGMFLTII
>1205898338
RQ----
LRMMEIREIGIYALFVLILFFLSYQYWNWLENKLIPNLYGLPRLRQLRVLKSAPYMGYKGGGYVALFLKKDVWLDVYTRGVFLEFAVYNPNINLFLLTEFVSSGGAVALVVIFQVIYACFSLYFFIRCYFGRFWNYVNFQSM
AMYDELFGWMMACVVFLATVQFLKLLQFNKKMNMLGDTIKVATKDLKVFSIAFALYFFAFTACGHLLFGLLGSAESMFAFTLGLGPLFFFTFVGVVYVLMSMFFTII
>762107655
RQ----
QRLIELKELSSYIFFVIILFFLSYQYWVWLENKLIPSLYGIPRLRQLRVLANAPYMAYKGGGYVYMFLKNQIWMDAHTRAIFLEFTVYNPNINLFMVIEFLHSGGAIPMVIIAEVIYTIFSLYFFIRCYFKGFWNYVNFQSLALY
DELYGFMCGFVVFLATVQFLKLLQFNKKMGMLGETVQYATKDLKVFSIAFLLYFWTFTATGFLLFSVVGAAESMFAFTLGLGPIFFFAFILVVYVLMSIFLTII
>71999568
LASGHEKSDGKFMEVGGYAVFLIVLVYVAFAIWDWLSQVLIPGIYGEPRIRMLKVLENLKTVGYGGGGGFVQRLLKANRWIDRGSRAIIVDFALYNANINLFLLFELPASGGVITRMMIFEGIFCGFILYFIFEEYLTQFWN
NAPFDDVTSSENSYLNIKACVVFVAWVKVFKFISVNKTMSQLSSTLTRSAKDIGGFAVMFAVFFFAFAQFGYLCFGLYNSAFALLRLILGFGPAFFIAYVFFVSFLLNMFLAII
>585644120
LMKAKELKKREFKEIFVYCIFVITVLLIAYGFYIWVDEVLLPGLYGPPRFRQLRVLEGVPLYGYSGGGYVAVLLFDNRWLDERTRGLFVEFTLYNPHVNLFLLLEYPTTGGALIVVAACEVLFIIMLIYYMVKAYFVKFWN
FINLQQMAAYDEAYGYVIASIVFFGTLQFINLMRFNARMALLSTTLKGAVKVLANYAIIFYIVFFAYGQLLFLTVGFFTTIKTLYLMWLGLAPLVFFTLMLFVYALTTLFVTVI
>585688456
MEKARRKRFKELKEMFVYGTYIIILLLIAYGYLEYIRGAVIPSLYGPARLRQLRVLDGVSLYGYSGGGGYTAELLESTHWMDIRTRAVILELTVYNAQTDLFLISEFPSVGGGVAFLMACDVIFICYIIFFIVEEYFKDYWNFLN
LQSTAIYEEIYNLVLASLVFCSILKFMKLMRLNHHIALMIGTLGAARRDLANFCIVFGVTFIAFGQSGFLLFSFISTIETLLSTLLGLGPIFFFTFIMMSMMLLNLFVTAV
>1126215987
LREIKETRRKEIKEFLAYTFFLCVLLLLANEFWKWAQGSLIPALYGPARFRQLRVTKGVPIIAYGAGGYVVELLFKETWIDHNTRAVVTEFSVYYNANINLFLTAEFPPTGAASTIIIAAQVLWVIFIVYFTIREYFKSFWNFVS
FQRLATMDEMYVYVMSTLTYVGVFKFMRLMRFNKRVCILLAIIKNSIPALGLFTIMFTICMTAFALSGLVIFGLLSTYEGLLAMVLGLGPVYFCVFVNLMIMVMNMFVSIL
>1087817330
LRKARKLREKEIWEVIAYFFYIMILLFLCHGFYRFVNQTLLPNLYGPARLRQIRVMRGVPFVGYPGGGGYSAELLYKDAWVDEQTRAVFCEFTVYNPNVNLFMVVEFPPVGSAVTMVLAQVIFFL-
HLLTTTINFLFNDPWNFVNFQQVSTIDMMVATTLALSAFFGIIRFLKLLRFNRKIGMLTSVLKNSARQWPGFFVMALVFFIAFIHGYLVFCVIQSVETMFGTMMGFGPIFYFIFCFTSAFILNILIAII
>1139813591
LAEARERKRMKEVREIAFYFVFVALIVLVASHFWKWLLEILLPRLYGPVRIRQQRVDAGLPFSGYTGGGYVADLLMLNNWINQYTRAIFIEFTVYNPYVNMFIAFEMPTMGKMLIFVVLCELATLAIVIYFIVREYFKCFWNF
YNFRKVAMWDEVFGYFMAFTIFVSIIKLLHLFRFNKRMSMIAGTLKYSTKELSAFSVVITILIVAFASWGYIMFGMFQSFETLLAFSLGFGPIYFFTFILCIMILLNMFVTIL
>585651646
LEALRKERLNELREISIYAVFVWY----------
GPARIRQLRVLNGLPVFGYGGGGGYLAELLRNNSWYDIYTRAIIIEFTVYNANVNLFFLIELPTHGGGLLFVVGAQVLFVGFVAYFVIHEYFKSFWNFVNFQHLAYWDTIFSMIIALELFVAMLKFLKLLRFNKRMSMLGMT
LKYCFKDLLYFSVIWIIVFLAYCQVFYLVFGMSITMASQISMMLGFSKVFFFFFILCLYWLMNMFLTIL
>1139824617

LKTIREERQKELKEIAICMVFLYIVLTIAYAFWNWSENALLQGLYGGARLRQLRILSGYPYQGYSGGGYVAELLRDNSWVDKRTRAVFIEANIFNPNMNLWYLCEFLQTGAIEPFVMAIEISALAFIIFYTVKEYFKAFWNF
VNMHFTVMLDDLSSWFIAFLVFFLTVRFLRLLRFNKKISLLSSTLRHSAGMLISFSFVFNIAFLSFACTAYLFFMYIVTLETLLSMMLGTGPLLFGFFSICFSFLINFFLTIL
>919082289
VKAARDLRIKEIREIVFYFVFLLLLMMCAYGFFDWAERTLIPALYGVARLRQLRVLKGFPYWGYGGGGYAFELLRSSRWLDIYTRVVFVEFVLYNANLNLFTTVEFPASGGVFLVIMGAEILTAAFFAYFLYREYFKDTWN
FHNLQYSILWDEIYGYMMGFIIFVANVKFLKILRFNKTMNLLGRVMKGAAKPLGYFFISAGIVFFAFVQLAYIMFHIITTVETLLVMLMNFGPIFFMIYTVTVSFLISMMLSII
>PdumPKD916
LQEARQMRFKEIWEIIFYFLFLWLLMMVAYGFWDYVSQVLLPGLYGQPRLRQVRTLSGYPYWGYSGGGYTADLLRSNRWVDSNTRAIFIEFLVYNPSVNLFVAMEAQESGAMYLVVMAGEILTLVFIIYFLVKEYFKDSW
NFHNLQYAALWDESFTYLMAFMCSLSTLKFLKLLRFNRRMAMLSSTLRYAAGPHQFLIMFLILFLAFVLASYIVFGFARSMATLFSAMLNFGPTILFLFTLMITFLINVFITLI
>919098890
LKSAREQRFKEIREILVYLFFLYMLFTVSYGVWKYINQTVIPGIFGAPRLRLVRTTDGYPVFGYGGGGGYVAELLYDNRWVDRYARALFVEFVVWNPNVNLFIAIEFPESGGAWPVVMACELITIIFIIYFLVREYFKKFWNF
VNFQYVGYWDEMYGYFLAILVFLAVLKSMRLLRFNKRMSIIGVTLKEAGKPMITFGIVFFVLFFAYAQLMYNIFGFITCCETMFSMMLNFARIVFVSFMVIVSFMVNLFLTLI
>918996724
VEHMRGIRMKELVELALYVTFTLALVFVTYGMWEYLEETFVPSLFGAARLRQLRVLRSLPFAGYNGGGYVKLLLQEFKWIDQRTRAVFLEFTLYNPQVNLFLLFEFHPTGAIFPFVAVCEVMFILFILTSTYREYFSDPWSF
VNFFPAAFWDYILGYTLAAMLVLTVLKFFKLLRFNVRMNMLSATLKKSWKPTLHFYLVLAFLMMAYAFFGQCAFGLIDTLQTLFTLIMGIGIVFLITFAMIIQVLINMFIAII
>1067095036
TKNLRVKLTKEVRGVLVYCIFLAVLLIIVTDFWLWAKLAVMNITRGYPILRQIRNLKSFTSVGYGAGGYVIRLLQKVHWIDKRSRAVIFEFSVYNANVNLFIYFEINEGGGIVTFLTLCELTFVVSTVFFTIRQYFLMYWNYL
RMDYPVLIDQYYTYILGLIMFLSILKLIKLLQFNKRMDVLALTIALCWDELSYFFIAFTVIFFAFCCLFYLIFYVISAVQTSFSMMLGLSPLLFFIFSVLNSMLVNIMLAII
>196000368
VNAARSLRLKQFREILGYVVFLTVTILICYGFWRWTNQVFLPGLYGVARVRMLRMTGSYPYWGYYGGGYVVELLQQQGWIDHQTRAILIEFTIFNAQVNLFLLAEFPATGGIMPLVIVCEIIFVSFIAYYIYREYFSGFWNF
INFHYVATWDNLYIHLFAFLIFFGTIKLIRLLRFNRTIAILSITLRSAAKEILMFLVVFTVIFLTFAQFAYLIYGFIKTLESQFNMMIGVSPLYFFCYTFLVSWLINMFLTLI
>1005480473
LIAMRTARMKEIREVITYYIFLTLLLTIAYTFWNWTHSELIPGLFGGARMRLIRILHGYPVWAYSGGGGYVIVLLEKNMWIERHTRAIVTEFTIYNAQTNLFLVAEFPATGSVIPFVFTIDILYVLFILFFTYREYIKEFWNFKSF
QFAAYWDDTYTLCLAVIVMLANIKLNKLLRFNKRFSLISSTLKYAYYPLLMFSIVWGIVFIAFTFVAALVFGIMASLTSLLSLMLGLGPIFFFCFMFMMSFLINMFVSIV
>957833931
LNAARETRFKELRDLGIYIFFLWIVWVVSYSVWNWTQNVLPGALVGPAWMRQVRVMKSLPYWGYPGGGYYVFFLERNRWIDKYTRAVFIEFSTYNAYVNLYILFEFPSSGGIWKLKIICQIIYAIFVLIFIIKEYFRDIWNYI
TLQYINYVDEILGYMLGFTCFLATLKMIRLLRFNRRMSLLGQTLASCGKDLFAFTVMFGFVFVAFTMMFQLLLGLLRSAESSFSMILNVSPVFFFVFVVLCTFLMNMFLSII
>1009565777
LDAARKLREKELREMGAYILYLWILMILSYGFWNWTRHSLIPELFGYPVLRQVRALNGFPFWGYGGGGYIVPLLEKTGWIDKDTRAVFVEFGTYNPQVNLFIVAEFLPGGGVVPLILLCQIGFILFTIYYTICAYFKEYWN
YIKLQEALLLDEAFCYLIGFLMSLATLKFLKLLRFNKRIGGMISTVRLCVQELKGYSVCLIITFIAFVCLFWMTLGFLASFESSISMLLKLAPIAFFTFALTASVLVNILLTII
>1078802969
LDHVREERMKELREIFSYLLFLWLLTVLSYGFYNYLRETVVPQLKGYGVLRQIRALHGLPFWGYGGGGYVMRLLEVSDWIDGGTRAVFVEFSVYNAQVNLFIIAEYHPGGGVVPFVLLCEISFIGFIVYFTVREYMRSYW
NYVKLQRVAAVDEVFSYLMSFLIFISILKFTKLLRFNKRIGVLYSTLSQCSKDLSSFFVVFWVIFLAFTQVFYVILGFVTSAETCFDMARGIGPMVFFVFAFITSVLLNLFVTLI
>675368552
MERAKLERLKELKEIISYAFFIWILTILSYGFYKWTREVLIPELIGYGMLRQVRILNGLPFWAYGGGGYVFPLLEKSDWIDKRTRAVFAEFSVYNAQVNLFIVAEFLPGGGVVPFVLVCELTYVIYTLYFLYREYFKSYWNYI
KMQFVASVDEVFGFLMAFLMFIAILKFIKLLRFNKRMGILYSTLAMCAKDLKSFFVVFVVIYLAFVQFFYLVFGFVTAAETTFAIMNGLGPLSFFVFCLVTSILLNIFVTLI
>1238854278
TQRAKEERLTEIREILSYLLFLWILIVLSYGFWRWVREAVVPELLGYGVLRQVRTLDGLPFWGYSGGGYVFPLLETSDWVDSRTRAVFVEFSVYNAQVNLFIVAEFFPGGGIVPFVLLCEVFYVVFTLFFTVREYFQSYWSY
VKLQFVAAVDEIFGYISSFLIFIAILKFTRLLRFNRRMGFLYSTLSQCSKDLKSFCVVFLITYLAFVQVFYLIFGFVNSAETSFNMMRGVGPLIFFFFAFITSILLNIFITLI
>684373764
LERARLYRLKQIREIFFYVIFLILLLVITMSFYAWARTTLIPRLRGYGLMRQLRVLDGVPYMGYSGGGYSHMLLESVHWIDFCTRAVIVQLTVYNPNINLFVLIEKPGVGTLIPLEIGFQVAYILALLCYLVKEYFREMWNFM
NFQFLAYWNEALTYLLAVICFFATLKLVRLFRFNQRVGLLGAVLRYAAHDLKYFMIVFMVIFMAFVLVFYLLYTFLSSVETSLQIILGLGPLMFAVFTLCIIFVISMFLAIL
>Ptor35851
IEKAREFRLKEIREITFYLVFFMTLLLVTYSFWNWALNVLVPALRGYAVMRQVRVLDGIPYMGYNGGGYVYQFLQNNRWLDFQTRAILIQFTTYNPNINLFFILELPGSGGMWPFEIFCQALYGIFILVMIVKEYFKQFWN
YMNFQFVGFWNEYMGYLTSFIVFAATLKFIRLFRFNRRIGMLGAVLSFCSKDLKQFFIIFMIMFLAFVMVFYCLFMFLYSMESSMQIALGFGPIVFACYVIFIIFMLNMFIAII
>847028803
IEKAREFLRLKEIREVGFYLAFFCTLLFISYSFWNWILRVFVPRIRGYAIIRQVRSLDGLPYIGYSGGGYVYKLLRDNKWIDFQTRAIFIEFTTFNPNINLFYVAEMPGNGGLWTFEIFCQVLYGIFILVIFVKEYFSQFWNYLNF
AQVAFWNEYFSYLVSLIIFCATIKFIRVLRFNKRIGMLGAVLSYTATDMKQFFIIFFIMFLAFVMMFYSLFMFLFAIESSMNIPLGFGPIVFAGFVLFIVFLLNMFVAIL
>780039266
LERAREQRVKEFREIFFYVVYLWVVLVISYGYWAYLHSTLVPAIYGYPVLRQARVIESYPVLGYAGGGYLAKLLFEEAWFDRYTRGVFLEIAFFNAQVNLFLVAEFLPTGGVETVRLACEILFLGFIVFFIVREYFKDFWNRI
NLQYVAYWNELLTYLLGVVVFIGTLKFIKLLRFNKRIGILSTTIGSCTKDLFHFGIMFGIFFLAYALAFYKMYAFIFTLETLSAAMLGLGPIFFFFFMITTTFLVNMFLTIV
>1229191925
LERIREQRMKELREIAFYIVFLWTLMVFSYSFWRWVDTGLVPNLYGKVVMRQLRILEGYPFLGYGGGGYVVVVLYNETWIDRYTRAVFIEFSCYNPHVNLFLTAEFLQTGAGQPVRIICEVMFFCFILFFIVKEYLRNFWN
YIKLQYIAYWNELLGYMLGWIVFLASLKFLKLLRFNRRIGVLSSTIHNCSSDLMYFGIMFGIVFFAFATAFYLIFGFIYTCESLIAAILGLGPFFFFSFMMSMSYLINMFLTII
>291224264
LEAARENRIKEIREIIFYVLFLWILLVISYGFWSYAHDVLVPGLFGYPVIRQLRILDGYPYLGYWGGGYIARLLEAENWIDKYTRAVFIEFSTYNAQANLFLLMEILPTGGGYPFQVICEGTFFAFVLYFLFKEYFKGFWNYI
KLQYVAYWNELLGYMLGVVVFLANLKFLKLLRFNKRMSLLSSTLRAASKDIISYFFMFAIVFLAFAEAFYLIFTFVYTLETLFGTMLGLGPVLFFFFVLTITFLINMFLTIL
>918302224
LGKMREQRLKEIKEIIFYAFFLWILMVISYRFWKWAKNDLVDGLRGYAVMRQLRALDGYPFYGYGGGGYVARLLQNTSWIDKLTRAVFVEFTIYNPQVNLFLLAEFTRIGGVVTFQTLCDVIFVIFLIYFLIKEYLKEFWN
YMKFQYVAFWNEMFCIIVGWLVFLATLKFMKLLRFNKRMSMLSSTLTNSAPSLLNFSLVFWIFFLSFVQVFYLLYFFITAAETGFLMMMGLCPLIFAMFVVTITFLLNMFLSIL
>1207917521
LEKARAVRMKEIREIIFYSFFVWILMVISYQFWTWAKTGLATGIRGYATMRQLRSLNGYPYWGYAGGGYLVRLLEREKWIDKYTRAIFVEFTVYNPGVNLFLLAEFRPSNGIFPFQLACEVIYLLFTLYFVIHEYFKQFWN
YMKFQYVGYWSEIFSYIIGWLVFFGTLKFLKLLRFNKKMSLLASTLKHSCKDLMHFSIIFNIVFLAFIQLFYLVYFITALESGVVMMMGLTQVSLFLYVVAITFVVNMFLSIL
>762156788
LEKAREVRLKEVREILFYSFFLWILMVISYRFWVWAKSGMVDGIRGYATMRQLRVLDGYPYMGYSGGGYTVRLLEKEGWVDRYTRAIFVEFTVYNPGINLFLLLEVSPSQGIFPFQIACEIVYFAFTVFFIFQLYFKQFWNY
MKFQYVGAWSELFSYIIGWLVFFASLKFLKLLRFNKRMSLLASTLKNSCRDLMHFSIIFNIVFLAFIQLFYLVYAFITACESGLVMMMGMTQIFIFAYVVTITFVVNMLLSIL
>871215788
LERVRSKRIKEVREILFYCIFLWVLLVVSYRFWEWAKTGLINGLLGHAVMRQHRILNGYPYWGYSGGGYVAHLLEREKWVDRYTRTVFVEFTVYNPQVNLFFLAEFHSTNGVVPFQIVCEIAFVVFIVAFSFREYFTKFW
NYIKFQYVGYWNEIFAYLIGIVVCFATIKFLRLLRFNRRVSMLSSTLRHSARNLFHFSIIFWIVFLAFCQLFFLTYMFIESLVASVLMMMGLTQIFVLLYVVSVTFIVNMFISIL
>676474191
LEKARIQRLNEIREIVFYSFFLWILMVISYRYWKWSQSGLLNGIRGYATMRQLRVLNGYPYWGYSGGGYVVQLLEDDDWIDQYTRAVFVEFTIYNAQINLFIVAEFDSAGGVTTFQIACEVIYLIFTAWFIIRAYFTQFWNY
MKFQYVGYWNEVFSYMIGWLVFFATLKFLKLLRFNKRMALVAATLKNGSKSLMHFGIIFLIVYMAFMQLFYLTFMFINSMESGVLMMMGLTQIFVFFFVVTITFLVNMFLSIL
>871254812
LEKHRKRRMNEIREVIIYSFFLWILMVISYRFWIWAKSGLVNGLRGYATMRQLRILNGYPYWGYSGGGYVVNLLEREKWIDRYTRAVFVEFTVYNAQVNLFILAEFHPSGGLVQFQLVCEVVYILFTLFFIIREYFKSFWNY
VKFQYVGYWNEMFSYMIGFLCFFATIKFLKLLRFNKKISMLSATLKHSAKGLMHFGVIFVIVFMAFAQLFYLTYMFISSVVASILMMMGLTQIFVLLYVITVTFIVNMFISIL
>908400046
IEKLRAQRMKEIREVIMYSFFLWILMLISYRFWTWANSGLLSGLRGYATMRQLRVLNGYPYWGYSGGGYVVPLLQSEQWIDRYTRAVFIEFTTYNPQVNLFILAEFDQTGGLVPFQLVCEAVYMLFTLFFIVREYFLSFWN
YVKFQYVGYWNEMFSYMIGFLIFLATLKFLKLLRFNKKISMLSATLENSARSLLHFGLIFTIVFMAFSQLFYVTYMFLSTLISGVLMMMGLTQIIVLAYVITVSFIINMFVSIL
>919074653
LEEAREKRIKELREIIFYCLFLWLLMVISYGFWHWYKEGLVPGLRGYATMRQVRVLDGYPHYGYGGGGYVAELLEREKWIDKYTRAVLVEFTCYNPGINLFMVAEFPETGGVFTFQIFCEVSYFVFIFFFMYKEYFSEFW
NYIKFQYVGYWNEMYSYMLGFICCLGTLKFLKLLRFNKRMSFLAATLRHSTKHLFAFAVMFFIVFFAFVQFFTLDMGVLYASESCMQMMLGLAPYFFVLFVVLVSWTINMFLSIM
>443713927
LEEARLERLKELRELLVYSFFLWILLVISYGYFEWCHNTMIPGLRGYATLRQLRTMDSYPFWAYSGGGYVVELLIKNQWIDRHTRAIFLEFTTYNPMVNLFFVAEFIVTGGIDTFQLVCQIIYLVFIIVFIVKEYVKEAWNY
MNFRYVGYWNELLLYMVGWLVFLATVKFIRLLRFNKRMGVLASTLRYAAKSLFMFGIMFMIVFMAYAMFFYLIYFIIFTMETLLQMLVGLGPLFFFCYVVTVYYLINMFLTIL
>PdumPKD12

LEKARQQRMNEMREIIFYFIFLWMLLVISYSFFDWSENVLAPELRGFATMRQLRVLDSFPFWAYGGGGYVRELLHDGGWFDHYTRAVFIEFTVYNAQVNIFLVAEFLPTGSLFTFEIICQIIYMLYILFFIISEYFTEWWNYM
KFQYVGYWNELLLYMIGWLVFLATIKFLRLLRFNRRMSMLASTLRNCAKSLANFALMFLVVFFAYVQFFYLLFFFLVASETCLQMILGLGPLFFLAYVISVALLINMFVTIL
>260785879
LAAARELRSKQMVEVIFYLLFVWVVVSVANGFWTWLEEGLIPGLTGVTRLRQLRV--------
LNGGGYVAELLKNNMWIDDNTRTVLVEFTIYNAYVNLFLALELPVSGGVFPYIMACEILYVIMLLFYLYREYFGEFWNFVNYSEAAFWDAVYGYVIAIIVTLVIAKFVKLLRFNRRMSLIGDTIKHAAPKVLGFIVIYIVVI
MAFAQALFFVLGFLACLKTLWNIQLGLGPAIWFTYVVTSSCLVFTFVAIL
>999987444
LQQARALRLKEIHEIVLHFLVVILVLFVAFAFWQWAESVLLPNMYGVGRLRQLRNLKGFPYFGYRGGGYVVDLLWKDNWIDPYTRGIFTELNVYNVDINLLLLYEFLPSTGGVYAVMGFEIIFLFLTFYWIYRDYFKSVW
NFVNFQYVVTLNEVLNALTAFIVLLCTLKFLRLLRFNRKVSMFSDTLRHSANTLFTFVVMIIMIVLXFCXLSNLXFGIIRSAVSLFSMMLGIGPLLFFTFMVSVQFLVNLFIGIL
>1005458745
LETARKLRLKQIKEIITHFVVVCIVLTVAYGFLGWARNSLTPALFGVGRIRQLRIYDSFLYMGYRGGGYIIELVKDHGWVDRYTRALFTELSLYNVNVNLLLLFEDLPTGNGKLALLACEVILAFMLVRWIVRDFWRNLW
NFVSFQYVVLLNEGVNISLALAVMLLNLKFLRLLRFNRKISVLSSSIKASARKLVSFMVMFLVIFLAYCFLVLLVFGIVRCMVSMMAMVLGIGPAVFFSFMVIFQFIMNMFIGIL
>999981318
LEIARQLRMKEIRELVWYLIFLLVILTISYGYWNWMNETVIPSIYAVVRLRQLRILKTFPYWGYNAGGYAKELLKSKDWIDRYTRAIFTELAIYNANTNFFLLYEQLPTGALFNFVMACEITYXLFTLFFIFRESLKNIWSFVS
FNYTAFLDEWINALMGFVVFLSFIKIIRLLRFNRRMSLLAQTLRVCIPRIAAFLVIYGSSFFAFSLFGCLVFGILRSATTLMDTILGIGPFFFYFYTVXMVFLINMFLSII
>1005492071
LEAARLLKTKQIREVIQYLVFLSVLLVIAYGYWKWVQDTLIPTLMGTARLRQLRVLKGTPYWGYWGGGYIGDFLESQKWIDRYTRAIFAEFALYNAQTNFFILTEILTTGGYHHFVMACEITYMAFLLFFLYKQYFCDPW
NFVSFQYVAFVDEWVSVTTALAVFFSILKFLRLLKFNRRISLLIQTIKIAASPLMSFMLMFFILFLAFCQFAFLVFGFPSTLGSVMSLTLGLGPIFFFGYVVCVYFLMNVFIGIL
>1207182030
IEKMRNNMIKEIREILIYVGFLWMLLLVAYGIFIWANSSLLRNLYGNARIRQVRVLRANPVWGYKGGGFVVDLLFDNTWLDVYTRAVFVEFTVYNANVNLFLIFETTGVGTFQYFVMASEAIYYIFILYYMVVQYFKSKW
NFVSFNETATADAVLGYLIAFLVLLATIKLWHLLRLNPKLHMITSTLQRAWTDISGFIVVMTIMFLAYSIACNLMYGLLEAAQTIICLQLGLGAFLIGSCIIFMTFVLNLFISVI
>942189627
IERMKKNYIKEIREILAYLGFLWMLLLVAYGFFTWTNTTLISNLYGSARIRQVRVLRGHAIWGYRGGGYVADLLFNTIWLDVYSRAIFVEFTVYNANVNLFLMLETNALGAFFTFVVAAEVIYFLFIIYYMVLQYFHSKWN
FASFYETATADAVLGYLIAFLVLLATVKLWHLLRLNPKLNMITSTLRRAWGDISGFITVIVIMFLAYSVASNIMFGLFDAAETMVSLQLGLGSFLIGSCIIFMTFVLNLFISVI
>116006951
IEKMKTTHLKEIREILAYLGFLWMLLLVAYGFFKWANTTLVSNLYGSAQIRQVRVRQGYPIWGYRGGGYVVPLLFDNTWLDALTRAVFVESTVYNANVNLFLTLETSALGTFFTFVVAAELIYFLFLLYYMVVQYFCSKW
NGISFSETAAADAALGYIIAFLVLLSTVKLWHLLRLNPKMNMITAALRRAWGDISGFMIVILTMLLAYSIASNLIFGLFDAAETMVSLQLGLGSFLIGSCIVFMTFVLNLFISVI
>156406861
LLAARQKRLKELKEILAYFIFIAVLLLVAYGLWHWARETLIKNLYGISRLRLERVSFGFPYWGYSGGGYPADLLFHSHNWVNSNTRGVFIEFAIYNANTNLFLYVEFLPTGGAFPFVLAMEVIMLLFMFYFVYREYFKGFWN
FVSFQYALASDETLNYVIGILVFFVTIRFIKLLRFNRRMALLGNTIALIAKPIGQFLISFTLVFISFAFFGNMIFGFTTTMMTQLNICLGLGRIYFFSFMYFIVMLMNMFLAII
>828228329
IAQLRAYRLKEMRDFVIYVMFVIALCQVTYSVWPWLNSTVVNALFGGARLRQLRVLQTYPFIGYTGGGYVADLLHQNKWIDGRTRALFLEFSTYNPQVNLFLLLEFSPSSAVEFLTLVCQIFVVFFFLIAMYKEYFRGFWN
FVSFNRIVQWDQLFSGLTSVLVLLTCIKSIRILQYNKTISLFVLTLKKSASPLAAFFLIFAIFFTSFTAWAYLMFIYISASESVMSLLLGLGSLWFTLIMIFGVMIMNVFLTIV
>1005450057
LEEARKLRFKEIKEVALYVFFIACLCIVSYSFYTYAKTTLMNGLYGMARLRQVRVLDTLPFMGYGGGGFVADLLESNAWIDPQTRAVIIEIATYNPIVNLFLLVEFLPTNGVFLFLIVCELFIVLFFGIFMYQEYFREFWNFV
NFGQVASWNEVFMYVVAMVVFVSSIKGIKLLRFNRRITILAQTLRGSAGPLAAFSVVFVVFFLAYSLFAFAVFGFVSTCEAVMGILLGLGPLFFFSFMVFGTFLMNMFLTIV
>1191059136
LKKDREQRFREIKEVLLYVFFVTCLCIVSYSYYYEWAKSTLMDALFGGARIRQLRVLDTMPFIAYGGGGYAAELLENNTWIDAQTRAVFTELATYNAVSNLFLVVEFLPTNGVFLLVVVCEFFILIFFSVFIYQEYFKEFWNY
VSFVRVSSWSEAFMIMVALLVFITWLKGMKLLRFNARIQVLSQTLKGAAGPLATFSVVFIVFFLAYALFAFAVFGFVTTCESVMGLLLGIGPIFFFTFMVFGNMIMNNFLTII
>156395095
LEKARDRNDKELTDIGTYLTFLWVLLVIAHSIWLWTNHVLLPALYGFARWRQLRKLNGYPYFGYSGGGGYIVPFARNNSWIDRYTAGLFTEFTLYNPYSNLFVLVEVLPTGGFHVFTLVCDVVFLLFTSYFTFREYFRDYW
NYTNFQYISAWQMSFESILGILVFISTIKFIKLLRFNKKMFLLAYTLSRSAKDLAHYGIMFLIVFLSFAQFYYLLLGFVKSIEKLINVLLGMGAFGFFLYMTLTNFLLNMFIAII
>1005483664
LQLARFKAAKELLDFGTYILFTLIVLLIAYGMWDWTEQALLPSLYGVARLRQHRVLKGYAYKGYSGGGYVLELARSKLWMDRFTRSVFTEFNLYNPNINLFLLLERPPTGGLLPFILACEVLFLVFVLYFTYREYIMESSNF
ISFQYLSAWQLTFEGVVGLTVFVTCLKFIKLFRFNRRIFLLSCTLRQAGGELLPYFIVFMINFLAFAQVVYHILLALIGSMQKLLSALLGLTAVIFIVYMIITNFFLNVLIVII
>999981301
LEMARKRLQQELKDIISYLTFMLIMLLIAYGMWKWLENLVLPVLYGVARLRQNRVLTGFPFKGYSGGGYVLTMAKAELWVDVYSRAIFVEFNVYNPSSNLFVLMEVLPTGGFILFIVAMEVFFLMFITYFTYREYFKEMW
NFTNFHYISAWQVVFENLMALTVFAGXLKMVKLLSFNKRIYLFAHTLRHALRDIINFCFIFIIVFIAFSQLYFSMLGMVGSMAKLMSVMLGLGAISFIAXMFIMKFLLNILIVLI
>156394387
LQRARLQRFKEIQEILFYALFVWVVLVVAYGFWNWTEAVLIPGLYGTARMRQLRVLRSLPYLGYSGGGYAFDLLKRDKWVDRSTRAIFVEFTVYSPYTNLHLLLEVLPSGGTSPFVLLCEILFVLFTVYFTYKQHFHEYWS
FVNFHALAASDQVFGYVYAFVAFFVSLKFLRLFRFNKRMSLLATTLKCSAKELAHFSIVFGLVFMAFVHFCYLIFSVLSTVETLISVMLGLAPFMFFFYSIVVVFLINMFISII
>1005475249
LHAARVRLFKEIREIVLYLLFVWMVLLIAYGFWDWMIQTLVPGLYGIARIRQLRVLKTRPIMGYRGGGGYNTDLLKSHQWVDKYTRAVFVEFTVYNVHSNLYLLLEFTAAGGVLPFVLICEVTFVLFSLYFTYREYLAEFWS
FVDFHHLALVDMLFGYVYAFVVFLTSVKFLRLFRFNRRMSLLGSTISASARELFHFGIIFGLVFVGFSHLCYLVFSFLTTIETLISVMLGLGPIMFFFYSIGVVFLVNMFLSII
>1229159547
LAEARDIRVKEIREIVFYLIFLHLTMLICFGFWNYTENVFFPSLYGRVRFRQLRIVSSFPYYGYNGGGYIVEFLQQDMWVDQHTRAVFLEFVVYNANVNMYLLVEFLPQGGATPFVLAGQIVYTLFIIYFLVREYFDSVWNF
LNFHFVAYWDQLYGYLSGFVLFLVNVKFIRLLRFNRRILLFSMTLRQCGKDLAYYGVVFCLVFGAYALFAYAIFHFVGTLETLFSTILGLGPTFFFTYVVIIVFMINMFLSII
>148539638
LGEMRETRFKELKEISVYMFYLYIMFLISYGAWQYLEETLVPSLYGTARLRQVRVLDSYPAMAYYGGGFTIELLQDLKWIDQQTRAVFMEFTLYNPTTNIHCVIEFSPTGGALPVVMAGEVIFVILQVYYTIREYFKSPWNF
NNFQYLAMWDDVYTTMVGLLLFIATIKFIRLLRFNKNMLLLTDTLTVFGYELALFMIMFFVIYLGFSSFAFLVF-FLRTLESLFGTLLGLGPLFFFLYVVIVMWMINMLLAIV
>1229145338
LNEARETRFKELKEIAVYMFYLYILLLIGYGFWKHMREVFVPSLYGTARLRQVRVLDSYPAFAYYGGGFTVEFLQQNLWVDQQTRAIFLEYVVYNPVTNTHNVLEFPPTGGVFPFVLAAEVIFALLLLFYVVREYFKDLW
NFHNFQFLAYWDDIYVTMVTLCLFMAVIKFTRLLRFNIKILMLSETVNEFAYELSLFMVMFAVVYMGYASFCFLVF-FLRTLESLFGTLLGMGPIFFFSYVMVVMWLINMLLAII
>470247587
EKATLLAATNIVKDLFTHFVFFIIFLVIVVLFKAYLTEQFGPVVTSVVRLRQLRVSHGSLVFGYDTSGYYVDILYDNDFFSENTRAVIISFVTLNLNYEARLLAEWTAAGSVDPFRAFLEITFLLFLTYYLYFFFTSFWNLGQT
ALVFY------QLSAINILLLAFKTFRFLMLHRRLYALWTILAHAKVQLITFTTMFIIMMIGFLMSGWLTFGFFNSLGTSLQYIIGLGPIYNILFTIFMFFLINMFIAII
>66823157
EKATLLAAVETVKDLIIHLVFFVVFVLIVVLFITYLNSTFCESISNVIRIRTARVSHGSYVFGWDRSGYYYIDILIDNGFFDIQTRSVIISFSTLNLNFQSRMLTEWTASGSVNPFRGFLEVSFFIFILYYFYFYISNFKNLGQTALV
LY------QISAINILLMSFKTFRFLVVHRRLYILWIAISQSRLQLITFTIMFCIMMLGFLFSGWLTFGFVTSLGTLLQFIIGLGPIYYLLFTIFMFFLVNMFVAII
>551658247
ALIARK--
KSKVRELIVFVLWVTCFSATITLFYSFLQEVLVKILAGPLRIRQVRVLSEMAYFQYPGSGYVLDVIQDSGWIDLMTRAIFVDFSFFNPNLNLFLVIEFLPSGTVTMLVIAAFSAVVLITLYYFLADFVKSFWQLQATGYMLDQE
K---NLAGIILIFMWIKLYKYFSMSRKLSTLTRTIAKAASNVIFILIVLIIPTIGFAMGCTLIFGFKTSLYTLLRAAYGLGPILLLFWVFFSNFLLNIIIAIL
>673034662
AILKKVQEHKDIKDMIQYLMFLALFIVVTVDLHQYLTGPFYDVVYGPIRIGQVRV--------
YGSPTFGVMVLKEHKYFDKATRAVFVDLNLYTANIDHTLFAEMLPGGGVTTVQAGCEGLIYVVILAQLAREYFGEFSNFRTPTIYFSLA---
RSCQSFNCFLSWLKLFKFLSFIPMFGQLTKTVQRAAGKVLELVVIFALTLVGAALAFYLAFGFLSSFYTLLHIVTGLGPFFFISFVFLMMFILNIFIVIV
>116008215
AKTNRILYTTDLVEFSVFIIFLILTSLVVLSWWDYLKYNFLVTLHGPPRLRQIRVTPIWTV--
YRTGGYTVNLLKDIHWLDRGSRLCLVEFNLFNENTDIFLIAEIPPTGGVIPLMTVIYIFWYIMVIYYTIYEYFCSMLNYQSLDVLCFWNIIYVDMMAILAFLVWIKIFKFISFNKTLVQFTTTLKRCSKDLAGFSLMFGIVFLAY
AQLGLLLFGFITSILTMIRMILGLGPIYFLTYILLVFFLLNMFLAII
>695457291
EKIKRIARN--LVEASSYAVFLILFILATTVMWAFLQGPFADVVYGGVRLRTLRVTEAQTVRGFPGSGYVVDLLRVDRYLDLPTRALIVDFTVYNAYLQLFIVVEFPASGGAFVITL------
EGFLVFYILIQLTASLWNLQGLATVLYAES---VIESINAALVWAKLLQYTTISKRMSLLLRMLGRAARDLAWFFVCFLVCICAFAQIGLLLFGLGLAIVTLLQAVAGAGPVFYILYYLLLLLLVNVFLAIL

>551655682
NVSLAYAIDK-
IREIVFYLFFLVIFTTSALIFWDYMTTVIPQYLYQAPRIRQLRVLATSEYRNYTGGGFAIDLLKSVKWSDLHTRAIFFTFALYNPVNGLYLLFEFLPFGEVRPATVLLDAMVYLMVFSYILQDYWTHFWHFEAIGWTYYQMV-
--NFQAFNNIFVWLRAFTYIKYLTDIANLSYTLTNAAKDCGLFLLIFAMVIFAYAQAFHIAFGLLESIFGLFKTLLGLGPLLFITFEVVCYFFLNMFLAIL
>923132353
QQVTGYVLRA-
SLEVVLYILFLVAFCLMAYGVWQYLNGPFLAGLAGMARLRQLRVVGGFEVLGYDGGGFVEYFLQTNNWLDVQTRVVVVDFSTYNANINMFMVFEFLPSGGVVPFQIFLELTVAAFVLFYVLQEYLSDGWNFTNLVTLS
SAEF---NILSLNGFLLFFKVLKYARFSKSALIITKTIKRAVLGIAALTVMVMFLLLGFAFFGNQSFGFTFSLLTMVRGLLGMGPIFYLAFVILFVFLMSVYISVI
>923138555
APSTGLSLSRLTFSTLRYFVFLTLFTIIVFGMFQYMEGPLLEGLFGLVRIRQVRITLFGGVNGYPDDGFVVDLLKQNKFVDLQTRLLAIEFSVYNANLNLFIAFELPEAGGVVPFVLVCELLLCAMVFMYVVEEFVDDMWN
LQGVALMATQEI---NINAINMLLCWIKTLKYLNHFDRIHQLSETLVRAAADTVLFMIMFLIIFFGYAQAGYMMFSIFRSMESLVLSILGLAPLFFFSYMLLVFFLLNVFIAII
>1207165955
LKLVRGQIKKQTRDLMLHLIMLWMIIFVAYGWWNWTTTSLLEKLYGSPVITKIKSKKLAPFCGYKGERVVYVSLLLETGWMTPFTQIVILQFVLYNGPSNLFILVEKTSTGALLPSVMICEVTVIICSLLLYICSFLKKEDFLPI
CQLPIYHLQLSQSLLGIVVNLLLLKCLHLLRLNKTVAAAVVTWKLIFSNLIWLLAPGVIIALALSNLCNLILHLTRLLFTRFLYII-VPPTFFGGLVFFTMIVKAVLIAIF
>19923084
LRGTRQRMRRELRDISMDILMLLLLLCVIYGWWDWSLTTLLDGLYGSSVIRQLKVLNGPGGCGTRED-
CVLSLLRASMWIDRSTRAVSVHFTLYNPPTQLFLRVEILPTGSLVPHLMLPQLVFLALSLIHLCVQYWRKPRNFMDLTLMASWNQRARWLRGILLFLFTLKCVYLPGIQNTMASCSSMMRHSLPSIFVAGLVGALMLAALSHL
HRFLLSAFPGLLFHFPRRSQA----MACYFGILLICFGMLRGFL
>1039738876
LRVTGERLRRELRDMTTHSIMLLLLLFIAYGWWDWTLSTLLDELYGPPVVKLLKIPGGPETCGVKKESYMHSLLRASKWIDHSTRAMSVHFTLYNPPTQLFLGTECLPSGGLVPLLMLSELLFLVLNVIHLCFQYWRKPRH
LVDIGLMVSWHQRARCLQGILLFLWMLKYVHLLSSLSTMTPFSAVT---CFPLFRVLLVGALLLAAHYHSRWFLLFAFPGLLLQFPGRSKV----MRCYYGTLFLGFRMLRATF
>1238871302
LIESRLQAQKAFRGILGFVTILLILVVVAYGWWDWIQHDFLSSFYEQPRLRLFQGWGTRKLWGYWGGSDIANLLRNSSWLNRNTRAVSVEFLLYNPVFSLFLLTEFPPSGGVTVCILACEFLLFLVIGWRMKQEYWLDPW
NLPTFIQLAEWNQIIRGLTGLLIFIHLIASISLLELDTTFKKSLFALRTVSKQLFLLGGLYLFLAIGYSLLSPLVSGLLVSLTLMEAILNRGEIIMLLLVFGILVTITSLIRVIL
>260823633
LRPAQEKSKQEIRQGWLYGTLFVLMMLIAYSWWRWSQGHLLDHMYGEVTLRQLRV---
SSHSGYHGGGYVVALLQSQQWIDEKTRGVAVEFSYYNPPTNLFLLLELPSLGGAYTFIMACELVMLAVVLYTIKSECYCNPWHFVDLSHIADMDQVVSALAGLMLLLLMLKMLHLLRFYSVFTTFSMVFNRASWEIFTFM
LWLLSLLVAFSHLGTLLFGFWLTLQTLCHTAVGQLGLYLAAFFALVMVLLGLIIGIL
>585678603
LQKAKERMMKELKGCLFYCILFTCVMLMCYGFWAWCENDLMDGLYGKVSVRQLRVIVRYCKWGYDGEGFVTNLLKEDNWIDVHTRLVIVEFTLYNPPTNLFLVLELPGTGGIYPFIMVLELMFLALVVYSAKCEHFL
RIWAFVDMSFLALWDQIWYTVIVAI----------YSGVGVIVFGFGYAVQTMTSTIR-AVPVFYISYFLTLLILLG------
>780160480
LQKAKEKMLKELRDVILFLILIFCVLWMAYGLWHWLQEDMVDTLYGKPSLRQLRVLLRYCKWGYEGDGYAVDFLRNNSWIDAQTRAVILEFTLYNAPYNLFFVAEIPGSGAVYPFIMACELLFVLLVALISKQEYFRSA
WSFVDLSFLAWWDQRLRDLIGLILFIVTLKSLHLLRFNRMFAMFGAVFEAALGHLLLHMLCYTILETAFACLGNLLFGVYPAYEVISSLIVGLWPMFYITYVICMVIMAGLMKAIL
>1229127695
LQKAKEKMLKELKDVIMYSTLFACLMWMAYGLWSWMNGDLLDNLYGKASLRQVRVLKRYCKWGYEGDGFVISLLQNNSWIDEQTRAVIVEFTLYNAPTNLYFLAEMPGTGAVNPFIMGCELLFLLIVAVSGKREFFK
SIWSFVDFSLLAYWDQILKDLLASLIFLASVKFLRLLRFNQTVAMFGAVIDGAIKEILTFALYMLVLEAAFACLGVLLFSLYVSTNLLNAMVVGLWHLFYVAYVVSMIIMAGLIKAIL
>918292252
LLSSSKKSIKEFRDAAMLGLMIGVLCVMAFGWWEWLQTSFLDQMYGKVHIRNVQH-------
TTNIRNTLTKLMKKDEIINEQTQQIIVEFTLYNPPTNLFLNLRILETGVAEPFILSCELLFVILMLFSMKKLFLGQAWKYVDISFITIWDEVLKCQIGIVLFLSVVKLLSLSRYKQLFSMTESVYSRASTEILIFAVTLFVVIIAYTS
LGCLAFGIWSGFLAMTALLTRGGKLFVLSYVALATTFTAYIFSIL
>871204620
LKSSRKVLIKQLRDCVSFVFMFVLLCLITFGWYTWSQTTLLDVLYGQVQLRQLRVEVDPGLWGYDSSGLVLSLLESQHWLSVQTRAVFVEMTLYNAPSNLFMLLEVTPAGKIFTIVLGCELLYIVIAMWIWCREYFTHFWN
FINLALLTFWDELLRNLLGMLVFCAMLKLLQVLRYHQLFWRLQTVYRRCRKEILTAGALYLSLVMAFSSLSCGLFGLWTSMLTISALSVRLPVVFLFGY-TLCHGVFTYIFAAL
>908481541
LRSSRKKLIKQIRGLGTFIFMVVLLMIIAFGWYKWSQTVLLESLYGHVQLRQKKLTWNPKIWGFDMSGQVIQLLESDNWIDLYTQAVMVEMTLYNVVTNLFLLLEVGPAGNILTCILGCEMLYIVVTLWLLSQQYFYHF
WNFVNLQFLTLWDECIRGLVGLVLFCVVLKALQVLRYHQTFWRFQTIYRRSKNELLLAGSLYFGFVMSFASLSTCVFGLWSSIYTISAVSVHLDLVLFFSF-TLIHGIVTYVFVVL
>1139740091
LIESRKKVMKELLDIVSLVATFLIVLVMAYGWYDWASTTFMESIFGHVQLRKVKASSPYLLFGYDNSGHVIKLLASSHWMDDFTKAVFVEFTLFHPSTGLFLYLEIPDTGGVTPFLLFCELLFILQTLWKLKIVFLHYGWKF
VDVSFVTFWDEMLRCMVGFILFVVILQSLRLLRYERLFLLFGKIYSRAHRELKVLAILLFVLLLGFAALGHALFGMWSSIFTVSAILSGATRMFIFFMTVFGLGSTIYVVSFL
>443717936
LLKSRNDILKQSSETLLFTVQFVLLVLIAFGMWKWMQGHFTDTMYGQVQLRQHRFTGAGRVQGYGGDGYLVKLLEEDGWVNSMTSAVIAEFTTFHVPSNIFILFEIPAFGGMTPLLLACELLLIVCTMLKVRHEYFHNV
WNFMDVSDVHEWDEVLRTIIGFILFLVTLQLVKALRHQPPFSVFAKVYHRAAMDILVFALMFFIWLASYTCLTMALYSVDNTILSLLALLTGVWRLVIVSFMMFVIGLTAYMLAIL
>PdumPKD1L1
LIEARKTTLKQFSETLSFSILLFLLLVVAFGWWQWSHNDLLDLVYGQVSIRQKRVTGASCVVGYDSAGHQVFLLYNYEWLGHQTRAVIVEFVLFHAPSNLFLVLEIPPLGRAFPFILACELLLIIFTMILLHSEYIKSPWNFVD
TSFIAYWDHLLNSFLGFIIFLLLVMSLRLLRNQPLFSKYGAIYKKTVHEIITFGCLFLVWVCAYTSLGVALYSWDRGLGTVVGLMTRGSRLLVISFLLLALSLTGYIVAVL
>919066838
LIKARRRNMKELWDILSFAFVFLLLIFMAFGWWQWGKTDMLDSLYGYAILRQLRVLQYSSIMGYDGNGFIVRLLQTAGWTSRQTRALILEFTTFHAPTSFFVICEFPPTGQATPLHLACELLFIIVMLYRAQKVFFCHPWNF
VDLTFVAFWDEVLRTLVGLMLFLMLTQFLRLLRYSRLYAKIGNVYKRARWDLLTFFLAFVIWVMAYSSTGVALFNLWRGIFSVSALMTGLAAIFLISFMIFALGLTAYVIAIL
>1087812391
LEHCKLNAKKSLVMLTLHGLFLWAVMAFCVWVWGWMEQVMSPLLFGNVRIRQQRSFTGIGYWGYDDSGYSKYIMQEDSWIDRYTRVCMVEFTMLNPASQWFVMFEIGSTGAVRPDMLFI-
SFIGVLGFMFWVYEYAKKPTNFTNFYKMAFLSNLVTYIATILLTILLIKLTRLIRFIKRWSVFGDTIYCACKHVLGTAIMLLVLMVGYALFGYLIFGFQNSLYSMFRALRGFSICFFFSFYFCVFLFCRLFSAIY
>443732271
LQQAKEESRKYIQHFVVNLILLWLLLSVNYSFWDWVEGTASCVVHGEARLVQLRSLNSIPYHGYSAGGYVKLLLRQMNWIDRSTRLVFLEYTKYNVNKDAFWAFEFTPSGGVIPFVIICQVLLVPISISLAASLFFCSFWHHI
SLLPLISAREVWRMVNAMALFLMMLILARQLRFLRPLSVFGRTLGHAAPALCGCAFMFTFVHIAFSSLANQWFGFRISLVYLVVLLRAFTSVFLLSWLSVVLVATGLVIAVF
>205360954
LFLAKEEARKVLRSLLVYMLFLLVTLLASYGLWPWMAHVLLPYVHGPPRLRQVRLLLGAWSWGYDSGGYVQELLQLHNWLDNRSRAVFLELTRYSPAVGLHLRLEFPAAGRALAPLL-
TSVCLLLFAVHFAVAERVLRLGAFTSFDQVAQLSSAARGLAASLLFLLLVKAAQQLRFVRQWSVFGKTLCRALPELLGVTLGLVVLGVAYAQLAILLVSVAQALLVL--CPGTLCV------GLWALRALRLGAVIL
>1207107604
LSQAREEARKVLKGFLLYMLFLMVVLLLNYSVWAWLSNSLLPRLLGRARLRQTRASSGVWHWGYSSSGFVQDLLNTHQWLDPLTRALFLEFSLYNINTDLLFLLEFPVSERAQLPLF-
LTLILLVLLIYFSVREYFLRVWNFTDFFPLAQQNQLLTQMSATLLFLLVLKASHQLRFLREWGVFGRTLRKSFWELLMVAVTLLVFLLAYSHTGHLLFHVSSTCFKLIGSSGRSCFIFYISFTLLLVLL----MSAL
>260801585
LLQAREEARKVIHQCIFYLLFILTILIAFNWWEYMDSVLLPTLLGGVRLRQIRNYNGLMYYGYSDNGYILDLLNSSDWIDRGTRAVFVEFSLYNPNVNLFLLLELPVTGVGRALLMIMQGLFALFLLYYIVVEYLKRFQTFI
SFFHSAYLGQVTAYLNGLVIFLLMIKVARQMRFRRQFSIFGKALRLVTWELAGCFLIFFILILTYAQLGYLLFSMWMSITAMVAMFRGLAFLFFTSYILVVFCVGRLFGAIL
>PdumPC1
LMQAKEETKKVLRQMTAYCLFFWIMVVLCYTQKSWYENVFFDVLQGIPRLRQIRSALGQTYRGYHGGGYIINLLKMQNWIDRSTIAVFLEATFYYPPKNTYLAAEFPLSGGAYYSMTVFLGLLAATIMYLLIHIYLKEFW
NFISLAKNLELDTRWSIINAFIMFMLTIWIVKQIRFIKSWSLHGNTLRVSAPALLGISFILTLVILAYAQLGYLWYGFGSSFCNLVVLIRGFTQLYVMSFVSITLVIMGLVIGIL
>585696579
LIQAKEEAKKILRQFIVWVLFLSLILLINFTFWIWSKNILARGLHGVARLRQLRTLDGMSFKGYDGGGYVQHLLLANKWIDRWTSAVFVEFTVYNSNVNLYLLVEFPASSGAFPASTVFMMLLGIFMVYFIIREYFASLWHF
TNFYRAVYYIEVVNTLNAIIMLLVTLKLLRLLRFNRMMSTFSRTMSEAGRYLFCCTMIFVIVLWAYVQLTYLVYGLDAGMGFLLALMRGLTPFIFSFFWFIFCVIAALFAAIM
>919023136
LIQAQEEAQKKIRQTVVYFIFIWLVLVINYAFWQWADTVLAPSLQGNVQMVQVRNIDSLPLMGYHGGGYVQTLLRDHSWLDERTRAVVVYFTVYNANLNLYLLVEFPLPGKVLTPVLIIQCLMGAVLLYLIIHTYFTWF
WNGLKIHRASFYNCVLTDLSAFLLFLLCLKLTRQFRFVKSMAVYGKTLNEAATYIMGCFILLVIFLISYAALGYLFFGFSTSLLTLLASIRGFTAIFFISFNVVAHGIGGLVMATV
>405970250

LLKAKEDGIKIFRQFVAYMFYLLLVMALCYTIWKWTQSVLISSLHGVARLRQVRGTGSKRREGYGGGGYVQSLLQSKDWIDLRTRALFIDFTLYNAAEDLTLMIEFPLTGGINTPLMVCEVILFAVVLYYLVHIYFCDFWN
FTNFERVADIHTAARYLHAWLLFFLMFKVVKQIRFIKFLHKFEKTLNIAFPKLMGVMLIFGILYLNHGMIAYLFLGYWTTLMTLLGMIRGFTHCLFYSFYIFTYGTIALIIAVL
>1205910066
LVKAKEEGQKVLRQFVAYMFYLWLLMVLCYIFWVWSQDVLASALHGTARLRQVRSTGSKWRQGYGGGGYVQFLLQAKDWINLHTRAVFVEFTLYNPATDLTLLVELPLTGSINTPMIVCEAILFVVVVYLWIHMFFTS
AWNFVNFETTVDVHLSMRYVQAWLLFLLMFKIVKQLRFIKILYIYERTLSESVGKLSGVALIFAILFLNHGMVSYLWFGYWNTLTTLLGAIRGFTHFFFFSFYIFVYGTVCLVVAIL
>871263454
LLQAKEEGKRKLREVFMFTLMLALFMAIVFVFWQWSETVLARSLHGPAWMRQIRSLDLYSKFGYPGGGYVQTLLHSEGWLDLLTRAVMVEFSLYYPSADLTLIIEFPISGGVLTPLLVCQVFFSLFVLYLTLVLFFSYFWNP
SRLGQVAVLHAVLRCVHACLLFLFVLKIVKQVQFSRPLFKFYTTLSTASGPLLGVFIIFVILLLTHAQFGYLIYGFQHSMTSLFGIMGSLTHFFFLCFGFFVYGTMALSVSVL
>908506947
LLHAKEEGKKSLREAVMFMLMFALFLSVVYSFWQWSQSVMARVIHGAAWLRQIRGLNMYSKFGYTGSGYIQKLLETDQWIDNKTKAVFVEFMLYSAGVDLTVLIEFPITGGLLVPILVCQVIFLMFVIYLSVLLFFLGLW
NSAHLAQVVLMHAVYRYLHACLVFLFCLKVVRQIRYIRPLYKFYMTLSVSSCPLMGVFFIFIILLLSYAQLGYLIFGFEQSLTALFGLMFGLTLIFFTSFCFIAYGVTALSISIL
>565304042
VTALQKKQMIQVKDMICHFIFLALIFTISHPVYRWVNNNFLPLIHGLPRMRQIRRHWDYITEGYPSRGYTFYFLETKSWIDNKTWAVILELTTFNADEHMFVIFELSLLGPVNASQIFFFCLTIYFLIIYIADEYFTLAANFIPF
HAASQFDQTVRGVLGILLFFLVLKTYRYVRFLFDLRLAQKSLRTALPGIINIVVMGGIFFSAYMALGYVGFGMMYSFMTVLSYCALLGGLFLGSFLLVMICFINLSQVII
>5174633
IRIFKRKKRIKLSYILTHFIFLALLLILIVLIYRWLNSVLLPLLHGLPLMRQVRAQWLYYSYGYGSGGYALYFLQESNWLDEKTWAVVLELTTFNPDINLFVIFEVSQLGVVNTAEIYLYVAILIFFLAYVVDEYVRSVYNFIPF
HAVSQVDHIMRIILGFLLFLTILKTLRYSRFFYDVRLAQRAIQAALPGICHMAFVVSVYFFVYMAFGYLVFGLIHSTQTVFSYCVSLGVLFLSSFMLVMICLINLFQAVI
>115583675
IQILRRENRIRLSYLVTHFIFLTLLLLLIFSIYTWLHGVFLPLLHGLPLVRQVRARWVYHSYGYGSGGYVFFYFLEGKNWLDELTWAVIVELTTLNPDTSLMVVFEVSPLGVVNSSEIYLYAAILIFFCAYVVDEYIRSVYNFIPF
HAVSRVDHFMRIILAFLLFLTILKTLRYSRFFYNVRLAQKAIQAALPGICHTALVVSIYSFMYVAFGYLVFGMIHATQTIFSYCVSLGVLFLSSFMLVMICFINLFQAVI
>734563164
VSTTRDRRMRDTREVVFVFVSVAICLGFAFS----------
GYGIVRQVRSLQGRPFYGYSGGGYIQTLLENESWINEGTRAVFIEFSVYNAQTNHFLVLEIPPFGTVFPTATLFEVLYILLCIVLALKEYFSSFWNYIRLDRQRDLQLIFLALLSAVVFLTCMKMINILRFNRRIGLFTQTLSHS
APTMVIFGAVFCIVSLAFDSALFVILALTTVAKTTIISLLG-GAAVYIIFILAGKILLNFFVMLV
>392891400
LRDTENRKMRDIRDMLCFFASLYIMVMLTYYFWDWARESLATALLGIGTIRQVRTLSTETVSGYSGGGYTISMLDSERWIDDHTRAVIIEFSAYNAQINYFLLVEIPKSGIYLPVVKYYEMLYIFFSVLIFVKEGNFSPWNYI
NLTEQRNWEIVFSYCLAGAVFFTSCKMIRILRFNRRIGVLAATLDNALGAIVSFGIAFLFFSMTFNSVLYAVLGLMATFQTALAGMLG-AFVVIMLYMIAGSKVLQLYVTII
>Pdum688
IEKQRKRAVLTVKSLVAHALLLWVLVSLSIMFFDWLHNEFIPNYFGPARLRQVTILDSLPIWGYGGGGFAASLLNEYRWLDRQTRAVFLEFSLFNANVGMLLLAEMPNSGAAITHLVICAAIFLIYVIYYIVSFHFTKVMTFV
SFAYVAALDASFRYCVAFTLSLAIIKLIFLLKLYRRISQLMLTLRKSIKPLASVLVALSIMVLSYAFYAYLYLSIISSIESMVGMALGFTWFFFTTYIILMNFMLSILITVI
>961088331
RKIRHRKIISEVLNGIIYSLLLISAGYIAYVYFDWMENVVFPNFFGPARLRQLRVINGIPIIGYGGGGYINNFLIRAKWVDRQTRAVFLEFTLYNANTHMFFLVEFPESGGAIPTITVFQAIFIFCIVLLLIKVFFSQFWNFVNFY
HIVIWNEIISIALALFLIFFAIIRLSCILGVSQTMLQISLILNRCVSHLFGVSFLLCIYAGIYAMLGYLIFGYSSSLISVFLIVLGLGKIYFFSVVFFMIFLLKMFESVL
>Pdum12681
ITRLRRKKLKEIQEVIVYVLFLLVIWTISYSYLQWIKDQLIPWLFGPARLRQLRMVSGVPMWGYSGGGYVANLLWDNLWLDWKTRMVFLEFTLYNANANMFLMSEFPEFGGITQFVIACEILFICYLIYLIVRTFFGNGWH
YVNFQIVAMWDEYFNIVLALLVFVYTMKLIGLLNVNKHMGSLLSVLRVGKEELISYTVMFIVVFLAFSQLGWLLFGMMRACEELISTMLGGAQTYYMLYILVSTTLANTLICII
>871245685
VVKMRNLKLQELVELVLYIIFMFCLFSFSYSVYTWLNTTMTTRMFGPVRVRQVRTENAIPVLGYSGGGYIIDLLYQNRWLDLQTRAVFVEFSLYNVDANLLLVVEFPVYGSGFPYILMCSIVLVTILIVFFFKVFLKSFWGFV
NFYALAIYDDIFGWILAFVVFLSTVRILRILGYNQRVTTIAAVLQRAGKDLFSFSFLYLLMIGVFVMLGYVLFHMWNTFVTIVYLLLGVVRIYFILLSFTLVFLLTMFQAIL
>676439278
LVRARRRQEQELIEIVIYTGFIIALCVIAYGFMAWTEKNLFPFAFGPVRIRQIRIDWGLPLIGYGASGYFFDMLYNNLWIDRQTRAVSVEFSLFNADANLLLIAEFPEYGSAVTMVIVCEIIVAIVMIVTLVKAYFQDFWHFVN
FEHIVWFDSMFLYALAFLLFFTTIGMMRVLGYNKKITHALVLKRSAKNLGSFGVLFILMFFGFISSGYILFHILYDALFTCFTYLLGLGALFFSLFVFYVIFLLTMFQAIL
>1139824403
LENMKRKRLQELVELLLYVCFLAILYFISYNFYDWLRNTLIPFVFGPVRLRQLRVIWGVPTIGYGGGGYVADLLYNNLWIDRQTRAVFVEFTQYNSNANLFLMTEFPQTGGVLTFIFLCEIVLMICALAFFVYFFFRDTWRF
VNFGHIALWDEVLNAFLSFLVFLSTIRILRVLNYSRSITHLAGVLSNARRNLLGCFFMFALIFIAYAGFGYLLFGMFVSITTIVNSLVGIAEFYYFTFVTFVIWVMTMMCATL
>762108557
LESQRKLRQNEFRRLVVYAIFVFVVFSIGYIFFHWMRHTFLPRLFGPALLRQSRMVWGMPMEGYAGGGYMIKFLERTLWIDRRTRAVFVEFTLYNPNVNLFYLLEFPDSAGVVTYALLCYFIYIIFFLISSGRMFFKSTWNFI
QFQHIAFWDNAFVVVMGILVFIATIRILQILGYNRRMTQLIAVITLASKDLSGFAIVFSIIYSAYVVTGYLLFGLFVAFGSLTNSIIGLGELYYLTYILIVILMISMFAVIL
>762157119
IDMKRAKRQQELTKLVISAVYLSALYSISYLFINWLNKTFVPTMYGPAQLRQLRTIWGTPKTGYDGGGGYILRFMIEQNWIDRHTRAIFVEFTLYNANSNLFYLTEFLETGGIAKFTILCYTTSILFILITTVQVFFLHPWSFINF
NHIIIWDNIFNAVFATLVFLSTLQILNILGYNKRFTQIISVVVNAGSDIGGFGFIFLIIYVMFVILGTLLFGLLTTCGTLANSFIGFAEFFYVMYMFCVIMLITIFAAIL
>405968482
VAIMRKKRQEEMITLLVYAIYVSAIYSISYMFFGWLKGSLLPTYFGPPRLRQVRMIWGIPKAGYSGGGGYIISLLLTEQWIDRQTRAVFLEFTLYNPNVNLFFLTEFTEFGRVLIYVIACYSIFIAYVIYSFVKFYLKHPWHFINFA
NVVIWDNTFSIIFAILVFVSIIQFLRILGYSKKVTEILSVTINVGRDLYGFLIMFSIMFFTFVGTGYLLFGFFSVIGSLTNTFIGVARMFYFMYILLVIMLFTIFAAIL
>1101394602
VAAVRKMTFLTLTDIGWFVLFMFSVFVFTYGLWLYLESTFLVSYSGPARMRQVRILGIGHFMAYKGGGYVAVILRNSVWIDDLTRAIFLQMNTYNPSTNVFLVFEFHQSGLIIPAAIVFGICAVIFLCFLTVSLYFRDPWNFV
SFYFAVLTELAFLDMAALFLFVCSVKLLKIFRINRKVNSLFYVMNDARVPLVEYGISLLLAMTPFIAGHVIFGFGASIVTLVRVMVALGPVFFLTAVVITIFYTNAFVSLI
>919100788
EERKRLRRKHDLREIFHHIYLILLIFICFAIFAWLSDRILPIVFGPLRIRQKRVDKDMAATGYEGGGYFVEILKADSWIDRLTRAIFVEFTINNMNVNVFIKVELPVVGGIFTVLLVIQVIWALVVFYFLLVKEYFRELWHFVGF
DIVANWDYGYSITIAFSLFLATLQLFRPLSFNKQLAILRASLSRSSIYLFTYGLIIVIILTSFGIGTHILLSGFGAALFNLCGVMLALLKVFFMMFAMVMNFIINMFISVI
>Pdum14759
IEEAKKKKEMDMREIFMQAFFLMMLLIIAYGIHAWLKTNFTDEVYGPPRLRQLRVLRGLPFWGYDGGGYVANLLNSSRWIDQYTRAVFVEFNVFNMQSGLFISFELPPTGAIHFIAIVTEILCTLVFIAMTIYLWFKSFWN
VANFQIVTTADNIFLTFLGLATFLVMLQFLHILRYNKTIGVLNATVAQSASDLIGMGFLVFVIFAAFLSLGYLGFGIKNTAIALCSNLLGPGQVFLITYLCTMIYMINMFITLL
>Pdum15290
DENVLKLRRLDIKEIISNFLVLFVMMCVAYVVWPWVEKVLIPSLYGPPRLRQLRVLKILSYWALPGGGYVANLLKAMQWIDHRTRALFAEFNIYNANTELFLAFEFPSNGGVHPFSLIMEIISLFVFVGLLISSYLKNLWNTV
SFQHVTLGDQIFVCLFGLVIFLLTLAFLNILRFNWTIAVLAQVLVSSAKDLTLMALNILIMVMAFASMGNVLFGLPDTMTTLSTNFLGAGRVFLIVYLVTMMFVVNLFITLL
>260824333
LRSAREVRIKELRTTALHLAFLSVIVIVCLYLWMWAEEILLPNLYGTVRLRQLRVLTSAAFYGYRGGGYAVILLKKTGWLDKYTRAVFVEFSLYSPSSSLFVVIEFTTVGDVTKVVITCKALYCMFVLYFIGREYFRSFWNFV
NFQSAAWWNDVYSYVISCLVFASIVKFLTLLRFNGRMAMLSYVIRRATNELVQFTVMFFVIFLAYCQVTYLLLGFQASLEMMFAILLG-SCVFLLTFILTLVLMVNMFLAIL
>828232853
IESMKAAHLTETWELLMYFLFSIVVFYLGYVFWPWMDSFFLPLLHNGIRIRQVRVLDGYPLSAYSGGGYVVEILKQKKWIDRQTRALIIEFALFNAATNYFIVFEIPSFGGMVPVTIGLQIIFLLLMFVFTTREYFHEFWNFIN
FQFASYWDIMYTYIVSLLVFFVTLKFIKLLRFNRRVSLLSSTLNVAFYPLSMLGITFVVALIAIISFCTIVFGYLGTISSVISLFLGLGPLFFLGFNIFMIWLMSMFISIL
>260798008
LLKARQQRFLELMELSFFITYVFILILVANGFWNFMKDKVVPELYGPVQLRQLRVTTATPHYGYYGGGYMLELLEMEQWIDSYTRAIFIELIVFNPNANLFLLAEFPGLGKVYTVVLAFQIIFVLVTFVFVYREYFRLFWNY
SKYKQAASWDQLNTYVQAWLVCAATLKLMHLCRFNKHVERSAHTLTRSAKPMLNYMIIFVIYVSAFCMLTYLVLGYTNVETMLGVMLGLGPVIFFGYLTFVNFLLMMFCAIL
>1126169400
LAEQRTESKQKIVEVLMFGLFVTVIMLTSYGFWTWVITGLIPATHGPVQLRQVRLTDAFPYFGYLAGGYVTSLLQRQDWLDDKTRAVFIELTLYNPHVNLFMAVEFTNLGATYKLLLALRGLFALFVLVYALKEYLSEFW
SFEGYKNAVSWYQVYTYLLGLLICCSTFKFVRILRFNSHVYALSMTLRRSLKPVTQFMFTAGIVIMAFTQTASLIFGIASSLQSLLLMMLGLGPLMFFMYQCMMQFLLSMFMAII
>641798600
APPRRGRTWKKLGEIVAQLVLLASLMVVCYTFYRWAHGTLLPSLDGSVRLRQLRPLGEHPVWGYPGGGYLAHLLERSRWLDGCSCAIFVEFTVYNANVNLLLLLETPGTGAFLPQLTCAQVTYLLLLLFYLWVQYFRSQR
NFVSFYEAAQVASALTYLLAFLVALATGKLWNLLRLSPRMQLISGTLRRAWQEVLGFLLVLLLLLVGYSMACNLLFGFFSSSVTIVGLLIGLGSLLITTSVISMLYVINLFVSGL
>115583681
LAK-PTRKAWKTGGTLVQILFLTLLMTTVYSFYPWANGTLLPNLYGNVLIRQTRILGGYPIQGYSGGGYVVRLLEQRRWLDHCTKALFVEFTVFNANVNLLLILESSGVGTFLTWI-
VSQVVYYYLLVCYYAFIQFFTRKRNFISFYEALRVNSAVTHLRGFLLLFATVRVWDLLRHHAQLQVINKTLSKAWDEVLGFILIIVVLLSSYAMTFNLLFGFFRSIVTVVGLLMGLGSLLVLSSIILMGLIINLFVSAI

>31559825
PTKHPERTLKKTGDILVQILFLTLLMTAIYSFYPWANHILLPSLYGVQLIF--QI------EGCDNG------
LHLRRFSYICSPRPMVLIPTDELHERLTNGFSYIMRGAFFTWSIISQVIYYLLVCYYAFIQFFTGKRNFISFYEAVKVNSAATHLVGFPVLLATVQLWNLLRHSPRLRVISRTLSRAWDEVVGFLLIILILLTGYAIAFNLLFGFFSSAVTVVGLLMGLGTFLILTSVILMVLVINLFVSAI
>48094207
VNQVRERREAELKELLGQILLVVLALEVAHFVYGYLEESFVPELYGRARLRQVRISHGSYHMGYTQGGYIQVLLEDRKWLDELTRALYIEWTVFSADAGLMVLFEATTMGNLVSINLALQSVYLLVILGLIHLEYFQ-FWNFISFFYPVIMDQTFMAVTGFLVFINTLRCAKILRFNSSIQTIHHTFSNITSEFFIFTLMVLIFLFAFALLFCVTFGVAFSFQTIFLFLEGLGPLYFIFMTIALFFLFNFYILLI
>1229186594
VEQARERKRHAEIKAFLGHLVFIWLVLQVSIAVYTWIKGEFAPSLYGSPRLRQQRVAKGSYFLTYNRGGFLLDLLQDLGWVDQYTQAIILEWVVYNANTNLFVVFRAEGTTTVVPFVVILEVLYVMVVTVIMGMEYWA-FWNFILSFADIVVLDSVLVTSLASVAFCSLLKCLKILHYNDSMYLLHSTFRRLTKTLLIYLSMLTVTFVAFAYYGYLSFGLLRSIQCVFMLILRQAYFFFALANFFLIAIVNLFIAML
>260827108
LEKVRAARKKNVYETAAFMLFAFLILTIAHTVYLWLEEEIFMNLYGSVRMRQVRVV--W----
YGTGGYVADFLEDQKWIDRQTRAVVVDFTIYNANVDLFVLIETPDTGNARAVVLIFQVAYTLVIIYDIYAAFFRNVWNFLNLVPAIVADEALVAVVAIIAFVGCMRCCRLLRGNNSLYLIEATLTRSMAPLLSHGLIIVIMMIGYAVFGHLAFGFPRALVTVLNFAIG-GSIFFISFIFSMGFLLNFFAAIM
>919052864
LKSAKQLRYQQIREVIGYIIYCWLLMLMAYGFWVYVKDTAITNLYGSPRLLQIRVLQYMWQSGYHWGGYVADLLQDATWLNEGTRMVVLEFTAYNANVNQFTLVEVSVGGGLVLFVMACEIIYCIYVLYFIVAEYFISGWTFVNFEQTVLLDKALGYLIAVLVMLSSLKFISILRFNPVVHLLTQTFAKAAGQILWAVALVVAYFIAYAFLSNLIFGFRVSMESMFGFIRGVAPFILMSFVIVVIYFLALFIAIL
>585652993
VHQSKEAKIRVMREIFGYIFFIWLLLIIAHAFFNWTSDRLMPGLYGGARIRQIRSNPYWG---
YSGDGYVCVLVEDALWLDSHTRALFVEFTLYNANSNLFLLIERPDIGGLLKFVMSCEIAYIGAILFFSVIQYFMCAWNFVSFHECALLVQTFESIIALLVVLGTVKFLHLLRINPKVYLLTAVLASSGREMAIFTLMLMVIFTGFACMAYIVVGIAVTYETIFSIMLG-AAFLSISFVTIAVFVLNIFISLL

Tree file (Newick format)

(((919074653:0.38979406,(918302224:0.33053333,(((676474191:0.20094166,(908400046:0.17522307,(871215788:0.33172675,871254812:0.06827018):0.150000:0.02933241):0.989000:0.09990901):0.569000:0.04793804,(762156788:0.16765294,1207917521:0.11017855):0.981000:0.09977293):0.965000:0.09609497):0.953000:0.07855474):0.829000:0.05081624,(443713927:0.33909243,PdumPKD12:0.29234623):0.954000:0.12368775):0.058000:0.03424052,((((1229159547:0.49288212,(148539638:0.20531707,1229145338:0.23556990):1.000000:0.43666691):0.963000:0.18049853,(((((1087817330:0.84865810,1126215987:0.71313402):0.607000:0.17465564,(585644120:0.61900470,585688456:0.68467157):0.735000:0.12122586):0.274000:0.09536684,((585651646:0.50031640,148539556:0.68933196):0.917000:0.14493618,(1139813591:0.71686103,((871242653:0.48645780,(1205898338:0.18114693,762107655:0.23237343):0.459000:0.10779839):0.980000:0.15296607,(919018042:0.17442258,PdumPKD11:0.48302762):0.627000:0.07115052):0.997000:0.35243608):0.525000:0.07319933):0.294000:0.06391446):0.796000:0.04112858,(((Pdum688:1.17114430,(Pdum12681:0.76175837,((1139824403:0.49835320,(762108557:0.47945800,(762157119:0.48305867,405968482:0.43934531):0.927000:0.15397768):0.986000:0.29921254):0.833000:0.14587945,(961088331:0.96141646,(676439278:0.67481985,871245685:0.60657353):0.958000:0.25103283):0.204000:0.03928805):0.971000:0.22673050):0.735000:0.05304772):0.953000:0.14950796,(((116008215:0.91662873,(((71999568:0.49840308,1253289448:0.22811452):0.999000:0.28751462,(260802732:0.24908472,(47551289:0.25002426,(291234667:0.21360076,(((((685230872:0.27797002,(1564010113:0.09011251,(Chem014425:0.16150375,828192580:0.32510283):0.944000:0.08601326):0.931000:0.06093523):0.968000:0.08192114,(1067079595:0.19617922,(926626589:0.13930123,(675370064:0.15546914,241171514:0.15845318):0.817000:0.03967445):0.967000:0.07732506):0.990000:0.10238715):0.858000:0.02571418,(198420844:0.13922306,(((Pmar7053:0.10538127,Pmar01872:0.16265907):0.972000:0.06229505,(359465612:0.08590505,528498254:0.13934932):0.982000:0.08691814):0.847000:0.02436635,((556956196:0.15347385,(1160116546:0.37067822,(5669804:0.24847798,1033394660:0.09415394):0.660000:0.04264957):0.335000:0.04599642):0.960000:0.07651533,(1160100767:0.13499141,(4505835:0.11750539,(557013709:0.07483225,50539686:0.22244767):0.114000:0.03719950):0.913000:0.06434456):0.993000:0.10168329):0.586000:0.03718946):0.928000:0.04434182):0.726000:0.01231048):0.853000:0.02439261,((PdumPKD22:0.14029409,(9190766784:0.25084890,(PdumPKD21:0.09750558,(919076662:0.13981491,Bner206993:0.36218276):0.882000:0.06077596):0.081000:0.01118391):0.865000:0.027170840.889000:0.02675046,(((918287730:0.12883469,((405953847:0.06800565,1242857867:0.35002356):0.948000:0.06695499,(871233656:0.07104406,908475875:0.27370973):0.910000:0.03661231):0.973000:0.07191833):0.724000:0.04463960,(1236532622:0.20288088,(675875990:0.29916220,(Ptor32544:0.17804277,(675439102:0.27683787,(847028821:0.32374312,847028815:0.12526846):0.469000:0.04342791):0.522000:0.05339054):0.949000:0.09226190):0.411000:0.05506511):0.945000:0.06004495):0.288000:0.02212216,(675861632:0.25465940,(1174757226:0.42132009,270002831:0.73951989):0.793000:0.07844840):0.905000:0.08722501):0.779000:0.01409663):0.973000:0.05565832):0.562000:0.01830858):0.805000:0.01853921):0.763000:0.02090518):0.966000:0.13281938):0.848000:0.12630665,(551655682:1.38304007,((923138555:0.82902087,(551658247:1.27894516,(695457291:1.00920609,(470247587:0.28423787,66823157:0.20006608):1.000000:1.27835095):0.710000:0.17363545):0.778000:0.10011675):0.898000:0.12329366,(923132353:1.14739825,673034662:1.42796591):0.000000:0.02149203):0.806000:0.10474314):0.933000:0.19820145):0.037000:0.09059135):0.997000:0.32867356,(((565304042:0.57322795,(5174633:0.08775955,115583675:0.17266816):0.974000:0.32255399):1.000000:1.35930219,((1229186594:0.91131598,48094207:0.96894318):0.939000:0.34009540,260827108:0.87403848):0.000000:0.01111743):0.952000:0.36202977,(919100788:1.19782507,(Pdum15290:0.63065770,Pdum14759:0.61027520):0.988000:0.36595686):0.775000:0.08555485):0.941000:0.16166374,(((260823633:0.74353971,((585678603:0.48968739,(780160480:0.36306555,1229127695:0.28647123):0.975000:0.23371890):0.994000:0.29758867,(((919066838:0.57161388,(PdumPKD1L1:0.59051503,4437179360:0.74009918):0.932000:0.18466373):0.439000:0.07579216,(1139740091:0.66193818,918292252:0.88735529):0.377000:0.19504195):0.780000:0.11377474,(908481541:0.41336332,871204620:0.28665657):1.000000:0.59377660):0.990000:0.30558664):0.665000:0.05161463):0.794000:0.13128779,(1238871302:1.42067508,(1207165955:1.69675380,(19923084:0.30006521,1039738876:0.47798689):0.991000:0.67953230):0.821000:0.42784250):0.291000:0.18254237):0.992000:0.36548307,(585696579:0.70849030,(((1087812391:1.31049381,260801585:0.71415181):0.080000:0.05446895,(205360954:0.73974516,1207107604:0.69659001):0.990000:0.43461212):0.908000:0.14730518,((919023136:0.70252465,((405970250:0.27037192,1205910066:0.24180516):0.996000:0.33370228,(908506947:0.37553259,871263454:0.29516245):0.999000:0.44739232):0.985000:0.27083986):0.835000:0.07213905,(443732271:0.91430461,PdumPC1:0.78948019):0.914000:0.20177490):0.897000:0.10493269):0.848000:0.09546256):0.997000:0.37008808):0.878000:0.14197172):0.748000:0.05190226):0.000000:0.00680416,(Pdum9951:0.87882559,676479791:0.55190498):0.987000:0.32322408):0.854000:0.03854690):0.867000:0.04251857,(260824333:0.90566615,(919052864:0.81166887,(585652993:0.89143087,((1207182030:0.19800449,(116006951:0.16236260,942189627:0.08824090):0.819000:0.04439707):0.932000:0.15215114,(641798600:0.37602842,(31559825:0.45125608,115583681:0.21404520):0.995000:0.31298930):1.000000:0.45225868):0.989000:0.33342229):0.587000:0.09570737):0.982000:0.28087451):0.724000:0.07114948):0.916000:0.06820375):0.868000:0.04289251,((((1101394602:1.54625691,918996724:0.57635382):0.915000:0.27729246,(260785879:0.70445602,(260798008:0.72711776,1126169400:0.87471339):0.942000:0.28291951):0.686000:0.05953344):0.897000:0.07784682,((((196000368:0.64114210,(828232853:0.85747448,1005480473:0.37231245):0.937000:0.17009291):0.842000:0.08536096,((828228329:0.74936927,(11910591360:0.28155162,1005450057:0.16423642):0.954000:0.19575096):0.996000:0.37362203,((1005492071:0.49546067,(999981318:0.52906473,(1005458745:0.52395447,999987444:0.41891688):0.995000:0.29722824):0.395000:0.07858671):0.929000:0.12959257,(156406861:0.62264340):0.253000:0.04443763):0.318000:0.06696903):0.019000:0.04520984,((156395095:0.43574185,(1005483664:0.43318890,999981301:0.44800023):0.925000:0.15635646):1.000000:0.36875595,(1139824617:0.67717288,(156394387:0.24594716,1005475249:0.34685195):0.998000:0.28708939):0.879000:0.08383515):0.581000:0.01507451):0.907000:0.05341677):0.829000:0.03131982,(1087817335:0.89246704,(919098890:0.52636298,(919082289:0.53493244,PdumPKD916:0.43975305):0.085000:0.07867135):0.906000:0.08815688):0.785000:0.04326223):0.745000:0.00942799):0.850000:0.03552171):0.983000:0.13735263,(957833931:0.66451242,((392891400:0.81150295,734563164:0.64595219):1.000000:0.77205828,((1067095036:0.98661095,1009565777:0.48422227):0.722000:0.07045055,(675368552:0.21348730,(1078802969:0.26497617,1238854278:0.16172011):0.911000:0.07896523):0.950000:0.13899192):0.865000:0.17672975):0.869000:0.11451914):0.549000:0.10041857):0.296000:0.05931130,(291224264:0.28092660,(780039266:0.39975267,1229191925:0.29556095):0.904000:0.11138548):0.973000:0.14971667):0.940000:0.07832542,(684373764:0.53908044,(Ptor35851:0.08365531,847028803:0.30968071):0.976000:0.22965163):1.000000:0.37652874);

PKD2 phylogeny

Sequence alignment (FASTA format)

>470247587
NIERGSRMVKDLFTHFVFFIIFLVIVVLQRNVYTINMALKNSFMDFKAYLTEQFGPVVT--GMDN--AMNKVISVVRLRQLRVTDHSCF----DLPCYYGTHTRETEPIIGTNYYFT--
KDSHGSLVFGYDTSGYYVDILTDGIQQLYDNDFFSENTRAVIISFVTLNLNYEAATVVYLLAEWTAAGSVDPYYTMRTYRIEMYNDRFRAFLEITFLLFLTYYLYFFINEARIDFITSFWNILEVINLVLFIVSVILYLIFLGDNDRRTDLEILGQTALVFYQLSAINILLLAFKTFRFLMLHRRLYALWITLAHAKVQLITFTTMFIIMMIGFLMSGWLTFGTDIASFNGFVTSLGTLLQFIIGNPDYEAMSYTNRALGPIYYLLFTIFMFFILVNMFVAIISDSYQNISNTFEQKTK-R-----RWERTMKSLFTPNS---TDGIDPSRFKDEIIQ----QVHK--KDLDTWKKEAKIKEKKK---FELNT-------------KLDLLLQAM-QQSIPQ
>673034662
KDFSAHREIKDMIQYLMFLALFIVVTVDVSGPYRVTAMLNAQLNELHQYLTGPFYDVVY--GAFDIGGNSRLVGPIRIGQVRVKAEACVMA--PVQCFYSAATDSIEPFGVEPRFYSHSHHWYGSPTFG--------
VVYDLLASLKEHKYFDKATRAVFVDLNLYTANIDHTTTARLFAEMLPGGGVTTQVEFLTYRLYPWHDFVQAGCEGLIYVVILAQLAREATAVRRVYFGEFSNVAVVCNALLFLFVLAFRVLSYATLPASINFRTPTIYFSLARSCQSFNCFLSWLKLFKFLSFIPMFGQLTKTVQRAAGKVLELVVIFALTLVGAALAFYLAFGNFAANYHTFLSSFYTLLHIVTGEMSLADLRLANRVLGPFFFISFVFLMMFVILNIFIVIVSEAYTDTKKELRD--EMASLSKAIAHHFLHRVLGPRV--AKVNLN------AEVTRRLR---------GRSSVVAA--TEVMKHETACQAKLQDIEQLMPACLE-------HL------N
>551658247
SKRNAIRFVRELIVFVLWVTCFSATITLTRDSYQFGFLFKNWLIDFYSFLQEVLVKILA----PLFLYGNKRIGPLRIRQVRVPYTTC--
QYLNIHCYFSLSTNLDSKFIEPWFTFYTAKQLSEMAYFQYPGSGYVLDVVNTSLKTIQDSGWIDLMTRAIFVDFSFFNPNLNLFLVTRLVIEFLPSGTVTMYPTFRIVDPFRFSDLLVIAAFSAVVLITLYYFLADMLDIAKGFV

KSFWQIFSFLNVLMMLALVGLTVWDEMAARALFDLQATGYMLDQEKNLAGHILIFMWIKLYKYFSMSRKLSTLTRTIAKAASNVIFILIVLIIPTIGFAMGCTLIFGSDDFYFSNFKTSLYTLLRAAYGDFNFQD-
WSSNRYLGPILLLFWVFFSNFLLLNIITAILCDSFAVVMAENERARGVKGVFGKRITNIIE--
KQRQVALEKMDADGDGKTDMELESWLSEDIMEKYDKDGSGLLDAEEMEQIKRWVASEERILSIENNVKEIKEKVEEILNKLPRLELNQ
>695457291
RNPVPFGSLVEASSYAVFLILFILATTVVEEAFYFANRLKTALVQMWAFLQGPFADVVY--
KEETVSSSNRLLGGVRLRTLRVKAGSCLSQIPVPFCYFDSSVEEKQGYGVHAYNWSSAKVTEAQTVRGFPGSGYVVDLSRERLAELRVDRYLDLPTRALIVDFTVYNAYLQLFNIVEIVVEFPASGGAFVQLSDAVLRLFRY
SGRITLEGFLVFYILIQWWR--
MVKDMHRDLTASLWNILTVLHLVVFAVTIAVRLYTIDLAYGTMNLQGLATVLYAESVIESINAALVWAKLLQYTTISKRMSLLLRMLGRAARDLAWFFVCFLVCICAFAQIGLLLFGLDVRGFSSLGLAIVTLLQAVAGELD
YAGMADSHRVAGPVFYILYYLLLLLILVNVFLAILNDAYMQTITEQEEDEELA-----AHNHAMGWD-SGDV--ANADISPSEERALAARREME--------TTACRLL-VADVKRVVSEIRVSRRVELQQSQIENLRHTVDDVAARL-
--KTQR
>923132353
RARDVVMASLEVVLYILFLVAFCLMAYGARNSYHYNALMKDLFLEVWQYLNGPFLAGLA----
DNIYGVNRLIGMARLRQLRVKPGKCLLGAFVQGCLYKEQYESRETFSTAPFRYASESEVGGFEVLGYDGGGFVEYFDASKMAFLQTNNWLDVQTRVVVVDFSTYNANINMFFVGQMVFEFLPSGGVVPSSVFRVIRFY
RYPGRFQIFLELTVAAFVLFYVLQELFEMKQKYLSDGWNFFDIVNLLLFITLFVFRIITVEFNKLQNFTNLVTLSSAEFNILSLNGFLLFFKVLKYARFSKSALIITKTIKRAVLGIAALTVMVMFLLLGFAFFGNQSFGTDVD
SYRTFTFSLLTMVRGLLGELDFESLWANNRLMGPIFYLAFVILFVFILMSVYISVIDDAYGRVQDKYERPNPHADYAAGHMPAKVKSWKQKEIGVLAADTNNDDKIDREAQALLADDLFAMYDADQDGQLDLEEQAAI
MARLELGTRVESLETLVASLHGKLDVALDFL-QARARR
>551655682
IDKFRRDKIREIVFYLFFLVIFTTSALIQRNSHYYMQTVKDAILDFWDYMTTVIPQYLYKDKWYNVLHVNRLVQAPRIRQLRVNSQICVPPRLIKTCYFSTSAENRTSIPGTLVTYKTASQLATSEYRNYTGGGFAIDLALDF
FNYLKSVKWSDLHTRAIFFTFALYNPVNGLYLSVLLLFEFLPFGEVRPSSQFRVIRM-
GLSDYATVLLDAMVYLMVFSYILQDVRKIAAMYWTHFWHWFNWVNYTIFMVSLYFKISFYMASLRYVDFEAIGWTYYQMVNFQAFNNIFVWLRAFTYIKYLTDIANLSYTLTNAAKDCGLFLLIFAMVIFAYAQAFHIAF
GTDIGTYQTLLESIFGLFKTLLGDFNFNAIKSSNAYLGPLLFTIFEVVCYFIFLNMFLAILNKSYSDVIDKGVDPMTVENTLAANFRRFIDRLKSEDGRNQQE----DGDEDEEVKKIEK----------
ALGMMLNTIQVQPDINRNGQKVRS----------TSKLHD
>923138555
RLLYIRSWTFSTLRYFVFLTLFTIIVFGSRPAYYMSAQLRDILVEMFQYMEGPLLEGLFPTVDYAIYGYDQIVGLVRIRQVRIRKDTCVADKFVEHCYFKSSVVSKDPFGGNEWVWQADRQTLFGGVNGYPDDGFVVDLA
TQQISFLKQNKFVDLQTRLLAIEFSVYNANLNLFCVIRIAFELPEAGGVVPFAHFRTVKLLRYLDYFVLVCELLLCAMVFMYVVEEVGELVQAFVDDMWNYLDVLNLSLFVVVIIFRIIMTITLEQMFNLQGVALMATQEI
NINAINMLLCWIKTLKYLNHFDRIHQLSETLVRAAADTVLFMIMFLIIFFGYAQAGYMMFSADVEGFRTIFRSMESLVLSILGEFDYSQIERSNRYLAPLFFFSYMLLVFFVLLNVFIAIISSSFDSVREAAARDAEQRQLLSGK
VQRLSPWAKRAMLRSGKGDINLDGMISR--DEALTAEAFDQYDENEDNQLDADEIEAWLADLRYKKRVARMGSSGDPGLRPIDPNILRLTAAISDS
>116008215
TDEEVREALVEFSVFIIFLILTSLVVLSVRHMFYFNDTMKKLFTDWWDYLKYNFLVTLHMDFKYNFLYENLLLGPPRLRQIRVRKESCVNDAFFNTCYYSSGAEDRKPMHGSP--
FRTMHDLDSTPIWTYRTGGYTVNLNVKHINDLKDIHWLDRGSRLCLVEFNLFNENTDIFQSIKLIAEIPPTGGVIPQAHLQTVKMYSFFSMLMTVIYIFWYIMVIYYTIYEITEIRKSYFCSMLNILDCAILLGCYLALVYNIW
HSFKVMSLTSLDVLCFWNIIYVDMMAILAFLVWIKIFKFISFNKTLVQFTTTLKRCSKDLAGFSLMFGIVFLAYAQLGLLLFGTKHPDFRNFITSILTMIRMILGDFQYNLIEQANRVLGPIYFLTYILLVFFILLNMFLAIIMET
YNTVKGEITGRS--HSYIYRKLSGMLYWIGRKRRASETEDKD--AEHDV-------R----------KNMTPAEQQYFK----LNNRVGLLEEILEKLINNMDDILKRVYHNKKK-
>71999568
KIKLTARSFMEVGGYAVFLIVLVYVAFAQNSSYYYSKVMSDLFVNIWDWLSQVLIPGIYWTETS-
IYYENRLLGEPRIRMLKVTNDSCVMKSFIKECFYEEKLEDKTMVGVDAFIYATAKELENLKTVGYGGGGGFVQRLAQSAIATLKANRWIDRGSRAIIVDFALYNANINLFCVVKLLFELPASGGVTITPKLMTYDLLTYQG
TRMMIFEGIFCGFILYFIFEELFAIGRHYLTQFWNLVDVVLLGFSVATIILSVNRTKTGVNRVPFDDVTSSENSYLNIKACVVFVAWVKVFKFISVNKTMSQLSSTLTRSAKDIGGFAVMFAVFFFAFAQFGYLCFGTQIADY
SNLYNSAFALLRLILGDFNFSALESCNRFFGPAFFIAYVFFVSFILLNMFLAIINDSYVEVKAELAKKDGEGDWFMNK----VRGLRGKRP---AP----GEDATYDYKLMLYNEAFTRFNVTSMTEHVPEKV---
AEDIALNRRVDQMQESVFSIVDRIEGVNATLEKQRVQQ
>1253289448
KVSL-
KSALKDLISHFIFLVIVCIIAFSTCSTYYFTNVVNNLFTDVWEYMNNQLMAGLYWSPLTTIFYENRILGKPRIRMLKVTNKSCVVESFIKACYYEKGKEDRAPFSLDAFTYSTADELENSDLWGYGGGGYVQELSMDSIAFL
KANRWIDRGTRLIVIDFAVYNGNLNLFSVIKLMFELPATGGVIPKAQFNTLRLIRYQDYFIGGCEAAFCVLVIILVIKEVIDIWNTYLNKFWNVIDLVVIGLSITCIILSLRRTNLARSRIPMDDVVNAENMYNNIAALLLFFA
WIKVFKYISFNKTMNQLSATLSRSAKDIGGFAVMFFVFFFAYAQFGYLVFGTQIQDYSTFYDAVFALFRTILGDFDFHALERANRVLGPIFFITYVFFVFFVLLNMFLAIINDTYVEVKAELAQPDDFQSFVSKSWYRFLRAI
TKE-I---NM----ESNIDFAWKTALRNEKFTKFEVQSWKEMDNERR---D-ELEISRRVEGIETAIFAIASRIETVAGRLEWDKVRI
>847028821
KNLYANTTVRELVVYMLFLVLLFIITFSMRPMFYFTNFLNNLFLDFWKFLMGPLLNGLYWNVWYNIYFENKLLGVPRIRQIRVKNNSCIPNILVNQCYYSRNIELKESFGSSAWNYTEDSILQSGSFQGYGGGGGFYVDLSL
NIILSLFNNRWLDRNTRAVLVDFTIYNANINLFCIIKLVGEFPPVGGVLPSYEFRTVKLIRYVDFFVLACELIFVLYLLYYIEELIEIDKHYFCSLWNVLDVIVLSLGVICVSFHIYRMYMIQTMLNFEYLSFWQLQYDNLSAFL
LFLAITKIFKYISFNKMMTQLISTLNYCLKDLVGFSLMFFIVFLAFAQAGYLIFGSQLANYREFYICIFTLFRIILNDCNFKELLPLNPILGPLYFILYVFFVFFILFNMFLAILDDAYGDIKEELEKNENEFEDFLKEKSRKVFEKL
KREKIAMKTVDTNNDGVMDFEWKTELKESLFAKFDTDGDRVLDEQEQLKMLKELNLSRHIDDLENIIEILTQKFDIILDNFIKTKEKS
>675439102
RDLYIQTTLRELILYIIFLAILMIIAFGMTTMYWYTNLLSNLFLDFWNFATGALLNGLYWNTWYNIFYENKLIGLPRMRQVRMKNNTCIPLEFVNQCYYKMSDESTDSFGSTAWNYTSGSVLGSYPFQGYSGGGGFYIDLAL
GVLNFLYTNLWIDRGTRAVIIDFTVYNANINLFCVIKLVAEFPAVGGVIPTYDFRTIKLIRYVDYFILGCELAYLIFLIYYTIEELIEIKINYFHSVWNCLDVIVISISFVCVIFNIYRTIAVQTILNFDNLSYNQTTFNYLLAVTVF
LAWVKIFKYISFNKTMTQLSTTLGRCSKDLAGFAVMFFIIFFAFAQLGYLIFGTQCSDFRKFEISIYTLFRIILGDFDFNQLQTANAVLGPLYFIVYVFFVFFILFNMFIAIINDTYCEVKEELSQANEFEDFFKQRAALMMKKLK
RDKIALNIADANNDNKLDFEWKAELKEAVFARYDTDGDRVLDENEKKNMQSELELNVRVDHMESTVSSIVTKIDGVLQNLEKTKSKR
>828192580
REIILKSLLRELAIYLIFIILSIISFGMTSMFYFTKSLKDLFTDFFNYCDKVLIEGLYWEKYYNIFFENKLLGLPRLRQVRVRNDSCVHKFFVSKCFYSPESESKDDFGDIIWKYKSESTLEGRVFSGYGGGGYVSLLSSQILSNL
KHNLWLDRGTRAVFIDFTVYNANINLFCVITLVFEFPPTGGCVTSALYRPLKLLRYVDYFVMGCEVIFVIFILYYLIEELLEIKKHYFFDVWSIFDVLVLCIGIFCIVFNIYRTIIVNKLLNFDFLSDWQCRFNDFVAVGVFIAW
IKIFKYIKFNKTMSQLQSTLARCAKDIAVFAVMFMIVFLAYAQLGNLLFGSKLNDYNTIIDCIYTLFRIILGDFTFSDLKNTNPILGPIYFVTFVFFVFFVFLVNMFLAIINDSYAEVKSELAQEDEFQDYFKKGYNKLINKLKK
VRIALKSSDKKQD-NLSFEWRSELKEAVFAKYDMDRDFVLNSEERGKMMIDLELARRVDRIEHSVGTILSKVDRLLLKLDKLKVKR
>1067079595
REVYVRTTLRELVVYLLFLVILCIMTFGSASMYYYTSVLTNLFLDVWDFYENVMLDGLYWENWYDILYENRLLGLPRVRQVRVTNNSCVHADFIRQCYYDDNKVDKQPFGTSAWSYSSQQELNGFAHWGYSGGGSYK
NLSKAILADLRSNLWIERGTRAVFLDFTVYNANINLFCIIKLVFELPATGGVIPSWTFRTVKLLRYVDYFIMACEFIFCFFILYYLVEEAIEISVLYFKTVWNILDLLVILLSLGCIGYSVYSTMTISGKLNFEVLGYAAVQYNSV
VAVCVFFAWLKLFKYISFNKTMTQLSSTLSRCAKDLAGFAVMFFIVFFAYAQLGYLLFGSQVEDFSSFLKAIFTLLRVILGDFNFNEIEEANRVLGPLFYLSYVFFVFFVLMNMFLAIINDTYSEVKVEIAQRSEFEDYFKRGY
NNMMGKLKRDKLALKLSDGNSDGHLTFEIRHNLKEMFFAKYDLDGNRTLEVDEVKMMLKDLELQRRVDRLEHSIGSIVSKIDAVLVKLEKTKNKR
>926626589
KEWHVRTTLRELVIYIFFVVVLCILTFGMMSMYYHTKVISELFLDFWRFVDNVMLDSLYWENWYNVLYENRLLGSPRLRQLRVRNDSCVHDEFIFACYYAPDIEDKRPFGGTAWQYFSEKQLNGSSHWGYSGAGSYIDL
AKHIMNELRQNLWISRATRAVFLDFTVYNANINLFCIAKIVFEFPATGGMIPSWSFRTVKLLRYVDYFIMGCEILFSVFIMYYIEEAIEIKKNYFKSIWNVLDIVVIVLSLVCIAFSIYRILMVDTMLDFEFLGFWQTQFNNAV
ALAVFFAWIKIFKYISFNKTMTQLSSTLSRCSKDITGFAVMFFIVFVAYSQLGYLLFGSQVQDFSSFLKAIFTLLRVILGDFNFNEIEEANRVLGPLFYLSYVFFVFFVLMNMFLAIINDTYAEVKTELAEKNEFEEYFKKSYK
NLMSKLRRDKIAMKQVNGDQIRKFSFEIRQTLKEMIFSKYDLDGNRVLDEIEMKQMLLDLELSRRVDRMEHSIGSIVSKIDAVLVKMERAKAKR
>675370064
KEWYVKTTLRELFTYIVFLVLLSIMTFGMMSMYYQTKVMSELFLDFWKFAKDVLVENLYWENWYNILYENRLLGSPRIRQLRVRNDSCVHSDFITQCYYSPHIEDTEPFGGSAWVYHTEKEMNGSNHWAYSGAGSYRD
LALARLEILRQNLWISRATRAVFVDFTVYNANILNFCVIKLVFEFPATGGMIPSWSFRTVKLLRYVDYFIMACEVCIVLFILYYIVEESLEIKKNYFYDVWNILDLIVILVSVICIAFSAFRTVVVDNMLDFEFLGFWQMQFNN
AVAVDIFIAWIKVFKYISFNKTMTQLSSTLSRCSKDIGGFAVMFFIVFFAFAQLGYLLFGSIVKEFNTFGTSVFTLLRLILGDFDFEALENANRVLGPIYFLSYIFFVFFVLMNMFLAIINDTYAEVKAEIASNSEFEDFFKKGY
NNILGKLKRDKIALRVADFDGDKKLSYEVRQQLKEMLFAKYDVDGNRELDEKEMQQMFADLELTRRVDRMEHSIGSIVSKIDAVLVKMEKAKVKR

>241171514
REWYIRTTLRELLIYIVFLVILCVLTFGMMSMYYQTKVMSDLFLNFWSFVEDVMINSLYWENWYNILYENRLLGSPRIRQLRVRNDSCVHPDFITQCYYSPHFEDKGPFGGSAWTYHSEKELDGADHWGYSGAGYYADL
ALQVMAELKDNLWIGRATRAAFLDFTVYNANVNLFCVIKLVFEFPATGGMIPSWSFRTVKLLRYVDYFVLACEVVFVIFILYYIVEEALEIKSNYFKSIWNILDILVILISTVCIAFSIYRTVMVNQLLDFGKLGFYQTQFNNA
VALAVFFAWIKVFKYISFNKTMTQLSSTLSRCSKDVAGFSIMFFIVFFAFAQLGYLLFGTQIQGFSSFVNAVFTLLRLILGDFDFEEMEAANRVLGPIYFLSYVFFVFFVLLNMFLAIINDTYAEVKSEIAQKNEFEDYFKKGY
NNILGKLKRDQIALKMADSNNDSKLTFEVRRTLKEMIFSKYDVDGNRELSEQETKEMMAHLELSRRVDRMEHSIGSIVSKIDAVLVKMEKAKAHR
>1242857867
RKLRVKTTLKELVVYFLFLVLLSVVAFGMTSMYFVTKVISELFLDTWKVLKGSLPSALHWEHWYNVFFENKLLGLPRIRQVKVRNGSCVHEYFIFDCYYSKSNEDTESKVYTGFTYQSEASLDGSSSEGYSGGGFVQNLT
QRILEHLFTESWLTRGTRAVFVDFTVYNANVNLFCVVRLTVEFPPTGGAVSSWSFRTVKLLRYVEYFIMACEFFFALFIVYYVIEELLEIKIHYFKSVWNILDVTVIVISIVFLTFDIYRLLAVERQLDFDFISLWQIRFDNAIAI
TVFLAWVKLFKYVSFNKTMSQLSSTLARCAKDLAGFAVMYFIVFLAFTQLGHLLFGSKLKEFSDFTSSFFALFNIILGSFDFYALREVDRYLGPTFFMLFIFFIFFILVNMFLAIINDTYSEVKSELAQKDDLKDYVLTRYRKM
LGKIKEDRVVLKYADANADGVIDFEWRQDLIEALFCKYDADGSRYLDQEEQRRLRQDLELCRRVERMEHAVGGIIGKIDAVLAKLEKAKIRR
>908475875
QELLVQRTFRELVIYVLFLIFVCIIAFGMSSMYYYTKVMMDLFLDFWKFAKGPLLNSLYWEKWYNIYYENKLLGVPRLRQLRVSTNSCVNEEFITDCYYSKDDEDTSDFGETAWKYQSQDTLKGFMYWGYSGGGYIKD
LSNQIINQLIDKLWIQRGARAVFISFTVYNANINLFCVITLLVEFPATGGAIPSWTFHTVKLIRYVDMFILGCEVVFVIFVLYYIIEEIMEISFTYFLDFWNILDVVLLFVAIVCITFNVYRSLTIMSKLPFENIGFWEVRFNNAVA
ITVFIAWVKVFKYISFNKTMTQLSTTLSRCAKDLLGFFIMFFIVFLAFTQLGFLLFGTQVNDFSTFMNSFFTLFRIILGDFDFHQLEQANRILGPIFFMLFVFFVFFVLINMFLAIINDTYSEVKSDMAQTDQFIDYIRQGYQRI
WSKLKKDNKKILSGELS-GKAVHFDWKKELKEEVFAKYDIDGDRVLDPVEQAKMLEELQLTRRVDRMETTLKGMINNIDAVLVNLVQR----
>685230872
RELVIKTTLRELITYIVFVIIICIVTYGMTNTYYYTRAMRNLFVDFWKYAEGPLLDGLYWEKWYNLYYENKLLGRPRLRLLRVRSDSCIHSEFTTDCYYAKGIEDKSSFGETAWTYQTSSELKSRDYWGYDGGGFTHLLAA
RVIAELKDNMWLDRASRVVFLDFTVYNANINLFCVVKLIVEFPATGGAISSAVFRSVKLLRYVDYFVGFCEVVFCLFIIYYIVEEVLEIYLHYFKNVWNCLDILMILIAIVCMAFNIYRTVTVGGKLSFENLGFWQVQFNHM
SAVLIFFAWFKLFKYISFNKTMNQLSSTLSKCAKDVLGFCVMFFIVFFAFAQLGYLVFGAQLRDFSTFGDSVFTLFRIILGDFDFIALEDANRILGPIFFVTYVFFVFFVLLNMFLAIINDTYGEVKAEIR-
DDEMTDYFKKSYNKLLGKLRRAQVALKDADHNKDKNIDFEWRTEMKEEAFAAYDENGDRVLSEDEQRHLARDLE----------
>Chem014425
REFQVKTTLRELIIYVAFIVILCILTFGMTHMYYFTKVMKDLFVDFWTFADDGLMNGLYWEEYYNIYFENKLLGRPRLRQVRVRNDSCVPDDFIKECYYSESVEDKRAFRGSAWEYSTEDELNGRGFRGYGGGGGFTKLL
SKDIIAELKANLWLDRGTRAVFIDFTVYNANINLFCIVTLLLEYPATGGCIPYYKFRTLKLLRYVDQFVFACECIFVLFLYYAIEEVIEIHKHYFKDFWNIVDILVVLMGIVCVAFNVYRTMEVGNLLNFDTLGYWQETFN
DFIAVAVFFAWIKVFKYISFNKTMTQLQSTLARCAKDIAGFAVMFFIIFLAYAQWGYLIFGTQVRDFSTYLNSVFTLFRIILGDFNFREIEDANRFLGPLFFISYVFFVFFVLINMFLAIINDTYAEVKSDLAQEDDFADYFKK
GYDKMMNKLKREKIALQTADTNQDNKLDFEWRTELKEAVFAKYDVDGDRILNEDEQRKMQEDLELSRRVDRVEHSIGSIVSKIDAVLLKLDKAKIKR
>156401013
REMHIKTTLRELIIYLVFITILCILTFGMTSMYYFTKVMRDLFVEFWMYTTETLSDGLYWEQYYNIYFENKILGRPRFRQLRVLNDSCVHDDFIKECYYSPNVEDMNTFGGSAWTYQTEKDLQGRSYWGYSGGGGFTQLLT
QQIIEDLKQNLWLDRGTRAVFIDFTVYNANINLFCIVRLVFEYPPTGGCIPSFNFRTVKLIRYVDHFVMACEGIFILFIIYYTIEEGLEIKKHYFKSFWNVLDVVVIMLGYVAVVFNLYRTMAVGDLLNFDSLGFWQTQFN
NMVAVAVFFAWIKVFKYISFNKTMTQLSSTLSKCAKDVAGFAIMFFIIFFAYAQLGYLIFGTQVRDFSTFEDSIFTLFRIILGDFDFHQIENANRVLGPLFFMTYVFFVFFVLLNMFLAIINDTYAEVKSDIAQESEFEDYFKK
GYQNMLNKLKREKIALKSADVNQDNKLDFEWRANLKEAVFAKYDLDGDRILDEDEQRRMQEDLELSRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>260802732
KEQYVKTTLRELFTYIIFIVILCILTFGMMGMFYFTKVMEDLFLDWWKFTRGPLLDGLYWTKWYNIYYENKMLGVPRIRQIKVRNDSCVHEDFILECFYSEAVEETEPFGWDAWFHKDSDEIDGSGHRGYAGNGYVQD
LSEAKILELFNNLWTDRGTRVVFIDFTVYNANINLFCVIRLTAEFPPTGGLIPSFNFRTVKLLRYVDFFIMACEGLYVLFHIYYIIEEALEIKRHYFKSFSNILDITVIMLSLAAVAFNVYRTLKVEELLDFEFLAWAQGIYNDM
VAITVFFAWIKVFKYISFNKTMTQLSSTLARCAKDIAGFLVMFFIIFFAYAQLGYLVFGTQVKDFSTFPDAIFTLFRIILGDFDFNALEEANRVLGPMFFITYILLVFFVLLNMFLAIINDTYSEVKEEIAQKSEFEDYFKKGYN
KMLSKLKRDKIALAAADVNRDKQVDFEWRQDLKEAVFAKYDLDGDRVLDADEQIKMQADLELSRRVDRMEHSIGSIVSKIDAVLVKLERAKLKR
>198420844
RELYIRTTIRELVIYVFFIATLCILTFGMTSHYYYTKVMQELFVDFWRFSEGPLMDGLYWETWYNIYYENKLLGVPRIRQLKVRNGSCVHPDFITACYYAKAKEDTSPFGGTAWVYHTEDELDGSSHWGYSGGGGYYQDL
SLAIIKNLKENLWLDRGTRVVLVDFTVYNANINLFCVVRLVVEFPATGGAVPSYQFRTVKLLRYVDYFLMGCEVLYMLFILYYIVEEVLEIRKNYFKSVWNCLDVLVILLSIVAFSFNIYRTITVDSLLDFEFLGFWQMQYN
NMVAIVVFFAWIKIFKYISFNRTMTQLQSTLSRCSKDIAGFAVMFFIIFLAYAQLGYLVFGTQVKDFSTFQESIFTQFRIILGDFDFHALEAANRVLGPIFFLTYVFFVFFVLLNMFLAIINDTYSEVKAEISQKSEFEDYIKRG
YNKVLEKLKKDRIAIQTADLNQDRHLDWEWRLDLKEAVFAKYDQDGDMILNENEQKKLKADLELARRVDRMEHSIGSIVSKIDAVIVKLERAKLKR
>47551289
KELHVKTTLRELAIYLVFLVVICIVTFGMTSMYYYTKVMSELFLDFWRFAETPLMDGLYWEKWYNIFYENKLLGVPRIRQLKVTNDSCVHTDFINTCYYSESLEDTATFGGTAWVYSTEEDLDGSGHTGYGGGGGFYAD
LSLDIILTLKENLWLDRGTRALFVDFTVYNANINYFCIIRLTVEYPPTGGAIPSYKFRTVKLLRYVDYFILACEGIFCLYIMYYMGEEILEIKRHYFKSSWNCLDVIIIMISVGCVAFSVYRYLTVANLIDFEYMGYWQDLFNNI
IAVNVFLAWIKTFKYISFNKTMTQLSSTLSRCAKDVGGFTVMFAIIFMAYAQLGYLVFGTQIKDFSSIGSCIYTLFRIVLGDFNFHELEAANRFMGPIFFLTYVFVVFFVLLNMFLAIINDTYSEVKSDIAQKSEFEDYFKKGY
EKMVGKIKRDKIALAHADANADQHLDFEWRQELKEAVFAKYDVDGDRVLDAEEQMKMAHDLELSRRVDRMEHSIGSIVSKIDAVLVKLERAKLKR
>291234667
RELHVKTTLRELIIYILFIIVLCIVTFGMTSMYYYTKVMSELFLDWWRFAEGPLVNGLYWEQWYNIYYENKLLGVPRIRELKVRNNSCVHSDFIKYCYYSESVEERTPFGGTAWEYSTEEELDGSSHWGYGGGGYYLDLSL
EEILELKEYLWLDRGTRVVFIDFTVYNANINLFCVIRLVTEFPATGGAIPSWDFRTVRLLRYVDYFIMACEGIFCLFIAYYIVEEILEIKMNYFKSVWNCLDLLVIGLSLVCIAFSVYRTMEVAEMLDFEYLSYWQSMYNNML
AVLVFFAWIKVDNLFYA----------KF---------D-----
IFTGFTLFRIILGDFDFHELEQANRIMGPIFFCTYVFFVFFVLLNMFLAIINDTYSEVKADIAQKSEFEDYFKRGYDKMLGKLKRDKIALAQADINNDQQLDYEWRHELKEAVFAKYDVDGDRVLDSDEQARMAADLQ
LSRRVDRMEHSIGSIVSKIDAVXVKLERAKLKR
>675875990
SNIKIKTTLRELVVYILFLIVLCVMTFGMTSMFYYNQAINSLFIDLWTFMQNNLLNSLYPDSWYNVMHENKLLGVPRIRQVKVRNNSCIPDDFIRECFYNSDNEFKHSFGGTAWQYRDSSEMKSSSYSGYGGGGFAQDLS
QALVQNLFDNLWIDRGTRAVIIDFTLYNGNINLFCVVRLVAELPATGGLFPAYTLQTSKLLRYVDYFVLACEIFFILFLIYYTVEEIIEIRMHYFKSFWNILDVLNILVGVCCCIFDVYRTISVNNELNFVFLSYCQDMANSAL
AIMVFFSWIKIFKYISFNKTMTQLSSTLSKCSKDLGGFAVMFFIFFLAFAQLGYLLFGAQVPDYNSLGTCIFTLFRFILGDFDFASLQQANRVLGPIYFVCFVFFMFFILLNMFLAIINDTYSEVKSELAAKNEFEDYFKQGAL
KMMDKLKKEKLAVKSGDLNNDGVVDFEWRHNMKEAVFARYDSDGNRNLDGAEMDNMAADLE----------
>675861632
QDLRIKTTLRELVIYCVFLAILCITTFGMTSMYYYTNVMNSLFIDVCIYAEKVLLPGLYWETWYNIYYENKLQGVPRLRQVRVRNDSCAPDNMIKDCYYSSSAEDTRPYGGTAWTHSN----
LGSNVRAYGGGGYYADLSEQVIDNLKTNLWVDRGTRAVFIDFTIYNANINLFCVIKLIAEFPPTGGIIPSYDFRTVKLIRYNDFFVFACEIIFCIFILYYIGEEFLEIKKYYLKEFWNVLDIVVICISLCCIAFNIYRFLEVKKLL
DFDYISYWQSVFNSALAIMVFFAWIKIFKYISFNKTMTQLSSTLGACAKDLGGFSIMFFIVFLAFAQLGYLLFGAEVSDFNAYETAIFTLFRIILGDFNFQALQQANSILGPTYFILYVFFVFFVLLNMFLAIINDTYSEVKAEL
AQPNDIDGYFKQRANNVLTKLRKNKIAITTADANGDRMIDFEWRHDLKEAVFAKYDINGDRLLDESELYKMQADLE----------
>Bner206993
SETKAYVLLRELLIYGIFLTILLVLTFGMTNMYYYTKVMEELFLDFWRFSEGPLINGLYWETYYNIYFENKMLGVPRLRQIKVSNSSCIHSDFIKQCYYTASVEDTNPFGGSAWTYHKADDLESGNVKGYSGAGYYQNLSL
AIIQDLKQNLWIDRGTRVVFFDFTVYNANINLFCIVRLLVEFPATGGVIPSWTFRTVKLIRYVDYIVLGCEGIFVLFLYYIVEESLEIKKHYLYDIWNWLDLLILGISLICVAFSAYRYWEVSDKLDFEFLSYWSVQFNYIIAG
AIFLAWIKVFKYVSFNKTMSQLTTTLGKAKKDIVGFLVMFMIIFFAFAQLGYLLFGTQVRDFSNLINAMLTLFRIILGDFDFHELESAHRVLGPIYFIFFVFFVFFVLLNMFLAIINDTYAEVKAELAAKNEIEEFFKQQYNK
MMSKMKKDKIALRSADKDNNKKIDMEFRTELKEAAFAKYDLNGDRVLDEEEQRKMQDDLELSRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>1236532622
AELKIKTTLRELVMYCLFLVVICILTFGMTSMYYYTFVMNSLFLDWWGFAQGPLISGLYWDTWYNIYYENKLLGVPRMRQVRVRNDSCIATDFIRACYYSRSAESREPFGSTAWNYSTEQQLDGRGYSGYNGGGYYIDLS
TAKIDELFHGLWIDRATRAVFIDFTVYNANINLFCVIKLIAEFPATGGGFTSWEFRTVKLLRYVDYFVLACECIFMLFICYYIVEEIIEIKRHYFSSFWNWLDIVVIVIALVCAAFNIYRTISVGNKLDFDFLSFWQEQFNQAIA
ITVFFAWVKIFKYISFNKTMTQLSSTLGACAKDLLGFAVMFFIVFFAFAQLGYLIFGTKVEDFREFQQAIFTLFRIILGDFDFASLQNAHRVLGPIYFILYVFFVFFVLINMFLAIINDTYSEVKSNLAQPNDFEDYFKQRAGR
VLDKLRRDKIAMQSADMNNDRQLDFEWRTELKEALFARYDMDGDRVLDEEEQRKMQWELELSRRVDRMEHSIGGIVNRIDSVLLKLERAKLKR
>918287730
RELHVKTTLRELIIYIIFLAVLCVITFGMTSMYYYTKAMTDLFIDVWKFAKGPLMDSLYWEKWYNIYYENKLLGLPRMRQLRVRNDSCVHSDFIKTCYYSENNEDHEGFGANAFNYRSKDDLDGTSHWGYGTGGFVQD
LSMKLVNALFDNLWIDRGTRVLFIDFTVYNANINLFCVMRLVVEFPATGGALSSWRIASVKLIRYVDYFVMACEAIFVLFILYYIEETIEIKKHYFKCFWNILDMLVIFISVLCIAFNVYRTSAVGDKLDFQFLAFCQTHFNN
AIAVTVFMAWIKIFKYISFNKTMTQLSSTLGKCAKDLAGFAVMFFIIFLAFTQLGYLIFGTQVKEFSSFESAFFTLFRIILGDFDFHQLESANRVLGPIFFILYVFFVFFILLNMFLAIINDTYSEVKGDIAQKSEFEDYFKRGYQ
RMVDKLKRDKIALQSADMNNDKQLDFEWRQDMKELMFAKYDLNGDRVLDEDEQKKMNADLKLSRRVDRMEHSIGSIVSKIDAVLVKLEKAKIKR
>405953847
RELHVKTTLRELIIYLVFLIILCVVTFGMTSMYYYTKVMSDLFLDFWKFARGPLIKGLHWETWYNIYYENKLLGLPRLRQLRVSSNSCVHENFIRECYYSANLEDKTTTAYTAFAYKTEEELDGSSHEAYSGAGYVQNLSL
EVTIYLFNKLWITRGTRVVLFVDFTVYNANINLFCAIRLIVEFPATGGAIPSWDFRTVKLIRYVDYFIMACELIFALFIAYYIVEEALEIKKHYFKSLWNILDLTVIIIALLCIAFDIYRTVEVDNKLDFEFLSYWQTRFDNAIAIA
VFFAWVKIFKYISFNKTMTQLSSTLGRCAKDLLGFAVMFFIIFLAFTQLGYLLFGTQVKDFSNFGNSFFTLFRIILGDFDFHQLEAANRYLGPIFFMLFVFFVFFVLINMFLAIINDTYSEVKGDMAQMNEFEDYFKRGYQK
MVDKLKRDKITLKHADINGDDKLDFEWRQDLKEALFAKYDIDGDRVLDKDEQIRMQEDLNLSRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>871233656
RELYVKTTLRELIIYIVFLVVLCVITFGMTSMYYYTKVMSDLFLDFWRYAKGPLVNGLYWETWYNIYYENKLLGVPRLRQLKVHNDSCVHEDFISECYYAESIEDTKSFGDTSWKYFTEDQLHGSSHWGYSGAGYYKDL
SVEIVNHILFDELWIRRGTRAVFVDFTVYNANINLFCVVRLLVEFPATGGAIPSWTFRTVKLIRYVDYFILACEIIFCLFILYYMVEEALEIKKHYFKSFWNILDILVIIISIVCVAFNVYRTLEVDAKLDFERLSYWETRFSNAIAI

AVFMAWIKIFKYISFNKTMTQLSSTLGRCAKDLAGFFVMFLIIFLAFTQLGYLLFGTQVKDFSSFENSFFTLFRIILGDFDFHQLEQANRILGPIFFMLFVFFVFFVLINMFLAIINDTYSEVKSDMATKNEFEDYFKKGYMK
MLDKLKRDKIALQSADINQDKQLDFEWRQDLKEAMFAKYDIDGDRVLDEEEQKRMQADLALSRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>919076678
KDLYIKTTLKELIIYLLFLVVLCVITFGMTSMYYYTKVMSELFLDYWQFARGPLMDGLYWETWYNIFYENKLLGVPRIRQLKVKNGSCVESNFIDDCYYASGDEDKSTFGGTAWSYQTSEQLDGSGHYGYEGGGYVQK
LSLAIVDDLKQNLWLTRGTRAVFTDFTVYNANINLFCVIRLLVEYPATGGALSSWSFRTIKLJRYVDYFVMACECIFILFIIYYIIEEVLEIKKHYFKHFWNVVDILVLVISSVCIAFSIYRTIAVDTMLDFEFLSFWQMQFNSAV
AVAVFFAWVKVFKYISFNKTMTQLSSTLSRCAGDLAGFLVMFCIMFFAFAQLGVLVFGAVIRDFSNFWNAVFTLFRIILGDFDFNSLQDFDKFLGPLYFFCYVFFVFLILLNMFLAIINDTYADVKSDIAQKNDFEDYFKK
GYQKMMDKLKRDKIALASADMDNDMVVDFEWRTELKEALFAKYDVDGDRTLDESEQKQMAVELELQRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>PdumPKD22
RELHIKTTLRELVIYIVFLIVLCIITFGMTSTYYYTKVMSELFLDFWRFSESSLVNGLYWETWYNIFYENKVLGVPRIRQLKVRNDSCIHDDFITACFYDKGIEDRNPFGGTAWEYHSESELGGSSQRGYGGGGYIQDLSLNI
ILDLRENLWLDRGTRAVFVDFTVYNANINLFCVVKLMVEFPATGGAIPSWNFRTVKLJRYVDFFIMACECIFCLFILYYTIEEVIEIKKHYLKSFWNLLDILVIGISLVCIGFSIYRTIAVRDTLDFEFLSFWQVQFNSALAVTV
FFAWVKVFKYISFNKTMTQLSRTLGSCAKDLAGFAVMFFIIFLAFAQLGYLIFGTEVKDFSTFGNAVFTLLRIILGDFNFHELEEAERILGPVYFILYVFFVFFVLLNMFLAIINDTYSAVKEDLAQKSEFEIQQNEGETQLQ-
---ERQDAIQAADINQDRMLDFEWRRDMKEAVFAKYDQDADRVLDENEQSKINADLELSRRVDRMEHSIGSIVSKIDAVLVKLERAKSKR
>PdumPKD21
REMHVKTTLRELVVYLVFLVILCVVTFGMTSMYYYTQVMSQLFLDFWTFSEHALCDGLYWETWYNIFYENKLLGVPRVRQLKVRNDSCVHDDFIKACYYAQSIEDKAPFGGTAWTYKTEKELDGSGHWGYGGGGFY
EDLSAAIIADLKQNLWLDRGTRAVFIDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVDYFVMACEGIFVLFILYYMVEESIEIKKHYFKSFWNILDILVIISIVCIAFNIYRTLAVGKKLDFEFLSYWQVQFN
SAIAITVFMAWVKVFKYISFNKTMTQLSSTLAACAKDLAGFAVMFFIIFLAFAQLGYLIFGTQVKDFSSFDDAIFTLFRIILGDFNFQELEEANRILGPIYFLLYVFFVFFVLLNMFLAIINDTYSEVKAELAQKSEFEDYFKKG
YDRMLDKLKRDKIAINSADHNQDKQLDFEWRHDLKEAVFAKYDLDGDRVLDEGEQMRMNRELELVRRVDRMEHSIGSIVSKIDAVLVKLERAKAKR
>919076662
RELYVKTTLRELVIYLLFLTDLTILTFGMNSYYYYTKVMSELFLDFWRFAEGPLMDGLYWEKWYNVFHENKLLGVPRLRQLKVRNDSCVHDDFIKTCYYAESIEDKFPFGGTAWVYQTEDELDGWGHEGYGGGGFV
QNLTEPIADLKQNLWLDRGTRVVFVDFTVYNANINLFCVIRLVVEFPATGGAIPQWVFRTVKLIRYVDYFVLACEGIFLLFIIYYIIEEALEIAKHYFKSVWNILDIVVIVISVVCTVFNVYRTIAVGNMLDFEFLSWAQVLFN
QGIAFAVFFAWVKIFKYISFNKTMTQLSSTIANCAKDLAGFGVMFFIILFAFAQLGYLIFGTQVRDFRSFPDTILTLFRIILGDFDFHELEAAHRVLGPLYFLLYVFFVFFVLLNMFLAIINDTYSEVKADLEQENEFEDYFKK
GYDKMLDKLKRDKIAIVDADANADKQIDFEWRKDLKEAVFAKYDLNNDRMLDEEEIRKMQEELELSRRVDRMEHSIGSIVSKIDAVLVKLEKAKLKR
>270002831
-
HDVLFTTFREAVLYIIFVIAVTICTVGRRSMFYITQALKGQFLDFWHYTQSLMLENFYWEHYYNILFENKLLGVPRIRQVKVKNDSCIHEYFFTSCYYGPTDEDRAPFGGTAWTYNTAEKTESIPFSGYGGGGGFYLDLTAK
LIRDLKENLWITRGTRAIFIDFSIYNANLNLFCVCKLIFEFPPTGGHPSHSFHAVHLVHYVDYVTFVFECTVYAIAGFFLAEEIREIIYFYFLQFWTYIDVVIIGMAFANLITSIVVFPNVDDAINLEYLAEMRIFYNNFAASLLF
FSYIKLFKYLNFNKTMGQLNNTLKRCAFDILGFSIMFFIIFFAFALLGYLLFGSQVEDFSSFGVAMFTLLRTILGDFDYQAIEKANRVLAPIYFLAYIFFVFFVLLNMFLAIINDTYADVKTEIAAPDEMQEFLKKGFYKMLQ
KCNIKYF--QKA----EFNATIQIRDALKEMFFARYNIDPLAEVGDYDIKKIMKELEQQERLDQIEKTIQMLATKIDTLIKKLENVRKAK
>1174757226
KELYVRTTLRELFVYILFLSTTLVMTFGATNMYYLNNALQNVFVGLFKYLKGTMLDALFVEDWYNIYYTNKILGLPRIRQLRVKNGSCVPVEFIFDCFYSIGQEATEPFGTSPWIYQTATDTGMASYWGYSGGGGYILDLSL
KLIEKLEARRWLDQATRAVIVDLTVYNANLNYFATVRLCVEMPPTGVLVPTANIRVIKLIKLLDYAVLAFEVIFCLFVAYYIGEELIEISKLYFRSFWNILDIIVLTCSTTCIICAICLFMQINENLDFESLSFYQITLNNNALAITV
FCAWIKVFKYLSFNKTMMQLSSTLGRAAKDVASFAIMFFIVFFAFAQLGFFLFGTQLEAYNDVGTASFTLMRIILGDFDFDFALYAANRLFGPLFFFMYIFFVFFVLLNMFLAIINDTYSAVKEDLAKRNDFEDYIMRGFKRL
VRAFKRFERKILKN----KALVELRWEDDLKAEVFGKVHTEAAGDSA-----------LEGRVRKLETALSAILTNFDLALAKLERVKLTR
>1160116546
RNVNPISTLREFLLYLVFLLDLSILTFGMVNMFFFTKVMSVLFLNFWEYVEGPLLDSLYWKEWYNIYYENILLGHPQIRQLRVRNNSCIHEYFFKDCYYTFATEDQSPFGKLEWEYRVPNNEFELWHWGYGNGGYIFELT
AGEIKQLKSKLWLDSGSRAVFIDFTVYNANVNLFTAIRLLVEFPATGGAIPSSEFFTVKLLRYVDYFLAACEIIFCFICNYLIQEIIELHSYYFKNFWNCLDFLIVLFSLLAIFYNVFRTIKVSLLLNLYFLAYWQNRYNVIIAFT
VLIAWIKIFKYINFNMTMTQLSTTLSRCAKGIVGFAIMFFIVFFGYAQLGYLIFGTQVESYHTFGNCIFTQFRIILGDFNFHEIEDANQILGPVYFITFVFCLFFVLLNVFLAIINDTYSEVKDDKTKEVDFDDVFRKGYHSFL
ARLKKTSSAVYKG----DKETQK--EDDFENIYY----TDSAKE-------------NIRRTGAIGYGS----------
>5669804
KEVEITTTLQELLLYFIFLINLCILTFGMVNMYYLNKVMSSLFLDFWKFMEGPLLEGLYWDSWYNIYYENILLGVPRVRQLKVRNNTCVYSSFMSECYYTSANEDLSNFGNTEWRYSTSNTN-
SPWHWGYRNGGYIFTLTKNKFIDLRLNSWTTRGTRVFIDFSLYNANVNLFCIIRLVAEFPATGGILTSWQFYSVKLLRYVDYFIASCEITFCIFLFVFTTQEVKKIKEFYFKSIWNWLELLLLLLCFVAVSFNTYYYNVQIFLLL
DFYFLACWHIYYNNIIAITIFFAWIKIFKFISFNKTMSQLSSTLSRCVKDIVGFAIMFFIIFFAYAQLGFLVFGSQVDDFSTFQNSIFAQFRIVLGDFNFAGIQQANPILGPIYFITFIFFVFFVLLNMFLAIINDTYSEVKADYSRR
PDFEKMIKQSYKNVLEKFKKAQKT--KG----SGDLAE-RRE----------IQ------NAEQMKKWKERLEGTTTKYKMRFSLSA----------
>1033394660
RELEMKTSLQELLIYIIFLTDLCILTFGMVSMYYLNKVMSNLFLDYWEFLEGPLLDGLFWNTWYNIYYENLLLGVPQIRQLKIRNNTCIYPSFMKVCYYRYLYEDRESFGDTEWFYSSSQTL-
TPWHWGYGSGGFMFTLSLEKLIYLKENSWTTRGTRAVFIDFSMYNANVNLFCIVRXXVEFPATGGAIPSWQFHSVKLLRYVDYFLASCEVIFCLFIVAFIIQEITKMKKLYFRDAWNWLDMLLILLSVFAIAFNVYRTVEVS
ELMDFYFLGFWQTRYNNMISVNVFFAWIKIFKFISFNKTMTQLSSTLSRCAKDIIGFAIMFFIIFFAYAQLGYLVFGSQVDEFSTFQNCIFTQFRIVLGDFNFASIEHANRILGPIYFITFVFFVFFVLLNMFLAIINDTYSEVKA
DFSPSREFEDLIRQSYNKALVKLKRQTSYSIKR----KGQSTSDVTK---SSASG--VPDPVSHKEFQKLSKYIADLECGDTGNHSSKEIQE----------
>556956196
KESKLKTTLQELLIYLIFLIDLCILAFGMVSMYYLNQAMSNLFLDFWKFAEGPLLEGLYWNKWYDIYYENILLGVPRIRQLKVQSGNCVHGYFIRECYYSSRYEDRSSFGGTAWTYSFPSTY-
TAPHWGYSSGGYYLDLSAEKIAFLKSNHWFDHRTRAVFIDFTTYNANVNLFCVIRLTVEFPPTGGAVPSWHLYSVKLLRYFDYFVASCEIIFCLFTFAFLGQEFIEIWESYFKNTWNCVDVLLGLLSLVAIAFNIFRTTRVSM
VLDFYFLAYWQTQYNNMIAVIVFIAWIKIFKFINFNKTMTQLASTLSRCAKDIIGFAIMFFIIFFSYAQLGYLVFGTQVDDFSTFENCIFTQFRIILGDFDFATIEMANRVLGPIYFITFVFFVFFILLNMFLAIINETYAEVKAD
LSQNPEFEDLIRKSYNKALVKLKKAAVCVRKA----SGEMNLKLQQDMEKAIFNAYDIDKNEELHKTEGQKIDQKSQLVTQVSQLKQQINDMMQKMDEKTPKSKSGKKAY
>1160100767
RERYLKSMLREMVTYSIFLVILCVLTYGMVSMYYYTKVMASLFLDFWKFAEGPLLDGLYWDIWYNIYYENLLLGVPRIRQLKVKNESCIHEDLIKECYYSVANEDTSSFGKTAWVYTSEEDLKETSYWGYSGAGYYLDLT
ATKLKVLKENLWLDRGTRVVFITFSVYNANINLFCVIRLVVEFPATGGALSSWQFQTVKLIRYVDYFLAACELAFCIFVFYYIIEEILEIYIHYFKSVWNCLDVAIIVLSLIAICNIYRTSIVNEMLNLESLAYWQIQFNNMAA
VVVFFVWIKIFKFISINKTMTQLSSTMSRCAKDIIGFAIMFFIVFLAYAQLAYLVFGTQVDDFNTFHTCIFTQFRIILGDFDFAKIEEADRILGPVYFVTFVFFIFFILLNMFLAIINDTYSEVKSDMAQKNEMEDLIKKVLIFFF
F----------
>50539686
REMYLKTVLREMITYILFLLTLCIITYGMVSMYYYTKVMSQLFLDFWKFTEGPFLNGMYWELWYNIYYENLLLGVPRLRQLRVRNESCVHEDLVYDCYYSPANEDKAPFGGTAWRFKDESSLGESSYWGYGGGGYYQD
LSANQLQELKNNLWLDRGTRAVFLDFSIYNGNVNLFCIVRLLVEFPATGGAVPSWQFQTVRLLRYVDYFVGMCEVSFCLFVLYYLVEEALEIRILHYFKSLWNCLDVLIVALSVPAIIMNICRTSAVSHRLNFEPLARLQVHF
NNLAAIIVFLSWVKLFKFINFNKTMNQLSTTMSRCAKDLMGFAIMFFIVFLAYAQLAYLVFGTQVNDFSTFQACIFTQFRIILGDFDFSEIEEADSVLGPIYFTTFVFFIFMILLNMFLAIINDTYSEVKADMAQRSEMEDLIK
KSYNRAMVKLKKSSISLQQA----AGKLSFELRQDLREAIFAKYDLDGDQELTEHEHQQMRDDLELVRRVDRMEHSIGSIVSKIDAVIVKLERAKMKR
>4505835
REKYLKSVLRELVTYLLFLIVLCILTYGMMSVYYYTRMMSQLFLDFWKFTEGSLLDGLYWKMQPSIFYENLLLGVPRIRQLRVRNGSCIPQDLIKECYYSVSSEDRAPFGGTAWIYTSEKDLNGSSHWGYSGAGYYLDLTA
AQVASLKKNVWLDRGTRATFIDFSVYNANINLFCVVRLLVEFPATGGVIPSWQFQPLKLIRYVDFFLAACEIIFCFFIFYYVVEEILEIRIHYFRSFWNCLDVVIVVLSVVAIGINIYRTSNVEVLLNFEHLAYWQIQFNNIAAV
TVFFVWIKLFKFINFNRTMSQLSTTMSRCAKDLFGFAIMFFIIFLAYAQLAYLVFGTQVDDFSTFQECIFTQFRIILGDINFAEIEEANRVLGPIYFTTFVFFMFFILLNMFLAIINDTYSEVKSDLAQKAEMEDLIRKGYHKAL
VKLKKNTVSLRQG----GGKLNFELRQDLKEAIFTKYDQDGDQELTEHEHQQMRDDLELVRRVDRMEHSIGSIVSKIDAVIVKLERAKLKR
>557013709
REKYLKNVLREMITYIIFLVILCLLTYGMVSMYYYTKVMSRLFLDFWKFTEGPYLNGLYWDLWYNIYCENLLLGVPRIRQVKVRNGTCIHEDLIKDCYYSVGNEDTSQFGGTAWTYTKESDLNGSTHWGYSGGGGYYLD
LSAAQIKTLKDRLWLDRGTRAVFFDFSVYNANINLFCIVRLLVEFPATGGMVTSWSFQTVKLIRYVDYFLAACEVAFCFFVLYYIVEEVLELRIQYFRSFWNCLDILIIVLSVVAIGINIYRTASVDASLNFEPLAYWQIQFNN
MAAVIVFFAWVKLFKFVNFNKTMSQLSSTMSRCAKDVIGFAIMFFIIFLAYAQLAYLVFGTQVDDFSTFQDCIFTQFRIILGDFSFTEIEEANRVLGPIYFTTFVFFIFFILLNMFLAIINDTYSEVKADMAQKTEMEDLIKKG
CNKAMVQLKKTTVSLRQA----GGKLNFELRADLKEAIFAKYDQDGDQELTEREHQQMRDDLELVRRVDRMEHSIGSIVSKIDAVIVKLERAKLKR
>528498254
PELNVKANLKELLIYIVFLVDLCLLTYGMTSTYYYTKAMTDLFVDVWTYMEGPLLDNLYWTKWYNIYYENLLLGVPRVRQLKVKQNSCVPTDFISDCFYNIKKEDDGAFGGTAWQYHSEKEIKGSSHWGYSGAGFYQ
DLSAALLEDLRENLWLDRGTRVLFIDFTAYNANINLFCVIRLVVEFPATGGAIPSYQIRTVKLIRYVDYFIIGCEIVFCVFILYYIVEEILELFRIHHFTSIWNILDIVVILLAVVAIVFSAFRTIKVDGLLDFEFLAFWQTQYNNM
NAVNLFFAWIKIFKYISFNKTMTQLTSTLARCALDIFGFAIMFFIVFFAYAQLGYLLFGTEVETFSTFNKCIFTQFRIILGDFDYDAIDRANRVLGPIYFFSYVFFVFFVLLNMFLAIINDTYSEVKSELAQKDEFQDLIKQSYA
KTFMKLKKEKIALDSG----ASELEFDFRNALKSAAFSKFDQDQDGNQTLDKQEQENMKQELELVHKVSQLENSVSAIMSKLDFVMEKMTFNTESR
>359465612
RELYIKTTLRELLVYIVFLVDICLLTYGMTSAYYYTKVMSELFLDFWDFAQGPLLDSLYWTKWYNIYYENMLLGVPRLRQLKVRNDSCVHEDFILSCYYSPDKEEQLPFGGTAWTYHSQDELGGFSHWGYSGGGYYLDL
SAEALRALQEGLWLDRGTRVVFIDFSVYNANINLFCVIRLVVEFPATGGAIPSWQIRTVKLIRYVDFFIVGCEVIFCVFIFYYVVEEILELHIHYLSSIWNILDIVVILLSIVAVGFHIFRTLEVNRLMDFEFLAFWQTQYNNM
NAVNLFFAWIKIFKYISFNKTMTQLSSTLARCAKDILGFAVMFFIVFFAYAQLGYLLFGTQVENFSTFIKCIFTQFRIILGDFDYNAIDNANRILGPAYFVTYVFFVFFVLLNMFLAIINDTYSEVKEELAQKDEIQDLLKQG
YNKTLLRLRKERVVLQGG----EQEIQFDFTNTLRTATFTKFDRDGNRILDEKEQEKMRQDLELTRRVLQLETVLEGVVSQIDAVGSKLERKGWLA
>1160139095

PELHIRRTLRELLVYVVFLVDICLLTYGMTSAYYYTKVMSNLFLDFWSYAQGPLLNGLYWTKWYNIYYENILLGVPRIRQIKVTNNSCVHKDFINGCYYTADTERKSSIRETAWTYYSESELGGSAHWGYGGGGYYQDLSAAILQNLKDNLWLDRGTRAVFIDFSVYNANINLFCVIRLVVEFPATGGAIPSSQFRTVKLIRYVDYFIIGCEVIFCVFIVYYIVEEILEIRIHYIKCIWNCLDLVVILLSLLSIGFHIFRTIKVNILMDFEFLAFWQTQYNNMNAVNLFFAWIKIFKYISFNKTMTQLSSTLARCAKDILGFAIMFFIVFFAYAQLGYLLFGTQVANFSTFLKCIFTQFRIILGDFNFDEIDNANRILGPIYFVTYVFFVFFVLLNMFLAIINDTYSEVKADLSQKSELQDIIKQGYNKALMKMKKERITLQNG----SKKLDYDFRSNLKNDTFSRFDSDDNKILDEEEQTRMQNELELGHRVVQLERSVSNIVLKLDAISVKVEKNKDMV

>942132194

RELYVRTTLRELLVYIVFLVDISLLTYGMTSSYYYTKVMSDLFLDFWTYAQGPLLDGLYWTKWYNIYYENLLLGVPRIRQVKVQNNSCVHQDFIFGCYYREDKEDKSEVGGTAWKYHFEKELGGSTHWGYSGGGFYQDLSASILQELKDNLWLDRGTRAVFIDFSVYNANINLFCVIRLLVEFPATGGAFTSWQIRTVKLIRYIDYFIIACEIIFCVFIFYYIVEEILELRIHYFKSIWNILDVVVILLALVAIIFHIFRTIKVNILLDFEFLAFWQTQYNNMNAVNLFFAWIKVFKYISFNKTMTQLSSTLARCAKDIFGFAIMFFIVFFAYAQLGYLLFGTQVENFSTFVKCIFTQFRIILGDFDYDAIDDANRVLGPIYFVTYVFFVFFVLLNMFLAIINDTYSEVKEELSQKNELQDIIKQSYNRTLMRLKKERIALQSG----SKELEFDFRNCLKTAAFTRFDHDGNHVLDEDEQKIMQNDLELVQHVVQLEHIIGSIVSKIDTVMMKLEKSKVKR

>Pmar01872

RDKYVKTTLRELFVYIAFLINLIILTFGMTSMYKYTLAMDNLFLDFWKVVQGPVLNGLYWEKWYNIYYENFLLGVPRIRQLKVRNDSCVHADFVQGCYYSEDAEDKASFGGTPWKYSTSSDMNGSSHWGYSGGGYYLDLSAILLQDLFDKQWLDRGTRAILIDFSSYNGNINLFSVIRLLVEFPPTGGAVPSSQIYTVKLIRYVDYFILACEGIFCCFILYYIVEEILELSIHYFKSVWNCLDLLVIMLSVVVAIIFNVYRTVTVNKLLNFQFLAFWQTQYNNMVAVNVFFCWIKFFKYVSFNKTMTQLSSTLARCAKDVLGFAVMFFIIFFAYAQLGYLLFGTQVEDFSTFSNCIFTQFRIILGDFDFNAIENANRVLGPLYFTTYVFVVFFVLLNMFLAIINDTYAEVKNELAKENEFDGYFKKVYLDIRKKLNRGNIVLDAS----SQKINYDWRLKLKEAAFAKFDKDKNSVLDSNEMKLMRDELKLVQNVDNLEQSISTMFSKVDVIMNKVAK-----

>Pmar7053

REQYFKTTLRELIIYVVFLIDLCLLTFGMMNTYYYTKVMMDLFVDFWKVVQGPVLDGLYWEKWYNIQYENVMLGVPRVRQLKVRNGSCIYPLFIPGCYYSCGEEDKETFGGTAWKYWSEEELSGSSHWGYSGAGYYRDLSAPMLQELKDYLWLDHGTRAVFIDFTTYNGNINLFCVVRLVVEFPLTGGALPSYQFRTVKLIRYVDYFIMACEVIFCCFILYYVVEECLELRLHYFKSVWNCLDIIVIVLSMVAIIFNICRTIAVSNLLDFEYLAFWQTQFNNMLAINVFFAWIKIFKYVSFNKTLTQLSSTLARCAKDVLGFAVMFFIIFFAYAQTGYLVFGSQVEDFSSFVNCIFTQFRIILGDFDFHAIENANRILGPIYFTTYIFFVFFVLLNMFLAIINDTYTDVKAELATKDDFEDFFKQGYKNMLNKLKKDDIALDSA----GKHLEYIWRKNLIAAAFSKFDKDSDNVLDEDEQKQLRAELEFASRVAKMESSMDSLMSKVESLLTKL-------

Tree file (Newick format)

(923132353:1.17635470,(((923138555:1.19248149,695457291:2.05899041)0.923000:0.27892365,((1174757226:0.95068758,((((6852308720.41726932,(156401013:0.14338053,(Chem014425:0.15581544,828192580:0.51716541)0.928000:0.06690703)0.964000:0.06388437)0.818000:0.04132464,((1067079595:0.31885523,(926626589:0.20954043,(675370064:0.19966442,241171514:0.12894066)0.784000:0.03195506)0.989000:0.07917006)1.000000:0.14253553,((((918287730:0.19527346,((1242857867:0.48730094,405953847:0.10048898)0.992000:0.08863845,(908475875:0.46067677,871233656:0.06771334)0.897000:0.03485327)0.965000:0.05529909)0.929000:0.03286960,((675861632:0.33813201,1236532622:0.21791241)0.218000:0.03480474,(675875990:0.45797832,(Ptor32544:0.23746101,(847028821:0.51145712,(847028815:0.26429831,675439102:0.29482423)0.185000:0.04651541)0.908000:0.05803766)0.996000:0.13472769)0.820000:0.05959979)0.999000:0.10293918)0.859000:0.02941736,((919076678:0.28046322,(919076662:0.16779315,Bner206993:0.36980096)0.633000:0.04087755)0.703000:0.01060067,(PdumPKD22:0.23405831,PdumPKD21:0.11453906)0.457000:0.03745410)0.679000:0.02583727)0.965000:0.04586112,((260802732:0.25741119,(291234667:0.17067146,475551289:0.21832901)0.788000:0.05065302)0.533000:0.03338181,(198420844:0.17442687,(((1160116546:0.63007558,(556956196:0.30666048,(1033394660:0.28867731,5669804:0.36487358)0.997000:0.17370114)0.590000:0.09998765)0.976000:0.13980927,(1160100767:0.21869054,(4505835:0.15605734,(50539686:0.22494931,557013709:0.08376149)0.748000:0.02688463)0.589000:0.04059878)0.998000:0.12466002)0.991000:0.09795789,((Pmar01872:0.31755615,Pmar7053:0.24495527)0.978000:0.08160598,(1160139095:0.18735626,(359465612:0.22585947,(942132194:0.09756341,528498254:0.27091741)0.312000:0.02421756)0.368000:0.01972084)1.000000:0.13221104)0.951000:0.07301968)0.965000:0.05554778)0.969000:0.05961896)0.827000:0.01618821)0.196000:0.01290598)0.444000:0.04059366)0.933000:0.09051690,(1253289448:0.36544496,71999568:0.79957619)1.000000:0.56299737)0.799000:0.05721234)0.000000:0.05020835,(270002831:1.01255611,116008215:1.47891002)0.663000:0.07103494)1.000000:0.40472243)0.338000:0.07045737,(673034662:2.331234 58,(551658247:1.90492533,551655682:1.65773711)0.900000:0.29424141)0.676000:0.09290929)0.966000:0.39923578,(66823157:0.42442313,470247587:0.55616822)1.000000:1.85703798);

PKD1 phylogeny

Sequence alignment (FASTA format)

>684373764

AKSRAFVSTCWNSRGCLCYHLELPNTLDFFLLLFIWARRKDIKDKLGVTPLAENNPDDEYLYEVIVSTGMRRGAGTDSRVCFMLSGTDARVLLQRGNCDRFLLACPRPLGSLIYCRIWHDNSGASWYCNYIGVVDLQTRAKFHFIVERWLAVGQIDRLIAVSTREELMAFTHMFTSITHDLAEDHLWVSVVARPSRFTRVERVATCMLLLFLSMLTSCMFYSAQEVFVAFINNLITLISFIVTYIFRCPWQMRIFAWFLIVSLGVAVAFTTFYGVSFGEAVCQKWLSSLFLSFFASQPIKVVVLALFLSFACKLERARLYRLKQRRANDIIREIFFYVIFLILLVITMSFRAFTEMNTVEDFYAWARTTLIPRLRAKRWYDRTDRIIGYGLMRQLRCYATYYGVGWYSTASELDGVPYMGETNWYGGGYSHMLLESVHWIDFCTRAVIVQLTVYNPNINLFAIITVLIEKPGVGTNLFGQTFQVAYILAYFREMWNYIEFMNFQFLAYWNFFATLKLVRLFRFNQRVGLLGAVLRYAAHDLKYFMIVFMVIFMAFVLVFYLLYTLGPLMFVISMFLAILDESFHSVEMTQFILSQLVVWTG

>Ptor35851

YKVRIYSSGCWKTSGCLCTHLVAPNTIDFFVLLMVWARRKDIKDKIGVTPLVDNDTSDMYLYEIIVDTSIRNGAGTRSKVYFILSGTEARHILQRGNIDRFLLAVPRTLGPLQYIRIWHDNSGASWYCNYFGVTDLQTKEKFHFLVNRWFAVGQVDRLIAVASSDELLVFGELFARKMKHDLSENHLWLSIMARPSRFTRLERVACCMLLMFLTMLTTCMFYTPREAWVGLISGIIGFLLFIVIELFRFQWWVRIIAWILIVVFIVAFFFTTLYGIQFGDMACKKWISAMFISFFASQPVKVIVLACIFTLLCKIEKARAFRLKELKMFDVIREITFYLVFFMTLLLVTYSFRDFTNIRREDQFWNWALNVLVPALRAPLWYDRNNRLMGYAVMRQVRCYGTYYYDTGWYKTMKQLDGIPYMGYMNIYGGGYVYQFLQNNRWLDFQTRAILIQFTTYNPNINLFAICTFILELPGSGGNLLGVCQALYGIFYFKQFWNWIEYMNFQFVGFWNFAATLKFIRLFRFNRRIGMLGAVLSFCSKDLKQFFIIFMIMFLAFVMVFYCLFMFGPIVFMLNMFIAIINEGFSSVEILDFIMTKIILMTG

>847028803

FRVRIFSSGCWGTNGCICNHLVAPNAIDFFILLLVWARRQDIKDKVGLTPLASNDPKDMYLYEVIVDTSIRNNAGTTSNVYFILSGTDPTHLLKRGNVDRFLLAVPRSLGAMQFIRLWHDNSGASWYCNCLGIIDLQTKEKFNFLVNKWFAVGQIDRIIPVAGINELLYFQEMFSRKMKANILSEDHLWLSIIARPSRFTRVERVACCLLLFLTMLTTCMFYSAKELWTGFISNIIAFVLFVLIELFRFPWWMRIFAWFLITLLIVASVLTTFYGIQFGDLACKKWISSMFISFFTSQPIKVICLAIFFALLCKIEKARELRLKELQMYDVIREVGFYLAFFCTLLFISYSLRHFTDITSETLFWNWILRVFVPRIRAPVWYNRNNRLIGYAIIRQVRCYSQYYGFGWYSNESSLDGLPYIGRLSTYGGGYVYKLLRDNKWIDFQTRAIFIEFTTFNPNINLFAIHTYVAEMPGNGGNLLGLCQVLYGIFYFSQFWNFIEYLNFAQVAFWNFCATIKFIRVLRFNKRIGMLGAVLSYTATDMKQFFIIFFIMFLAFVMMFYSLFMFGPIVFLLNMFVAILYNGFTVVELMRFIFTKFVMWTG

>1067095036

YFFKITIVGCWSATGCGSNHLPTVNEINFFIVAMIWGHYKDKKDRRGVIPLPDNKPEDKYIYEVTFVTGPDSDAPCESQIYFMASGTEVRFMYRRYDRNTFVMTTARPLGPLHYLRVFLDNRGDSWQLERVVFRDLQKFEQYTFETNAWLALGLIDRTFLCANNEDTSTFSQRMYVNTNRNANQDHIWMSVFLRPSRYCRKERIVVCTTFLFVSMLLNAMFYALSQMFKGFMVIIFVFPTALLGVIFKLPWWTRLLGWICAVVMAGSLFLVWSYGITWGEITTVKWFSSFFTTFLVSQWIKVAFVSCFGSLCCATKNLRVKLTKERDMKFIVRGVLVYCIFLAVLLIIVTDKTNPNQLRNTDQFWLWAKLAVMNITRAQRWYDRANRMVGYPILRQIRCTGDRYCANWYQTSRQLKSFTSVGRLGTYAGGYVIRLLQKVHWIDKRSRAVIFEFSVYNANVNLFMTGTIYFEINEGGGRLLRLCELTFVVSYFLMYWNIAEYLRMDYPVLIDFLSILKLIKLLQFNKRMDVLALTIALCWDELSYFFIAFTVIFFAFCCLFYLIFYLSPLLFLVNIMLAIILKAFNEIDVMDYMWDNFKKYLV

>1005480473

YTVSVYFAKCWRDRGCLCNHAVEPNPISTYFIGLIWARRKDRADKVGVCPLPDNDPSDPYRYEVTLWTSMRRNAGTKSKVSMIVYGSEARSLFRRGSQDTFLMTTPFELGTIQFIRIWHDNAGGSWFLSRVMVIDLQTDERSYFICNRWLAVGQVERIVPLASQQELTEFGHLFVARTRRNLSDSHLWFSVLARPSRFTRVQRLSCCFCLLLMSMMASGMYYSWEEVLIGVISSLIVFPNLFVVQIFRLPYYFLYVAWLVFIFSVGSAVVVVFYGMQFQNKKSLQWLFSSSVSFVQDQPLKVLGLAVFFALVVKLIAMRTARMKERKMYSVIREVITYYIFLTLLLTIAYTHRSFDKICSVKEFWNWTHSELIPGLFPTKWYDGSSQMVGGARMRLIRCYGEYYNVGWYQNASVLHGYPVWARLNTYGGGYVIVLLEKNMWIERHTRAIVTEFTIYNAQTNLFSVVTLVAEFPATGSRLYNYIDILYVLFYIKEFWNCVEFKSFQFAAYWDMLANIKLNKLLRFNKRFSLLSSTLKYAYYPLLMFSIVWGIVFIAFTFVAALVFGLGPIFFLINMFVSIVIDAFWVVEIVDFIIERFKLWSG

>1139824617

YKISFYSASCWTTEGCLCNHLVPMNTINLYVLMVIWARKRDLLDMAGSSPLPDNDPRDKYLYEMTVFTGSRKGSGTTANVSFILTGTPPRYLLQRSNCDGFIMATPKSLGKLTHVRIWHDNTGPSWYFSRLQVQDLQTGDKYFFICEQWLSVGMIDRMVPVAGKAELTSFNYLFWQKTKHNLADGHLWFSIFQRPSNFTRVQRLTCCLTLLFTTMVVSLAWYSLTGVYIGIMSGIMSFPNFIIVLIFRLPWWSIYIGYLSFITVMVAFYFAVEFGGVFGLNKSVSWLIGFIMSLFESQPIKVIIFAIFYALVIKLKTIREERQKELQMYAMLKEIAICMVFLYIVLTIAYANRDFVDIDSRTQFWNWSENALLQGLYNNKFYDKQSILVGGARLRQLRCMPAYYKLRWYQSMWELSGYPYQGIFQTYGGGYVAELLRDNSWVDKRTRAVFIEANIFNPNMNLWGISLYLCEFLQTGAKLDRYIEISALAFYFKAFWNVFEFVNMHFTVMLDFFLTVRFLRLLRFNKKISLLSSTLRHSAGMLISFSFVFNIAFLSFACTAYLFFMTGPLLFLINFFLTILMEAFEAVEVIEYLMKRMKMMMG

>196000368

YSLAFYDLGCWSSSGCQCNHLVTPNILDFFIVISVWAWRQDKKNKFSNTVLIDNRPDDLYHYEITVFTGLRVGSGTTANVSLIISGTSPRLLLQTRDIDTMILTCSQDLGSINYIRLYHDNSGPSWFLSRIIIKDLQTNEQYCFICNKWLAVGLIDRTLYPATTEELRDFMQSFMSKAAISLSDNHIWLSIVTRPSHFTRLQRVGCCFAILLLSMLTNAMFYTWDGLVRGLISGAIVFLSFMMVQFFRLPWWCLYLGWLIVISCGVGTAFTLFYGLSFGSELCLQWLASLTISILQSQPIKVVVVAFIFALIVRVNAARSLRLKQRKMQSIFREILGYVVFLTVTILICYGNRDFAKIATINEFWRWTNQVFLPGLYVGTWYNRQSYLVGVARVRMLRCAGSYYTVGWFQTSSMTGSYPYWGQYHIYGGGYVVELLQQQGWIDHQTRAILIEFTIFNAQVNLFSVVTLLAEFPATGGRLFRDCEIIFVSFYFSGFWNIIEFINFHYVATWDFFGTIKLIRLLRFNRTIAILSTILRSAAKEILMFLVVFTVIFLTFAQFAYLIYGVSPLYFLINMFLTLIIKSFEKVEMADYMKERFKSLVG

>PdumPKD916

YAVKVWTPSCYKTDGCLTEHLVPMNSINLYLVMLIYVRKKDKTDHAGATPLLDNNPMDKYLYELVFFTGSRKNAGTSAKVHCIISGTEPRLILKRSGVDVFLMSVPQPLGPLTHLRIWHDNTGPSWYFERLQVTDLQITEDVYFYVCDRWLAVGMIDRVMPVAGMEELTHFNHLFWIKAKRDLTDGHIWFSVFQRPSRFTRVQRLTCCLTLLFSSMMASILFYTPQQIYVGFISSMIALPNIIIALIFRLHWGFKIFAYFSFCIVAVTFWVTVEVAGVFGKKK

SLQWLLAFVVSVVESQPIKVLVLAAFYALVIRLQEARQMRFKEIKMNKIIWEIIFYFLFLWLLMMVAYGNRDLQEVRSRAEFWDYVSQVLLPGLYDNKWYDGTGYLVGQPRLRQVRCEKGYYSFGWYRSWLDLSGYPY
WGKYNSYGGGGYTADLLRSNRWVDSNTRAIFIEFLVYNPSVNLFGISMVAMEAQESGARLDRYGEILTLVFYFKDSWNIFEFHNLQYAALWDSLSTLKFLKLLRFNRRMAMLSSTLRYAAGPIIQFLIMFLILFLAFVLASYIV
FGFGPTILLINVFITLICEAFSIVEIVDFVVDRFKMVTG
>919098890
YQTSTFFATCFVSDGCICYHLIPMNTIDFYMVMCIWARKQDKIDKAGTSPLPDNDPRHKYLYELLVYTGGRSGGGTTAKVSMVFTGTEPRLIFERGGVDCFLMACPRPLGQLAHIRVWHDNSGPSWFFSRLQVTDLQTK
QKFFFICDRWFAVGLIDRVIPLAGKEELTQFNHLFWTQTKKNLYDGHIWFSIFTRPSKFTRVQRLSCVLSLLFTTMLSNIMFYTLEGIITGHIASLIVFPNLLIVQLFRFPWWGIYIGYVCFISVVVSIYFTTEFGGPFGAEKSAE
WMASFFISLLESQPIKVFILAVFYALIIKLKSAREQRFKEMAMMQIIREILVYLFFLYMLFTVSYGNRDLGSVRTQGDVWKYINQTVIPGIFSEKFYDYQSYRVGAPRLRLVRCRESYYQKGWFRDSGETDGYPVFGLLTMY
GGGYVAELLYDNRWVDRYARALFVEFVVWNPNVNLFAISSIAIEFPESGGRLDRYCELITIIFYFKKFWNVVEFVNFQYVGYWDFLAVLKSMRLLRFNKRMSIIGVTLKEAGKPMITFGIVFFVLFFAYAQLMYNIFGFARI
VFMVNLFLTLIMEAFAWAEIVDFMVSRFKGLFG
>443713927
YTMDAFFASCWSGDGCCCNHLVAPNTLDFYILCALWARHADRKDTVGVTPMPDNKPGDKYLYEIILCTGLRKNAGTDSKVHFILSGTDVRILLRRGSKDAFLMSVPHSLGPLNYARIWHDNSGSGWYLNYFIVNDLQT
RQKHIFICNRWLAVGQVDRILLPVASEEQKTQFGHMFPNKAAKNLGDDHLWFSVIARPSRFTRLQRVSCCLCLLFTSMLANALFYTPAQIYIGVISNLVSFPNFLVAWLFRLPWWCRWVGWVLWLSTLLSMVFVMFYGIQF
GDLKTKKWITSMMISLLTSQPIKVFILALLMSILFKLEEARLERLKELQMYDILRELLVYSFFLWILLVISYGFRDFEKMRTWDNYFEWCHNTMIPGLRADKWYDRVSRIMGYATLRQLRCNKPYYGIGWYRTSSEMDSYP
FWAYHGFYGGGYVVELLIKNQWIDRHTRAIFLEFTTYNPMVNLFGICNFVAEFIVTGGNLLGFCQIIYLVFYVKEAWNWVEYMNFRYVGYWNFLATVKFIRLLRFNKRMGVLASTLRYAAKSLFMFGIMFMIVFMAYA
MFFYLIYFLGPLFFLINMFLTILSEAFSAVEMVDFIFKRFMQWTG
>PdumPKD12
YRTEFTCMGAWHGEGLQSTHLVQPNKIDWFIIMLIWARYKDKQDKLGVCPLPDNEKEDKYLYEILVYTGLRTNGGTESKVNFIISGTDVRTFFKRGAVDAFLMATPHPLGHLNYLRVWHDNKGGSWYLHYVIVSDLQ
TRHRYYFICNRWLAAGSVDRILLAVAGEEEMTEFGHLFAHSTKKKLDDDHLWFSVVARPSRFTRCQRVCCCLLLLYTTMLADAVFYTPQQFYVGHSNLIVFPNFLVVFLFRFPWWCVIIAWLLVLSVAVSVCFVTFYGIMFG
DDKCKKWLTSMLVAFFTSQPIQVLVVAALLSVICKLEKARQQRMNEKQMYAVMREIIFYFIFLWMLLVISYSFRDFKNLKTMKQFFDWSENVLAPELRAGKWYDRNNRIMGFATMRQLRCNVAYRGVGWHRSAKSLDS
FPFWAVHHVYGGGYVRELLHDGGWFDHYTRAVFIEFTVYNAQVNIFTICTLVAEFLPTGSNLLGYCQIIYMLYYFTEWWNWVEYMKFQYVGYWNFLATIKFLRLLRFNRRMSMLASTLRNCAKSLANFALMFLVVFFA
YVQFFYLLFFLGPLFFLINMFVTILNEAFADVEIVEFVVKRLKSWTG
>919074653

MVEFGHLFTKTTQKNLGDGHLWFSVVARPSRFTRLQRVTCCLCLLYMTMLTSAMFYSPEQIYVGVIGTIIVFPNLLIVTFFRLPWWFTIVAWLMWILILGSAFLVTMYGITFQETKAQKWTTSMIISFLTSEPVKVFLLAIFFSL
IIKLFEARERKRIKELKMYDILREIIFYCLFLWLLMVISYGNRDFTKLQNSTHFWHWYKEGLVPGLRASVWYDRTCRMMGYATMRQVRCNEAYFGVGWYRTSSELDGYPHYGLLTVYGGGYVAELLEREKWIDKYTRAV
LVEFTCYNPGINLFSIGTMVAEFPETGGRLFNYCEVSYFVFYFSEFWNLVEYIKFQYVGYWNCLGTLKFLKLLRFNKRMSFLAATLRHSTKHILFAFAVMFFIVFFAFVQFFTLDMGLAPYFFTINMFLSIMNESFTAVEMVD
FIVTRFKLWTG
>918302224
YTLRTFTSGCWVTDGCFCNHMVAPNAIDWYLGFAGWSFWQDKKDKLGLTPLKDNNPNDKYYYEIIVATGMGKDSGTTSKVYFILSGTEARSLLQRGSVDGFLMAVRRPLGALSYARIWHDNSGGSWYLKHFLVRDL
QTGMKYHFIANKWLAVGQVDRIIPVAGKEQLLNFSHMFAKRAQKNLGDGHLWFSVIGRPSPFTRLQRVSCCLCLLFTSMFANALYYSIETISMGFVSNLIVLPNLIVTLFRFPWWCRIISWLLWLTVAGSCAGVTFYGITF
GNEKCTKWISSMLISFFSSEPIKVFGMAILLSLILKLGKMREQRLKEIKMWEVIKEIIFYAFFLWILMVISYRNRSFSEITNTQAFWKWAKNDLVDGLRATKYYDKTSRIMGYAVMRQLRCHKEYYGKGWYQDSSKLDGYP
FYGSLNMYGGGYVARLLQNTSWIDKLTRAVFVEFTIYNPQVNLFTVCTLLAEFTRIGGMILLPYCDVIFVIFYLKEFWNLVEYMKFQYVAFWNFLATLKFMKLLRFNKRMSMLSSTLTNSAPSLLNFSLVFWIFFLSFVQVFY
LLYFLCPLIFLLNMFLSILNEAFASVEIVNFMVGRFKTWTG
>676474191
YKIRVFSTGCWSSKGCFCNHLVVPNTIDWYLIAVIWARREDKKDRLGLTPLKDNDPRDKYYYEVCVVTGMRRNAGTDSKVCFILSGTDVRMFFRRGAVDAFLVAVPRPLGTLNYARVWHDNSGASWYLKYLLVRDVQ
TNQRFVFISNKWFAVGQVDRVIPVAGKEQMTEFGHLFAEKSRKSLVDGHLWFSVVARPSRFTRVQRVSCCLLLLFVSMLGNAMFYSPEQIFIGIAVNLIVFPNFLVITLFRFPWWCSILGWVLWIGVGLSLAFVTFYGITFQ
DEKCKKWITSMLMSFFTSQPVKVFLTAIFFSLVVKLEKARIQRLNEIKMWEIIREIVFYSFFLWILMVISYRNRSFSKVRNTAEYWKWSQSGLLNGIRASNYYDKVSRIMGYATMRQLRCNEAYYLPRWYTTANELNGYPY
WGKHEMYGGGYVVQLLEDDDWIDQYTRAVFVEFTIYNAQINLFAIATIVAEFDSAGGLLLPYCEVIYLIFYFTQFWNIIEYMKFQYVGYWNFFATLKFLKLLRFNKRMALVAATLKNGSKSLMHFGIIFLIVYMAFMQLF
YLTFMLTQIFVLVNMFLSILNETFGAVEIVDFMVGRFKRWTG
>1207917521
YNVSISQANCWEANGCLCSHLVAPNTIDWYIVAAIWARRQDKKDKLGLTPLVDNDVVDKYYYEIVVVTGMRRNAGTNSKVNFILSGTEVRCLFKRGGVDGFLLAVKRPLGPLNYVHIWHDNSGASWYMSSMVVRDV
QTDDKFIFINDKWFAVGMIDRIIAVAGREQMTEFGHLFAEKTTKNLNDGHLWFSVVARPSRFTRVQRVSCCLCLLYVTMLSNAMFYSPTQVYIGVLCNIIVFPNFLIVYIFRFPWFFTIIAWLLWLTVLVSAAFVTFYGVMF
QDVKCKQWITSMLISFFTSQPIKVFAVAILFSLIIKLEKARAVRMKEIKMWAVIREIIFYSFFVWILMVISYQNRSFEKITNVQNFWTWAKTGLATGIRANKWYDKVSRLMGYATMRQLRCNGEYYQVGWYTSAEELNGYP
YWGQMALYGGGYLVRLLEREKWIDKYTRAIFVEFTVYNPGVNLFSISTLLAEFRPSNGTLLPYCEVIYLLFYFKQFWNLIEYMKFQYVGYWSFFGTLKFLKLLRFNKKMSLLASTLKHSCKDLMHFSIIFNIVFLAFIQFFYL
IYVLTQVSLVVNMFLSILNETFSAVEIVQFMLGRFKLWTG
>762156788
YTIQTVSSTCWDTRGCKCTHMVVPNTIDWYIVTVIWARYKDKKDMLGLTPLADNNKKDRYYYEIVVVTGMRRHAGTDSKVSFILSGTDVRLFFRRGSVDGFLMAVPRPLGALNYLHVWHDNSGASWYLRNVVVRDV
QTDQKYIFIANRWFAVGMIDRLIPVAGKEQMTEFSHLFTEKTTKSLSDGHLWFSVIARPSRFTRVQRVSCCLCLLFLSMLTNAMFYSPEQIYIGVVSNIIVFPSFLIILIFRLPWGFAIVGWLLWAATLVSAAFVTFYGVMFQD
VKCKKWITSMLISFFTSQPLKVILTAIFFSLVIKLEKAREVRLKEIKMWGIVREILFYSFFLWILMVISYRNRSFLKIKNKDDFWVWAKSGMVDGIRAGVWYDKQSRIMGYATMRQLRCNDDYYGVGWYTTANVLDGYPY
MGDRVWYGGGYTVRLLEKEGWVDRYTRAIFVEFTVYNPGINLFGIATLLLEVSPSQGMLLPYCEIVYFAFYFKQFWNLVEYMKFQYVGAWSFFASLKFLKLLRFNKRMSLLASTLKNSCRDLMHFSIIFNIVFLAFIQLFYL
VYAMTQIFIVVNMLLSILNETFGAVEIVDFMMNRFKLWTG
>871215788

MRRNAGTNSRVFFILSGTEVRMLLKRGMTNGFLMAVPESLGRINYMRVWHDNSGRSWFLNYIVVKDLQAETKQVFIANRWFAVGQVDRIIPLAGKEQLEDFSYQFKERSKKDLVDGHLWFSVIARPSRFSCVERVACCL
CLLFMTMLANAMFYSPSQISIGFISTLIVFPNFLLVFLFRLPWWCRPLAWILVVTTAVAVAIVTFYGISFQDSTCKKWITSLIISFFTSQPIKVILTAVFLSLIVKLERVRSKRIKEVKMWAVVREILFYCIFLWVLLVVSYRNRSFE
KIQNTDEFWEWAKTGLINGLLAGSYYDKASRILGHAVMRQHRCNTLYFEVGWWTSAETLNGYPYWGLLGVYGGGYVAHLLEREKWVDRYTRTVFVEFTVYNPQVNLFSVNTFLAEFHSTNGMLLPYCEIAFVVFYFT
KFWNWVEYIKFQYVGYWNCFATIKFLRLLRFNRRVSMLSSTLRHSARNLFHFSIIFWIVFLAFCQLFFLTYMLTQIFVIVNMFISILNDTFTTVEILEFMMSRFKKWTG
>871254812
FSLRQYVTQCWSTDGCLCNHLVVPNTIDWYITAAIYARYKDKKMQLGIAPLADNDPKDKYFYEVLVCTGMRRNAGTNSQVCFILSGTDVRAFFRRGQVDGFLMAVPRPLGYLNFMRVWHDNSGGSWYLNYIVVRDIQ
TDLKQVFISNKWYAVGQVDRVIPVAGREQLSDFSYQFGERSKKNLADGHLWFSVVARPSRFTCLQRVSCCLCLLYVTMLANAMFYSPEQISIGHISNLIVFPNFLLIFLFRLPWYFTIIAWVLWVVTLGAAAMVTFYGISFKD
DICKKWISSMLVSFFMSQPIKVFLFAMILSIIIKLEKHRKRRMNEIKMWAIIREVIIYSFFLWILMVISYRQLGFSLIRNTDEFWIWAKSGLVNGLRAGSYYDKVSRIIGYATMRQLRCNIQYFKVGWYTTAGDLNGYPYWGTL
GMYGGGYVVNLLEREKWIDRYTRAVFVEFTVYNAQVNLFAISTILAEFHPSGGMLLPYCEVVYILFYFKSFWNLVEYVKFQYVGYWNFFATIKFLKLLRFNKKISMLSATLKHSAKGLMHFGVIFVIVFMAFAQLFYLTY
MLTQIFVIVNMFISILNETFTVVEVVDFMLQRFKKWTG
>908400046

MRIWHDNSGASWYLNYVIIRDVQTDVKQVFIANRWFAVGQIDRVIPVAGREQMADFGYQFGERTRKNLVDGHLWFSVVARPSRFTCLQRVSCCLCLLYVSMLANAMFYSPEQIYIGVVCNLIVFPNVLLIYMFRFPWYF
QIIAWLLWIVTLGSVCMVIFYGISFKDDTCRKWITSLIISFFTSQPLKVFLFAMVLSLIIKIEKLRAQRMKEIQMWAVIREVIMYSFFLWILMLISYRQLGFLTLNNRDDFWTWANSGLLSGLRAGPYYDKVSRILGYATMRQL
RCNIMYFNVGWYSSADKLNGYPYWGVLSLYGGGYVVPLLQSEQWIDRYTRAVFIEFTTYNPQVNLFSVNTILAEFDQTGGMLLPYCEAVYMLFYFLSFWNLVEYVKFQYVGYWNFLATLKFLKLLRFNKKISMLSATLE
NSARSLLHFGLIFTIVFMAFSQLFYVTYMLTQIVIINMFVSILNETFAAVEIVDFMLQRFKKWTG
>780039266
YGFRVLSSGCWLQDGCSCTHLVPPTPLDFFIFILLWARRRDRKDKLGVTPLVDNNPAHNYYYEIVVMTGQRKAGGTDSKVSFIMSGTSVRDFFRRGASDRFLMAVPGPLGTLNYMRIWHDNSGQGWYVKYIAIRDIQTR
ERFYFIVNQWFSVGQIDRLIPVAGREQMQNFSHLFSNHTRKNFNDGHLWFSIIGRPSRFTRVQRAACCLMFLWLEMLVNIMFYSPAQISIGVQANLIVFPSLLVVQFFRFPWWCVIVAWVLLLSMAAATTTTIFYGIQFGNT
ETSEWLCSIIISFVFGQPIKILLLATFIACIVKLERARFEQRVKEIKMYSIFREIFFYVVYLWVVLVISYGNSDFNKINSHASYWAYLHSTLVPAIYPQTWYNQNAFILGYPVLRQARCNVAYYGPGWFTAWDEIESYPVLGRQA
VYGGGYLAKLLFEEAWFDRYTRGVFLEIAFFNAQVNLFGVIDLVAEFLPTGGRLFTYCEILFLGFYFKDFWNMAERINLQYVAYWNFIGTLKFIKLLRFNKRIGILSTTIGSCTKDLFHFGIMFGIFFLAYALAFYKMYALG
PIFFLVNMFLTIVIEAFSSVEMVDFMIGQLKKWVG
>1229191925
YTLCVYVSGCWSTEGCFSRHLVPPTPLDFFAFIFIWARRKDRKDMQGLTPLADNKPKDKYLYEIAVMTGQWKGSGTDSKVHFILSGTEVRMFLQTGSIDRFVMAVPRPLGTLNYMRLWHDNSGQNWFLSYAAVRDLQT
RQRFYFIANRWFSLGQVDRLLPVAGREQMQGFSHVFSSHTRKNLNDGHLWFSIFLRPSRFTRVQRTACCLMALWLEMLVNIMWYSPAQISIGVQSSHVFPTLLVVQIFRLPWWVVIIGWLCILSIVTATVFTVFYGISFGDQT
TKEWLTSLVISFITGQPIKILLVAMFIALIIKLERIREQRMKELKMYNILREIAFYIVFLWTLMVFSYSNVDFNKVSGIKHFWRWVDTGLVPNLYADRWYDRQSFILGKVVMRQLRCNEEYYGRAWYIPASELEGYPFLGKH
ALYGGGYVVVVVLYNETWIDRYTRAVFIEFSCYNPHVNLFAVVCLTAEFLQTGARLLTYCEVMFFCFYLRNFWNWTEYIKLQYIAYWNFLASLKFLKLLRFNRRIGVLSSTIHNCSSDLMYFGIMFGIVFFAFATAFYLIFGL
GPFFFLINMFLTIINEAFKIVEMVDFMIRRLKLWTG
>957833931

MTKGSATDSLVKCIISGTDVSTFFRSGAVDIFVLSVPGPLGQLNYVRIWHDNSGGSWYLKYTISDVQTGEKFKFIAEKWFAVGQVDRLLPVAGKEQLEEFGHLFQNKAQRDLKDGHLWFSVFSRPSRFTRLQRSTCCLSL

LMTTMLVSAMWYTSQEIGVGLMSNLIIFPNLLIVMIFRLPWWFVIIAWLAILTVPTSMFFILLYGIQFGDLLCQKWLSSLFISFLFDQPLKVFLLAVFFSCVCKLNAARETRFKEMKMFEILRDLGIYIFFLWIVWVVSYSQRDF
EGLYRLERVWNWTQNVLPGALVTTNWYDRTNRRLGPAWMRQVRCAD-SFSSGWYRPGVTMKSLPYWGRLDRYGGGYYVFFLERNRWIDKYTRAVFIEFSTYNAYVNLYTVCTILFEFPSSGGTLN-
ECQIIYAIFYFRDIWNWNEYTTLQYINYVDFLATLKMIRLLRFNRRMSLLGQTLASCGKDLFAFTVMFGFVFVAFTMMFQLLLGVSPVFFLMNMFLSIIIVSFNIVEMVDFIFGRMKTWVG
>291224264
YSLLVSFSGCWSSLGCLCNHLIPPNDIDFYFIGLIWARRKDRSDLIGVAPLLDNEPNDKYLYEVVVFTGRRKGGGTDSKVGFILSGTDVRMLLRRGAVDSFLMAVPGPLGTLNYMRIWHDNSGQSWYLNYIAIRDVQTGE
RFYFIANQWFSVGQGIDRTIAVAGKEQMQSFDHVFSSTTRKNLDDGHIWFSIFARPSRFTRCQRISCCAMALWLEMLVNIMWYSPAQISIGVQANLIVFPSLLVIQIFRFPHWCIYLAWLIFLSIAVAVTFVLFYAVQMGDRETR
QWFTSLCISFLTAQPIKVILIGIFVALILKLEAARENRIKEIKMYSIIREIIFYVLFLWILLVISYGNRDFLHVSTKHKFWSYAHDVLVPGLFADEWYDRVSHILGYPVIRQLRCNEEYFRVSWYTSASELDGYPYLGRHAFYGG
GYIARLLEAENWIDKYTRAVFIEFSTYNAQANLFAMVVLLMEILPTGGNLLNYCEGTFFAFYFKGFWNWIEYIKLQYVAYWNFLANLKFLKLLRFNKRMSLLSSTLRAASKDIISYFFMFAIVFLAFAEAFYLIFTLGPVLFLI
NMFLTILNESFGDVEMVDFMLNRFKLWTG
>1009565777
YFIALYTTGCWSGDGCLCNHLVAPNTIDFYVLLAIWARWKDRKDKLGATPLADNDMADKYFYEILVQTGFPLNAGTDSKVHFILSGTDVRTFFKRGSLNSFVMSVPRCLGQLNYLRIWHDNSGASWYLKHIVVKDIHT
GYKYEIIAERWFAVGLIDYLIPVAGKEQTTEFSHLFSNASRRNLTDNHLWLSVFTRPSRFTRLQRLSCCMAMLYLSMLTDAMWYSPEQIGIGLLANLIVFPSFIMIVLFRLPWWCSYIGWFVIASIACSIFFIWAYALQFGNEKT
RKWLTSLVVAFFSSEPLKIFLLALVMACVCGLDAARKLREKEVQMYQVLREMGAYILYLWILMILSYGNRDFNNVNTEVRFWNWTRHSLIPELFAMEWYDRNNFKIGYPVLRQVRCAGFGYKKFWYMTAKDLNGFPF
WGQLDWYGGGYIVPLLEKTGWIDKDTRAVFVEFGTYNPQVNLFVVLTIVAEFLPGGGKLLHYCQIGFILFYFKEYWNIAEYIKLQEALLLDSLATLKFLKLLRFNKRIGGMISTVRLCVQELKGYSVCLITFIAFVCLFW
MTLGLAPIAFLVNILLTIIIQSFEDVQLMEFMIRRLKLMLG
>675368552

MRPSRFTRAQRVTCCFAMLFLSMLVNAMWYSPEQIGVGVMANFIVFPTFLMITLFRFPWWCRIIAWLAIASIVVSCFFLWAYGVQFGDEKTKKWITSLLISFFSSQPIKVFLLAMFFSAIFKMERAKLERLKELQMTAILKEIIS
YAFFIWILTILSYGNRDFMAITTTADFYKWTREVLIPELIVGKWYDRVNRIMGYGMLRQVRCRSHSYQPGWYRTSKELNGLPFWAIRDVYGGGYVFPLLEKSDWIDKRTRAVFAEFSVYNAQVNLFGVMTIVAEFLPGG
GRLMRYCELTYVIYYFKSYWNWAEYIKMQFVASVDFIAILKFIKLLRFNKRMGILYSTLAMCAKDLKSFFVVFVVIYLAFVQFFYLVFGLGPLSFLLNIFVTLILSAFQDVEIVDFMWKRLKGFFF
>1078802969
YSIRIYTSGCWSGEGCKCDHLVVPNTIDFYLMLLIWARREDLRDKLCAHPLPDNDHRDNYLYELTVFTGDKDGAATDSRIFFILSGTDARSFFRRGGQDTFVMACPKPLGRLNYMRIWHDNSGRSWYLKFISIRDVQTNA
RYDFIANRWFAVGQVDRLLPVAGKEEATEFNHLFQLKSQKNLADGHLWFSVFLRPSRFTRCQRVSCCIALLYLSMLVNAMWYSPEQVGVGVLANLIVFPTFLMITLFRLPYWCRYVAWVCLISIFTSVFFIWAYGVQFGD
EKTRKWITSLLVSFFTSQPIKVFMTAIILSALFKLDHVREERMKELKMSEILREIFSYLLFLWLLTVLSYGNRDFMEISSTDGFYNYLRETVVPQLKIGDWYDRVNRIMGYGVLRQIRCRGQSYEPKWHKKAKELHGLPFW
GKLDVYGGGYVMRLLEVSDWIDGGTRAVFVEFSVYNAQVNLFGVVTIIAEYHPGGGRLMRYCEISFIGFYMRSYWNIAEYVKLQRVAAVDFISILKFTKLLRFNKRIGVLYSTLSQCSKDLSSFFVVFWVIFLAFTQVFYVI
LGIGPMVFLLNLFVTLIISSFQSVEIISFMGRKVKGFFG
>1238854278

MTVPRSLGRLNYMRIWHDNSGRSWFLSFIVIRDVQTGQKYEFIANKWFAVGAIDRLLPVAGKGEATEFSHLFNVTSRKNLSDGHLWFSIFMRPSRFTRCQRVSCCIALLYLSMLVNAMWYSPEQIGVGVMANLIVFPTFLM
ITLFRFPWWCLYLAWVSLSTIAVSVFFLWGYGVQFGDEKTRKWVTSLLISFFTSQPIKVFLMAMIFSTLCKTQRAKEERLTEMKMNAVIREILSYLLFLWILIVLSYGNRDFTQATTPLLFWRWVREAVVPELLVGKWYDRV
NRLMGYGVLRQVRCSGMTYSPGWYRSSSELDGLPFWGKLDVYGGGYVFPLLETSDWVDSRTRAVFVEFSVYNAQVNLFGVVTIVAEFFPGGGRLMRYCEVFYVVFYFQSYWSIIEYVKLQFVAAVDFIAILKFTRLLRFN
RRMGFLYSTLSQCSKDLKSFCVVFLITYLAFVQVFYLIFGVGPLIFLLNIFTTLIISSFQSVEIISFMKNKFKSMFG
>919100788
VTILSTAIKPWGTSSFKSNFFIPPNTIDFYFLLLIWARRKDKQDMWTHHHVLENERMDHQHFLISVFTGSKSPKDTTVVPFVILYGTRPKVLYQKGSVYNFVLSTPRTLGKLTHLHVWHCDLDADWFLDRVIVYDVQRKE
RSLFLCNQWLSLGQTDHLLSAADDGSLEDFNFLLKNNASQNLFDDHLWFSVAKRPSRLSRVQRLTACLAVLFLCMISNAMFYSFRQVYVGLISSLLVVPTFVILEIFRFPWWAIVFGYLSFLAIATSAFFVLLYSLDWGATKSS
NWLLSMFFAFTESQPLKVIAVAVVMSMILREERKRLRRKHDAQLSGSLREIFHIIYLILLIFICFANRD-
NSVSRFEAIFAWLSDRILPIVFPTVDYDIYGLRLGPLRIRQKRCNSDYYEPAWYTTPDGDKDMAATGMFATYGGGYVFVEILKADSWIDRLTRAIFVEFTINNMNVNVFSMVRIKVELPVVGGRPYPYIQVIWALVYFRELW
HYVEFVGFDIVANWDFLATLQLFRPLSFNKQLAILRASLSRSSIYLFTYGLIIVIILTSFGIGTHLLSGLLKVFFIINMFISVINDTLYEIELADHLTMKIDTWFR
>Pdum14759
YRVRIAEIKIWQEHVFKSNLFSGPNTIDFYVPLLIYCRKKDKQDKWTTLPLVANVASASYRYEIHVYTGIKRHSQTKSKVHFVLSGTGVRQMFCRGDINLFLLNTQQDLGELMSMRIWHDGSGCGWFLDKVIVIDLQRN
HWWVFHSNKWLDNRQTERVLKVSTFGSSSNTGVLLRNNLNRKFFDDHVWLSVAFRNSVFTRAQRLSVCWAILYLTMVTNAMWYTMQELTNSIMASLIVIPILAITLFFTLPHWCIHLGWLIFLAISVSALFTIFYALEWGG
VKSRAWLIAFFLSFFQSQPVKVLLISALLSFLFKIEEAKKKKEMDKKINNMMREIFMQAFFLMMLLIIAYGNQD---
IDKIPEIHAWLKTNFTDEVYYYKNYNTIGMRLGPPRLRQLRCTHHYHSIGWYNTSSQLRGLPFWGFLSTYGGGYVANLLNSSRWIDQYTRAVFVEFNVFNMQSGLFTQVIISFELPPTGAKLYRYTEILCTLVWFKSFWN
WTIVANFQIVTTADFLVMLQFLHILRYNKTIGVLNATVAQSASDLIGMGFLVFVIFAAFLSLGYLGFGPGQVFLMINMFITLLCNYMEAVLVVDHFFENLRGLAM
>Pdum15290
YSLYLYEIECWNSMSFKSNLFQAPNKIDFYVVGLVVCRKFDKMDKWQCLPLIGSVVGGEYQYEVHVFTGFKTKSETTSKVYIEILGTGTRELFRRGECNAFKMSTTNDLGKLYAINVWHDGSG-
DWYLSKVIVVDCHRNEWYQFGCNKWLSEKMIKRTLDVSVLESHVNTGVLLRNNMSRKFFDDHLWLSVAYRKSRFTRAQRLSCCMALLYLTMIGNAMWYSVSMLVASVYSSLVVVPIMLITFLFTLPYWCIFFGWLVFLSV
AASAFFTILYSFQWGGVKSTGWLVAFLLSFFESQPMKVMLIAMIISFMFRFDENVLKLRRLDNKVGIMIKEIISNFLVLFVMMCVAYVSYD-
PEVEHFLNVWPWVEKVLIPSLYYVEHYNGMAIRFGPPRLRQLRCIPRVYLTNWYQSASSLKILSYWATLEVLGGGYVANLLKAMQWIDHRTRALFAEFNIYNANTELFTQVILAFEFPSNGGQLYRYMEIISLFVYLKNLW
NWVHTVSFQHVTLGDFLLTLAFLNILRFNWTIAVLAQVLVSSAKDLTLMALNILIMVMAFASMGNVLFGAGRVFLVVNLFITLLCTFVDAVEVVDHFLDMVQGFVL
>260827108
LELLLFSTTCWEDEGCICDHLFPPNEIDIYLIAIFWARRKDKEDKVETMALHDLDPMDMYRYDVTIITGFRRHAGTTASVSLRLIGSETHMVLQRGAIDSFLITTRDSLGDLREIMVWHNNAGPSWFLSRIMVQDLQTKKR
WYFLCNSWLAIREVEKHICVATEAELHQASTLFATKAVFALQDDHLWFSILARPSRFTRVQRLTCCVSVLLCQMLTNIIFYTWKELVIGVESALIAFPNFVIVALFKLPWWCVYLAWLAILSCLTAAFFTLLYGMAYGRQKSI
DWLVSMIIGFLQGQPLKMLVIAAIVAVIIKLEKVRAARKKNDAMFFIVYETAAFMLFAFLILTIAHTMREVEEIQDRYDVYLWLEEEIFMNLYDFS--DGQSVVVGSVRMRQVRCSEDYFLTSWYQHTDSVWSF------
SGYTGGYVADFLEDQKWIDRQTRAVVVDFTIYNANVDLFCLVSVLIETPDTGNRLYQYFQVAYTLVFFRNVWNLIDFLNLVPAIVADFVGCMRCCRLLRGNNSLYLIEATLTRSMAPLLSHGLIIVIMMIGYAVFGHLAFG
LGSIFFLLNFFAAIMMDIFTVEEIVTYTIERVKSLLA
>828232853
----------MTTTVCLGDLSYMQLWHDNTG-
GWYLRDIEVVDLQTEEHFYFMANCWIAVSSLDLVLPVSGKENITSFNYLFFSNARNGFTDEHLWLSVVTRLSNFTRCQRISVSVCLLMTSMTASAMFFTTKQLYVGIVSAFITLPNILLTHLFRFPHWFLYISWICITTILGLGY
VILLYGMAFGNKKSLDWLASISLGLIHDQPMKIFLFSILAALASKIESMKAAHLTELKMQSITWELLMYFLFSIVVFYLGYVSRGFETLQSSDDFWPWMDSFFLPLLHSEPWFDLMSKVVNGIRIRQVRCLSDYFDFDWYQTS
EELDGYPLSATLEIYGGGYVVEILKQKKWIDRQTRALIIEFALFNAATNYFTMVTIVFEIPSFGGNLYTSLQIIFLLLYFHEFWNVIEFINFQFASYWDFFVTLKFIKLLRFNRRVSLLSSTLNVAFYPLSMLGITFVVALIAIISF
CTIVFGLGPLFFLMSMFISILNDSLVCVEMLDYFLTRFKAWFG
>260798008
YTVTILSTRCWKSDGCFCDHLVAPNSLNIYILMVLWARWKDKKDKVGATVLGTSPGGGCSVYQVVVFTGARANAGTTAKVSVIFYGSGPHELFCTGGVDSFVVTCAGTMGPLNYIHVWHDNSGPSWFLNKIIVTDVYT
ERKYYFLCNKWFAVGKVERVFPVASDRELKRFKTIFSSKTGTGFRDDHLWFSVLGRPSTFSRVQRISCCLSLLLCTMLANIMFFSWAEIMIAIESALLVFPNLAIVQIFRLPWWCVYIGWLVFSSCFVSAFFTILYGFEYGRQK
AEAWLFTFLTSFFFDQPIKILLLGMFFALIIKLLKARQQRFLELSLRSALMELSFFITYVFILILVANGNRDFSKITERQDFWNFMKDKVVPELYTSEWYDYSSYIVGPVQLRQLRCNGDYYTPGWYQSTDLTTATPHYGLH
GMYGGGYMLELLEMEQWIDSYTRAIFIELIVFNPNANLFSVVSLLAEFPGLGKRLYLYFQIIFVLVYFRLFWNLVEYSKYKQAASWDCAATLKLMHLCRFNKHVERSAHTLTRSAKPMLNYMIIFVIYVSAFCMLTYLVLGL
GPVIFLLMMFCAILDVAYHEVKITELAMERAGV---
>999987444
YSMKTFIGACWKGDNCLCNHLVAPNPINFYILLVIWARRADLKDKTGIIVLPDNDPRDSIAYEITVYTGFNQEAGTTADVYLQLLGSDPRVLLQRAGVD----------
RMQIRDCQTQNKWWFVCNKWLAVGKVERTIRTATNSDLTQFNVLFVNATRRNLFDGHLWLSVYTRPSTFTRVQRLSCCLSVLCCTMLTNAMFFSWSLIMIGVQSSVIIMPNVLLVTIFRFPYWFIYVAYLCFMSTLVSATFT
VFYGIMFGKDKSEQWLSAMVISLFQSQPAKVLMLATFFALIIKLQQARALRLKEKQMSAIIHEIVLHFLVVILVLFVAFANHDFHKITNYDKFWQWAESVLLPNMYPVDSRGKGTNIVGVGRLRQLRCNDEYFNPRWYRS
SWDLKGFPYFGKMGTYGGGYVVDLLWKDNWIDPYTRGIFTELNVYNVDINLLCVITLLYEFLPSTGKLFRYFEIIFLFLYFKSVWNVVDFVNFQYVVTLNLLCTLKFLRLLRFNRKVSMFSDTLRHSANTLFTFVVMIMI
VLXFCXLSNLXFGIGPLLFLVNLFIGILCDSFNTVEIVDFILGRVKSLVG
>919082289
YTMSLYSMACYSDEGCFCDHLVPMNTIRPYFVLLIWARKADTKDKVGVTPLPDNDARDKQGYEIMVWTGSRKDAGTTAHVSIILSGTEPRVLCQRGGIDSFLLTTTQSLGKLSHLRIWHDNSGPSWYFSRMLVEDLLTG
EKYYFMCNRWLAVGQIDRVIPVAGKNEMTNFDYLFWSKTKRNLSDGHIWFSVLKRPSKFTRVQRLTCCISLLFSTMATNIMYYSLHAIYIGVITSLIAFPNLIVVACFRLPWWFIFVGYLAFVTIVGGFYYCVEVAAVFGEK
LAMEWLVSFLTGIVESQPLKVLLFALFYALVIRVKAARDLRIKELEMHKIIREIVFYFVFLLLLMMCAYGNRDIWKVKSRKTFFDWAERTLIPALYNGEWYDKQNYLVGVARLRQLRCRSEYFLPEWFQSALDLKGFPYW
GLMTTYGGGYAFELLRSSRWLDIYTRVVFVEFVLYNANLNLFTVSQTTVEFPASGGRLDRYAEILTAAFYFKDTWNVVEFHNLQYSILWDFVANVKFLKILRFNKTMNLLGRVMKGAAKPLGYFFISAGIVFFAFVQLAY
IMFHFGPIFFLISMMLSIIMETYTSVEIIDFTIRRLKAFIG
>1005458745
YKLSSFFASCWEGSGCRCTHLVAPNPIDFYIIGVMWARWKDKKDKIGVTPLLSNDPRDTYGYEITLQTGFQQKRRTTADIFIQLIGSEPRMLFGRGEVDQFLLTVPQSLGRVREIHVWHNNAGPAWYLFRVLIRDLQTDI
KWWFVCDQWLAVGSVDRRLQVATKAELTKFNILFVNATRKNLLDGHLWLSVFTRPSTFTRVQRLSCCLSLMLSTMLANLMFYSLKQIIIGVQSSFVVLPNMVIVSIFRFPHWCIYVAYLVALSSLTCGVFCMFYGMTFGKE
KSEKWFTAMLVSFFQSQPFKVLLLAAFFALIIKLETARKLRLKQCEMKAIIKEIITHFVVVCIVLTVAYGNHD---------FLGWARNSLTPALFPLNAPGDVSTVLGVGRIRQLRCNEFYFGVEWYKKD---

DSFLYMGQISSYGGGYIIELVKDHGWVDRYTRALFTELSLYNVNVNLLGVVSLLFEDLPTGNRVFRYCEVILAFMFWRNLWNHDFVSFQYVVLLNMLLNLKFLRLLRFNRKISVLSSSIKASARKLVSFMVMFLVIFLAYCF
LVLLVFGIGPAVFIMNMFIGILCDTFEEVEILEFMTNRVKAFLG
>999981318
ATVMMFTSMCWLTDGCLCNHLVPPNAIDWYFICLVFARRADRRDKAR----------
FVRGAGTSAEVSIVMNGSDPRVLFRTGGVDAFLMTAPQSLGVLRNIRIWHNNAGPSWNLLRIMIQDLQSGKRWWFVCDEWLAVGKIERILYPATSNDLTKFNVLFATEIRKNLTDGHLWFSVVTRPSTFTRVQRLTCCLS
ILLTSMLANMMFYSLEQVSIGITSSLVVFPNLIIVQLFRLPYWFNYVAYIAFMSCAAATVVVVFYGITFGEEKTGKWLMSMTISLVQDQPIKIFILASMFALLIKLEIARQLRMKEKQMKSIIRELVWYLIFLLVILTISYGNRDF
TEVQKIDDYWNWMNETVIPSIYKEQWYTPTDYLLAVVRLRQLRCNDFYYLQNWYQTGSELKTFPYWGYKATYAGGYAKELLKSKDWIDRYTRAIFTELAIYNANTNFFCVITLLYEQLPTGAKLYRYCEITYXLFSLKN
IWSLSEFVSFNYTAFLDFLSFIKIIRLLRFNRRMSLLAQTLRVCIPRIAAFLVIYGSSFFAFSLFGCLVFGIGPFFFLINMFLSIINDSFAEVEIVDFMVHRLKENIG
>156406861
YSIRVYSSNCWVNNGCECNHLVAPNPIDFYLIAVIFARRADLRDKTGATPLPTNAPPGSTSYLLSIHTGVRQNAGTTADVSIVLTGSQPTVLFTRGADNNLVLTLPNSLGKLTHVRVWHDNDGPSWYLSRISVENMNTKER
WYFICERWLAVGRVDRILPIASKQELTHFNHLFVSKTVKELGDGHLWFSIVTRPSVFTRVQRLSCCLSLLYCTMITSCMFYDLRQMIIGLQSSLIVFPNFALVHVFRHPN-
WRYLAWTFVLSTLACSTFTVFYSLSWGPELANKWLLAFVMSFFQSQPIKVLCASIIFALIIKLLAARQKRLKEIQMSMILKEILAYFIFIAVLLLVAYGNRDTIIVGDGDSLWHWARETLIKNLYPGNLYDRVNYLVGISRLRLE
RCYGEFYEVEWSIAEQFSFGFPYWGKMNFYGGGYPADLLHSHNWVNSNTRGVFIEFAIYNANTNLFGVAFLYVEFLPTGGRLYSYMEVIMLLFYFKGFWNWLEFVSFQYALASDFFVTIRFIKLLRFNRRMALLGNTIAL
IAKPIGQFLISFTLVFISFAFFGNMIFGLGRIYFLMNMFLAIINDSFSEVEMVEFILDRFKSQIG
>1005492071
FNTTIFQLTCWQYDGCRCNHLVPPNKIDWYFLLILWARRADRKDKVGVAPLPDNDRRDTYMYEIVVYTGFGNHSGTSSSVRMVLVGTTPRRLFCRGNVDYFLLSVPQSLGKLKSLRIWHDNTGPSWYLLRIMVHDIQTD
HKTWFICDRWLAVGMVERTLTPASKEELTSFNILLFSTEARKNLTDNHLWFSVVARPSNFTRVQRLSCCLSVLLSTMLANAMFYSMKQLSVGITSSLVVLPNVLIVMFFRLPHYFVYISYLVFITCTVSATFTIFYGLTFGKEKS
EGWLSSMMISFWQDQPLKVFAAAMFFAVVIKLEAARLLKTKQKQMKTIIREVIQYLVFLSVLLVIAYGSRDIDEVADYESYWKWVQDTLIPTLMPRYWYHGTAVLVGTARLRQLRCNVPYFEPYWYQTQWKLKGTPYW
GYFATYGGGYIGDFLESQKWIDRYTRAIFAEFALYNAQTNFFSVITILTEILTTGGRLYRYCEITYMAFYFCDPWNYSEFVSFQYVAFVDFFSILKFLRLLKFNRRISLLIQTIKIAASPLMSFMLMFFILFLAFCQFAFLVFGLG
PIFFLMNVFIGILNDAFVEVEILDFIFHTFKKTVG
>828228329
YTLRIWETVCFVTKGCVCKMEVAPNPIDFYLILLFWARREDRKDRAGVTALKENDPSDIQFYEITVYTGFSGKGSTSANVFIQLSGGEVRQLFKKRGQDVFLASFPDHIGDLQYLRIWHDNAGPSWYLSQVMVHDLNTD
RKYLFINDNLIAVGSLERILPVAGPEDMTTFNHLFFSTSQKNISDNHIWFSVFIRPSRFTRVQRLSCCLTLLYCTMLTNAMFYNPRQIGIAVITSLIIFPNILIVGVFRLPWWMKYFSWLLVLVSFTSAFFVVLYGFTFGKEKSAQ
WLSTLFLTLFQDQPIKIFGLALFIALIVKIAQLRAYRLKEIKMYDMMRDFVIYVMFVIALCQVTYSFMDFRDIKSLNDVWPWLNSTVVNALFSPHLYDLSSHLFGGARLRQLRCKSWYYAPGWYQTSSVLQTYPFIGYLATY
GGGYVADLLHQNKWIDGRTRALFLEFSTYNPQVNLFGIVNLLLEFSPSSARLYTFCQIFVVFFYFRGFWNLYEFVSFNRIVQWDLLTCIKSIRILQYNKTISLFVLTLKKSASPLAAFFLIFAIFFTSFTAWAYLMFILGSLWFIM
NVFLTIVMETYASVEIVDFMIKKFLKLIS
>1005450057

VAPNPIDFYLIVAIYARREDLKDKAGVTPLEDNEPCDRYHYEITVYTGYGKRAATTAEVSFILSGSQPRLLFQRRGVDVFLATFPEGLGELNYLHIWHNNGGPSWYVSRVVVHDLNSDKKYLFICESWLAVGSVDKRIPV
AGHEEMTSFNHLFYSTSQKNLADGHLWFSVFIRPSRFTRLQRITCCLTLLYCSMITNAMFYSPAQIGIGIMSSLIIIPNILIVGIFRLPYWFLYIAWICFLTCFTTAFFCVLYGLQFGKEKSSQWISSMLISFFQDQPIKIFGIALIIA
VVVKLEEARKLRFKEMKMNAIIKEVALYVFFIACLCIVSYSHRDFSAIGSRETFYTYAKTTLMNGLYSAAWYDKISYIVGMARLRQVRCNGWYYDIGWFNTSDELDTLPFMGYVTTYGGGFVADLLESNAWIDPQTRAVI
IEIATYNPIVNLFCVMSLLVEFLPTNGRLFTFCELFIVLFYFREFWNYVEFVNFGQVASWNFVSSIKGIKLLRFNRRITILAQTLRGSAGPLAAFSVVFVVFFLAYSLFAFAVFGLGPLFFLMNMFLTIVLDVFSEVEIVDFMIRR
FKKFTG
>1191059136
YSMRIYLPACWVTNGCVCMPGPAPNPIDFYVVLAIYARREDKKDRAGVTPLEDNEPSDAYREITVYTGYGKKAGTTSQVSFILAGSEPRLLFQRRGVDIFLATFPESLGEINYLHIWHDNSGPSWYLSRVIVEDLSNDKK
YLFITTESWLAVGTVDRLIPVAGKDEITSFNHLFYSTTQTKLADGHLWFSIFMRPSRFTRLQRLSCCLTLLYCSMLANAMFYSPAQVGIGIMSSLVVVPNMFIVTVFRLPYWFLYVGWVCGMTCFICAFFIVLYGLQFGKEKSA
QWLSSMLVSFFQDQPIKVVAIALVFAAIIKLKKDREQRFREMKMNAMIKEVLLYVFFVTCLCIVSYSHRDFSSISTRENYYYEWAKSTLMDALFNEEWYDKMAFVIGGARIRQLRCSTWYYDIAWFYTSGELDTMPFIAKVSS
YGGGYAAELLENNTWIDAQTRAVFTELATYNAVSNLFCVMTLVVEFLPTNGRLFSTCEFFILIFYFKEFWNCVEYVSFVRVSSWSFTIWLKGMKLLRFNARIQVLSQTLKGAAGPLATFSVVFIVFFLAYALFAFAVFGIGPI
FFIMNMFLTIIMDVFAEVVIVEHMIKRFRRITG
>1139824403
YDVSGWTTGCWDKTSCDCQNPVPPNSIDFYVIGLLVGRRYDKKDKWKLFYMSDNLKTDKHVYLVTVFTGLSRNSGTTSNVSFVIASTGVRHLFETGSVFTYVMTVQDSLGDLTFLNIWHDNSGASWNLIKVVVEDAH
TGKRFVFFCNQWLSLSSIQCTVPVATSESLQTFNFNFNENTRENVTDSHLWLSMFIRPSTFTRCERLSCCVTLLFLTMISNAMFYALSSIYISFVGIIMTFPVLTISMIFKMPPKCLYVAYLVVLSILLSGFFVVLYSMEWGKAKSE
EWLKTFFMSFLESDPIKVIAVAVFLSIITKLENMKRKRLQELRVKSILVELLLYVCFLAILYFISYNNRDYLNVRSTEYFYDWLRNTLIPFVFPLTWYDQANYRLGPVRLRQLRCYADYNCQSWFTSATDIWGVPTIGLQGTY
GGGYVADLLYNNLWIDRQTRAVFVEFTQYNSNANLFTFVSLMTEFPQTGGRVYQHCEIVLMICFFRDTWRLFDFVNFGHIALWDFLSTIRILRVLNYSRSITHLAGVLSNARRNLLGCFFMFALIFIAYAGFGYLLFGIAEF
YYVMTMMCATLNRSIREIEIDDLLHDLLKDFWH
>871245685
YKLLVYTDSCWTNESCDCVDAIPPNLIDFFVFLALWARRKVKEDRWQVGYLSDNLAMDNYYYLLTVHTGLSRGSGTTSRVSFVLTGTSVRLLLSTGSVVTYVMATPYDLGELNYIRIWHDNSGQSWYLNKIYVEDLQTN
TRTVFFGNQWLAMGRLDRILPVSCKDDLMTFTRLFSEHSRVNLTDSHTWLSVILRPSMFTRVQRLATCLMTLHLAMIANALFYSLTTLYTSVVSGLITVPVLLVVWLFKLPPWCLYVAWILAITSLSCCFFLILYSMQWGKAV
SEEWLASFVLSFLESDPVKVVLLTCLLAFCLRVVKMRNLKLQEQVAVGVLVELVLYIIFMFCLFSFSYSHRDWSKVRTRQDVYTWLNTTMTTRMFPVSFLDMANYRVGPVRVRQVRFYWPYYCLGWHRSTLDENAIPVLG
KLHLYGGGYIIDLLYQNRWLDLQTRAVFVEFSLYNVDANLLATVRLVVEFPVYGSRIYPHCSIVLVITFLKSFWGMMDFVNFYALAIYDFLSTVRILRILGYNQRVTTIAAVLQRAGKDLFSFSFLYLLMIGVFVMLGYVLF
HVVRIYFLLTMFQAILNASATQIGIINVLKKMYDALVV
>961088331
----------GNIWPLCDNKINDTDFYLITVFTGLKADSGTLSNICFILCGSATRCLFPTGSLRKFLMSSSKNLGDLIYLRIWSDCSGSSWYLNKVEIMDTKSEKIFVFRCEGWIATSTIYNISC-
NRSNDHPIFKSRFVDHIQHSLSDGHLWFSTYVHPSTFSRTQRIVCCATVLYLTMITNAMFSNLRQIAVSIQSIAITVPTLFLVACFKLPFWCVYVGWWATLVIFTSGFFVILYSLEWGSTKSEYWLTSFAMSNLESDPGKILLLS
LAVAIFRSRKIRHRKIISEMWYKSVVLNGIIYSLLLISAGYIAYVNQDFTNVKTVKLYFDWMENVVFPNFFPTASIDGVNIKIGPARLRQLRCNKGYFCLQWYTPASKINGIPIIGMWNTYGGGYINNFLIRAKWVDRQTRA
VFLEFTLYNANTHMFAYMNFLVEFPESGGRPFGIFQAIFIFCFFSQFWNVIDFVNFYHIVIWNFFAIIRLSCILGVSQTMLQISLILNRCVSHLFGVSFLLCIYAGIYAMLGYLIFGLGKIYFLLKMFESVLGECISMKNVMLILS
KY----FR
>762108557
FSMAIVTSGCWTDNGCRCSDPVPPNKIDFYFLALLYLRRKDKQDKWGTRFLCDSENTDAYFYLLTVYTGLRRDGGTKSNVNFIVGGTGVRVLFPAGSLRKFLMGTKHPLGDLNYFRVWHDNSGRSWFLHKVLLDDLQ
TGTRFTFLCDKWLAVGMVERIISVTSHQTLTSYDQLISDHARFNATENHLWLSMIIRPSTFTRVQRLTCAYALLFLTMITSCMFFSLTTLYVSFISILIATVIFIVTFIFKLPHGMVYVGWILVLAIVSSGFFVILYSMEFGKTKSE
EWLSTFVFSFVESDPINVIVMAILMAILFRLESQRKLRQNEIAAQAVFRRLVVYAIFVFVVFSIGYINRDFESINSADKFFHWMRHTFLPRLFPVSDIGFSHARVGPALLRQSRCIDEYYCTGWHTSASEVWGMPMEGQRNTY
GGGYMIKFLERTLWIDRRTRAVFVEFTLYNPNVNLFFNGIYLLEFPDSAGRPFQACYFIYIIFFFKSTWNVVDFIQFQHIAFWDFIATIRILQILGYNRRMTQLIAVITLASKDLSGFAIVFSIIYSAYVVTGYLLFGLGELYYMI
SMFAVILNLSISVVGLTNVVADSVKNVVG
>762157119
FTLQIITTACWDSENCRCA--
IPLNLIDLYIVLAVILRREDMKDRWSVSFLSDHDADDSYFYKITVFTGLRRHAGTTSNIFFVLSGSRTRSLLSSGSIVTFVFGTRHNLGDLEYIRIWHDSSGQDWYLNKIEITDIQTDERVVFLCETWLSLGQIERILLPVSTIESL
TTFKTLLAQQAQINVTEHHLWLSLLLRPSRFTRVQRLSCLMALLFLSMISNAMFYSPTTLFVSLIGIVITTFILFITVLFKLPYWIVYVAWMLTIIVAVSAFFLLLYSMEWGSSKSEEWLSSFILSFLESDPIKVCFVAVLFATVFKI
DMKRAKRQQEKRAKKALTKLVISAVYLSALYSISYLERDFSSIKKIEDFINWLNKTFVPTMYPETNYDDASVRVGPAQLRQLRCIHGYYCEKWYTNAESIWGTPKTGEYGTYGGGYILRFMIEQNWIDRHTRAIFVEFTL
YNANSNLFVHAIYLTEFLETGGRPILNCYTTSILFFFLHPWSWVIFINFNHIIIWDFLSTLQLLNILGYNKRFTQIISVVVNAGSDIGGFGFIFLIIYVMFVILGTLLFGFAEFFYLITIFAAILNNSISHVGIKFYISKTIKDIFG
>405968482
FSMAITTTGCWTSDNCRCS--
VPLNTIDFYVIAALYLRRQDKKDKWATHFLSDTDGGDTNFYLLTVQTGMRKGAGTLSNVNFILSGTGIRTLFATASVNKYIMGTSEPLGNLTHLRIWHDNSGQSWYLNRVVVDDLHTQERYIFLCDKWLSVGMIIDRILP
VSSNDDITSFNNLFSDHAQSNISEQHLWLSMFIRPSIFSRVQRLSCILCMLFLTMITNAMFFSVTTIIVSIIGISITTPIMFITFAFRLPHWVRRIAWVVFMSVLSSAFFLLLYSMEWGIEKSEEWLSSFMLSFFESDPLKVLLISVVFA
LIFKVAIMRKKRQEERTAKAAMITLLVYAIYVSAIYSISYMERDFSKVSDTDSFFGWLKGSLLPTYFPQKNYDNVSIRIGPPRLRQVRCYDRYYCVGWFTSALDIWGIPKAGEYNTYGGGYIISLLLTEQWIDRQTRAVFLEF
TLYNPNVNLFAYVMFLTEFTEFGRRPVSSCYSIFIAYYLKHPWHYVDFINFANVVIWDFVSIIQFLRILGYSKKVTEILSVTTNVGRDLYGFLIMFSIMFFTFVGTGYLLFGVARMFYLFTIFAAILNKSITDVGVMDIVLKSAK
SALA
>Pdum12681
FAVQSAATGCWNSNSCHCEKPVAPNMIDFYIPLAVCCRRLDKRDEWGVTKLLDNKQGHSHYYLMSIYTGMRGGSATRSKIGFSINGTGFRLLLKRGSVKRFLLAVPKDLGKLNYVHIWHDNSGASWFLDRIVVQDLQD
DTGFVFVCDEWLAVGEIERILPAAGKDNLMSPKALFMDNVKLTATDDHLWLSLLTRPSNFTRLQRLSCVMTLLFMAMIANAMFYTLGQIFTSLMSILIITPNLLVINLFKFPWWTTIIGWLVVLIVLASSFFCILYSQQWGKE
KANEWIVAFFISAFQDDPVKVCMIAGFLALVFRITRLRRKKLKERRLYELIQEVIVYVLFLLVIWTISYSNRDFRQVRTPSDYLQWIKDQLIPWLFAYDNYDLDNYRLGPARLRQLRTYGAYHCIGWYTKATEVSGVPMWG
QFNIYGGGYVANLLWDNLWLDWKTRMVFLEFTLYNANANMFCSLTLMSEFPEFGGRVYVHCEILFICYFFGNGWHVIDYVNFQIVAMWDFVYTMKLIGLLNVNKHMGSLLSVLRVGKEELISYTVMFIVVFLAFSQLG
WLLFGGAQTYYLANTLICIIVDAINSVQLIDFIIYRVKEMFG
>1087817335

YEIYNFVTSCWNTNGCYCNHLVAPNPIDFYVVGLFYCRRQDKKDKWCVALLPDNKLCDEYFYQITIKTGTQRNSETQSEVFFVLGGTSVRRLFPRGSVKKFLMAVDDSLGALSYVRVFHNNTGASWYLDHVSVCDMQ
TNKNSLFFCDDWLAVGSIDRVLPVAGKQQISNFKHLFASSTRKDFSDGHLWFSLISRPSRFNRVQRLSSGLAVLLSTMLASAMFFTIQQLLISIQASFIVLPNILIVQMFRLPHWTIYFTWLVFIVSAVSASFVFMYSIDWGAAKS
VDWLTSIVLSFTQSQPAKVILTAIALSVFVNLREARLAREKHKIMLETIHEVIVYLLYLILLLMAANTAKDFDEVRTPGDFWSWCNNSLIPGLHASDWYNRETFVTGPARIRQLRCISTYFVEGWFQDSAKLNGHPYTGLLN
IYGGGYSVTLLKDHGWVDRPTSAIFVEFTVYNANVNLFASVVLLLEGSANGARLYEFTYIIFVLVYFHHFWNWIDFVNFQYLILWNFFATIKFTRLLRFNRRIRLLSMTLSYAAKPVVLFSVMFSIYFLGFLFLANLWFELGP
LFFLLNMFIVILDESIKRAEVVDYMKKQLVGYIH
>Pdum9951
FDFQALALDCWENDGCLCSHAVPPNKIDFYVILSIIARQIDKDDRWAALPLTDNTLHDSYHYEVTIYTGLRKGAGTRSKVFFVVTGTEVRALIERGSILRYILSVPGPLGPLEYVTVWHDNSGASWYLDSAVIKDLQTDEK
SVFLCGRWLAVGMIERTIPAATEEDLRAFGHLFISTTKHDMFDAHIWLSIVSRPSHFTRVQRLACGASLLYLTMIANAMFYAPSELCSE----------
LPHWVIYIAWLSLLSIFGSAFFVVLYSLQWGKDKSEEWMTTFFLSFFQSQPLKVLFIAFLITLIFRLDRMRARRATQLRAYAITKEIIVYCIYLLLISTMAYDNQD-
DRVHSMAEFWNWTETIFIPRMHADYWYDYNSFRVGAPRFRQIRCYDSSFAPGWHQSAFSLMTLPYVGRVGSYGGGYVAQLLKKHSWTDRLTQAVFVEFVLYNPNTNLFTAARLAFETPGSSGRLYPYLDTAVLIFYLKT
FWNIWEYVSFDAVALWGFTAMIKLLRLIRFSQLVHQLGALMKMVSAPLTSFFMVFFLMFMAFTHTAYLTFGVGPMFFLLNMMLSIINDTFLASEAAIFLVNQIKEVLM
>585688456

MISNCMFFSLKELYVSVVSSLIVLPSILLVQLFRLPHWCVYFAWLMFLSSAGSAFFTVLYSMEWGNEKSIQWLITFFLSFFESQPIKVIFLALFVALIWKMEKARRKRFKEVKMKGVLKEMFVYGTYIIILLIAYGARD-
KIDNGFGSYLEYIRGAVIPSLYADTTYNHTIYRIGPARLRQLRCEVDYYGASWFATAEELDGVSLYGQIAWYGGGYTAELLESTHWMDIRTRAVILELTVYNAQTDLFSMVTLISEFPSVGGRLYPYCDVIFICYYFKDYW
NYLEFLNLQSTAIYEFCSILKFMKLMRLNHHIALMIGTLGAARRDLANFCIVFGVTFIAFGQSGFLLFSLGPIFFLLNLFVTAVMLAYSVVEIVDFIWDNIKEIVG
>585644120
FTIDILTFTCWSSEGCSCSHLIPPNPIDFYLIVIIWARRADKKDNQGVMPLTDNNEHDEFLYKIEVFTDVRKEAHTKSNVFITVIGTCERQLFRRGGIVRFVMAASESLGTIQGIKIWHDNTGASWCLGKMKIEDVQTNDL
YFFTCNRWLAVGKIERFLKASDTETLGSARGAFVSKANQDLLDGHLWFSVIGRPSHFTRVQRASCCLSLLYTSMIANAMFYSWHQLYVGFMSSMVVFPNLAIVQLFRLPHWCLYIGWLVFLSSSGSAFFTILYSLEWGHKKA
SSWLVSLLVSFLQDQPIKVIILAIIITVIMKLMKAKELKKREETMWSFFKEIFVYCIFVITVLLIAYGQRD-
SLSSGIADFYIWVDEVLLPGLYPTEAYGGLSYRIGPPRFRQLRCNGGYYGYKWYNTMTELEGVPLYGIENLYGGGYVAVLLFDNRWLDERTRGLFVEFTLYNPHVNLFSMFTLLLEYPTTGGRLYNYCEVLFIIMYFVKF
WNLIEFINLQQMAAYDFFGTLQFINLMRFNARMALLSTTLKGAVKVLANYAIIFYIVFFAYGQLLFLTVGLAPLVFLTTLFVTVIVWQYEWADIVDFMYWKFMTLVG
>676479791
IDVRIYSVGCWSKKGCLTNHLVPPNTINFYFIIVLVVRRADRKDKWSVEPLSDNIDTDIYHYQVTVYTGIRRNAGTNSNIFIVLAGSEIRALFSTGSVSSYLMSVPRSLGDLTYLRIWHDSTGVDWFLNLVVVKDIQNGASYV
FLCDRWLSLGKTERFLPVASKEDLQNFAHLFFFKSRKDITDGHLWFSVVARPSPFTRVQRLSCCLSLLFTTMIANAMWYSAQELLISLFGSFIVVPNILIVTLFRLPHWCVYVAWLVLLSVFVSGSFTVLYSQEWGAERSGMW
LSAFFLSFLESQPLKVITVAIVVAVIFRLEKVREYRKKQSRLFGVFQEIIVYSFYVITIAILCYESKDLEGISSYENFWEWCDERLLPNLYRRNYYDGIHFRVGLPQLRQARCLDDYYEVGWYKGTLEVNGFPIAGDRATYGG
GYTAVLLKSQRWLDRQTRAVFVEVNLYNPSVNMFTVVNMMLEFPPTGSRLYRYLDLVFVAFYFQEFWNCIEFVNFQRIVLWDFFSMLKLLRLVRFNSKVTQLAGTLRRSVWPVFNFCIVFTIIFVAFAISGYIVFGIGPLYV
LINMFMSILNDTFTAVAMVNFMINEFKRLIG
>1087817330
YHILIATPGCWTGDGCVCTKLVPPNTINFYILMLIWTTRMDRRDKWSLKSTVDNRKEDRYLYQVTVFTGHSRNAATQSQVSCIIFGSKVRRLYTRSSVVHFLLREPKCLGPLTHLHIWHDNNGRSWYLDQVVVQDVRTG
KRFFFISDQWLAVGKVDRVLPVATPENMSTFHRLFVSTTTQKFTEDHLWISLFYRPSNFTRTQRLSCCAALLFMTMIANAMFFSPLQIWTSFCSTLMVVPTTIVMLLFKWPHWCVYIAWLTSVAVILSGFFSIMYSFEWGKDK
ANRWLSNMMLSFWQSQPIKVVLLSMFVATIVKLRKARKLREKEKAMSQIIWEVIAYFFYIMILLFLCHGNRDLNTVSSVEYFYRFVNQTLLPNLYPNLWYDEQSYRVGPARLRQIRCRGHFFAPGWHQTWMEMRGVPFV
GSLAVYGGGYSAELLYKDAWVDEQTRAVFCEFTVYNPNVNLFTSVLMVVEFPPVGSRLYQYAQVIFFLHLFNDPWNYFDFVNFQQVSTIDFFGHRFLKLLRFNRKIGMLTSVLKNSARQWPGFFVMALVFFIAFIHIGYL
VFCFGPIFYILNILIAIINESYGMTEIVDFMMQRFTKWSG
>260824333
VRVATFTSGCWRQTGCLCTHLVPPNTIDFYFIVVGLARRADRKDKWGVSPLVDNDIRHSYHYQLAVYTGMRKDAGTRSNIGFVLSSTGTRRLFDRGSVNNFLLSVPRCLGNLQYLRIWHDNSGASWFLDKIHITDLQTEK
RTYFLCGQWLAVGVVERLLPAAGEEDVLQFQNLFVSNSRKNMTDGHLWMSVLSRPSTFTRVQRASCCLLLLFTSMITSCMFYTYQELSISIQSTLVVLPNLLVVNIFRLPHWFLWVAYIVYGGVMLSGFFTILYSLEWGYEK
STQWLVTFLMTFFQDQPVKVLVMAALVSLIVKLRSARFVRIKELKVEDVLRTTALHLAFLSVIVIVCLYNRG-KEIHTADGLWMWAEEILLPNLYPDVTY--
TSFRVGTVRLRQLRCNVDYMSPGWYGDAVDLTSAAFYGDVHLYGGGYAVILLKKTGWLDKYTRAVFVEFSLYSPSSSLFIAVTVVIEFTTVGDRLHG-
CKALYCMFYFRSFWNCVEFVNFQSAAWWNFASIVKFLTLLRFNGRMAMLSYVIRRATNELVQFTVMFFVIFLAYCQVTYLLLGLSCVFLMVNMFLAILVDSIKHDELTDFMIHRLKTMFR
>1126215987
FKWIVFSPGCWSNEGCDCNHLVAPNTIDFYVLLLVYTRRADKNDKWGTLPLVDNAKNDRYFYKLDIYTGMRPGAGTKSRVNFILSGSGVRIVLRRGNIDHVLLGVPKSLGYLTFLRVWHDNSGQSWYCNKVVIQDIQT
GERYYFLCDRWLAVGLTDRILPVAGTENIIGFNHLFFSSSKRNLSDGHLWFSVASRPSNFTRTQRLTCCVALLFLTMIANAMFYSLTNIIISIESTLVVLPSMALVTFFKLPWWIFYVGWLAFFSILGPAFFVILYSLQWGKAKA
EGWLSCMMLSIFQSQPIKIVAIAMLVAIIFKLREIKETRRKEIMMWAIIKEFLAYTFFLCVLLLLANEHRD-
DKVSSVANFWKWAQGSLIPALYQTGEYDLNSYRIGPARFRQLRCYDEYFKPNWYQTSGETKGVPIIAEKSVYAGGYVVELLFKETWIDHNTRAVVTEFSVYNANINLFGCVILTAEFPPTGARLYSYAQVLWVIFYFKSF
WNWVEFVSFQRLATMDYVGVFKFMRLMRFNKRVCILLAIIKNSIPALGLFTIMFTICMTAFALSGLVIFGLGPVYFVMNMFVSILDETYGAVEMVEFMWMRFKRFTG
>1139813591
YTISFVTSGCWVTNGCLCNHLVPPNTINFYMIGLIWARFKDKRDKWGIAPLADNVVCDKFFYQLTVCTGMRRGAGTKSKISFVLAGTGVREILCSRGSINNFVLSTSEWLGPLTFIRIWHDNSGQGWYLSKVIVTDLQTSE
TYCFLCNRWLAVGMIDRLLPVASKEDMTNFQNLFMTKTKKSFTDSHLWISVFSRPSTFTRVQRLSCILSLVMTTMMVNAMFYSLSGIYISFVSSLIVLPNIIIDQLFRLPHSCVYFAYLVYISAFLSSFTIFLYSIEWGKEKSLSWL
SSILMSIFESQPFKIIIVAMLLSIILKLAEAREKRMKEIAMFDIVREIAFYFVFVALIVLVASHNRD-
NNRNHVPDFWKWLLEILLPRLYISKQWTRVAYRVGPVRIRQQRCSQEWYLEGWHRSLFDDAGLPFSGEFGVYGGGYVADLLMLNNWINQYTRAIFIEFTVYNPYVNMFSNVFIAFEMPTMGKRLFSYCELATLAIYFKC
FWNCLEFYNFRKVAMWDFVSIIKLLHLFRFNKRMSMIAGTLKYSTKELSAFSVVITILIVAFASWGYIMFGFGPIYFLLNMFVTILNESIAAVEIVAFIWCRFKDWVG
>585651646
YTLDIMSYGCWTGDGCLCNHLIPPNTIDFYCLLMVWARRKDKNDKWGVTPLIDNSPRDKYYYEIVVYTGMRQGSGTKSKVNFMLSGTGIRKLFERGSINHFVLATPRSLGPLTYLRIWHDDSGGSWYLSRIMISDLQTG
EKFYFLNDKWLAVGMIDRVLPVASKSELVNFNQLFFHSARQNLSEKHLWFSVVSRPSHFTRVQRLSCCLSLLYCSMIASCMFYTSYELYVSIMTNLIVFPNFIVVEIFRFPPWCRYIGWLCILCTCIASAFFVILYSMQWGRKKSE
EWLVSLILSFSLSQPVQVVVIAIFFSCFMRLEALRKERLNELKMQSILREISIYAVFVWYVH----------
HNLLLGPARIRQLRCSVDYFNPKWYHDAIELNGLPVFGYLSVYGGGYLAELLRNNSWYDIYTRAIIIEFTVYNANVNLFSLVTFLIELPTHGGRLYNYAQVLFVGFYFKSFWNCLEFVNFQHLAYWDFVAMLKFLKLLRF
NKRMSMLGMTLKYCFKDLLYFSVIWIIVFLAYCQVFYLVFGFSKVFFLMNMFLTILNQSIIAVEIVDFVWRRFTKFIG
>148539556
YNFSTYVSSCWKSDGCLCNHLVPPNKIDFYLLLLIYARRKDKQDKWGVTPLSDNHPSHTYGYQVTVETGVRPGAGTKSNVGIIISGTGARVLFQSGSINSFLMTTAGPLGLLAYVRIWHDNSGAGWFLSKVAVMDLQTG
KEFFFLCEQWLAVGDVDRTLFVASKKDLTQVGQLVFTSGRKNLSDQHLWFSLFSRPSSFNRCQRLCCCITLLYTSMIANAMFYTPQQISIGFISSLIVMPNLIIAQIFRLPHWTIYIAYLAFVACFISCFFSVLYSLEWGKEKAE
EWLKSMVVSFFSSQPAKVIVIVFIVSIIFKLQEARKERLRDMQVSGVIRDAFLYALFLMVVLYLAYINRDALSQSTPDEFWNYIEQTLPGLYGYKYYDWQSYRVGPARLRQLRCTLDYFAADWYQTAVDLYGVPYQSNLFI
YGGGFVAEFLRDHGWYDKYTRAIFIEFTVYNANVNLFSVVTLLVETPTFGGRLYQYFQGIFAVYELKNFWMILEYVNFNHIAKWDFLSILRFMRALRFSARVTHVAFALQRASKDMLYFLFMWCIIVIAFALWGYAVFSIG
PIFFLANMLITIIVEAFAKSPIGRHMMKRFKMRMA
>871242653
YTLLNLISGCWSSDGCFCNHLVAPNTIDFYIILLIWLRYMDKKDKWGAMPLDDNLPTDGYHYQVTVQTGVKKNAGTESQVRFIISGTGVRRLIPRGSILNYVMSVDAPLGPLTFLRIWHDNSGKSWYLDQIQVNDLQTG
EKMRKLK----------LPHWCIYIAWLCFFSIVASGFFTILYSMQWGKEKANEWLTTFLLSFFQSQPIKVLLLVAFIACILK---------L----------
GLAFSVLCAKFLVRNFRK-PLIKTVLDFWVWLEGSFLPSVYALQYADMESRRVGVVRLRQMR----------
ARHQDWLDLNTKAIFVEFTIYNPNANLFASVTALVEFLTTGSRLMSYMEVFFGLTYFKDFWSILEFVNFNSIALYDFLATIQFLKLLQFNKKMGMLGSTVKLASKDLKVFSITFFLYFFAFTATAFLLFGWGPIFFLMSIFLTI
IGDSFTTVEIVDFMWKKIKGLFK
>PdumPKD11
YFYNSFVSGCWSAAGCLCNHLVPPNTIDFYLGVIIWARRADKKDKWGAQPLEDNLPSDQYHYQITVYTGGNKNAETTSRVSLIISGTGVRRLLEKASINNYILSTEGCLGPLSFLRIWHDNTGRSWFLDQVQVKDLQTGE
LFCFLCDRWLAVGLVERIIPVAGLREMTAFKHLFNSSVKKKLANDHLWVSVVKRPSNFTRVQRVSACAALLFLTMITNCMFFTLHQLFISVVTTCIVFPSLLIITIFRLPHWCIYIAWLCITAILASAFFTVLYSMQWGRKKSLD
WLSTFVLSFFQSQPVKVLLLAAFLSCVMRLEEARQLRMKEMAMEAILKECIVYFFFILVVFFLSYQARD-
DSISKPDSVWTYLRNTVLPNMYSNNWYDRCTIRVGVARLRLQRCRNNYYLPGWYRNSIELKSAPYAGTHAMYGGGYTFTFMEQYEWVNPNARAMFIEFTLYNGNVNLFGSVILLVEWIAGGGRLNNYFEVLFLAFYFK
EFWNMLEYVNFVTIGVWDYLTTVKFIKMLRFNRRMGMLGDTIRIATKDLKNFIIMFFIYFFAFCMFAYVIFGIGVLFFLMGMFLTIIADAFTAVENVDFILGKIKGVFG
>919018042
YSWQLYTSGCWGTDGCLCTHLVPPNSIDFYFILLGWARYKDKKDKWEATPLEDNLPTDNYHYQITVTTGMRKKAGTTSKVSFILSGTGVRRLFERGSILNFILSVEHSLGPLTFLRIWHDNCGRSWFLEQVQLQDLQTG
EKFFFLCDRWLAVGMVERILPVAGLEDLSDFKHLFSSSVKKKLTNEHLWFSVVSRPSNFTRVQRISCCLSLLFCTMIANAMFFTLSQLYVSLISTLIVFPNLILVTVFRLPHWCMYFAWLCFCSVSCSAFFVILYSLEWGKEKAN
EWLTTFVLSFFQSQPIKVFLLAAFISCVLKLTAAREQRMKEIQMEAIIKEILIYFFFLLILFFLSYQTRS-
KDVHTIPQYWAWLNSVLLPGLYASSMYGFGANRVGVGRLRQMRCRDDYYYPLWYQSSLKLKNAPYMGVVNFYGGGYVQDLLKDDDWFDQYTRAIFVEFTLFNPNVNLFASVVILKEFLAAGGRLFPYFELMFGAFYF
REFWNLMEFVNFQSLALWDFLATIKSLKLLRFNRRMGMLGDTIRIATQDLKVFSITFFLYFFAFCQAAYLLFGLGPLFFLMGMFLTIINDSFATVEMVDFIVKRFKAAFG
>1205898338

YDYQLWTSGCFISNGCMTNHLVPPNTIDFYIILLIWARYKDKKDKWGASPLDDNLPTDNYHYQVTVQTGVRKLAATNSEVSFILSGTGVRRLFVRGSIMNYILSVESCLGPLNFMRIWHDNKGRSWYLDQVQITDLQTG
DRYFFLCDRWLAVGMVDRILPVAGLEDLAAFKQLFSSSVKKKLTNDHLWFSVFSRPSNFTRVQRISCCVSLLFMTMITNAMWFTLQQLFVSFASTMIVFPNFVLMTFFRLPHWCIYIAWLCFFSITSSGFFTILYSMEWGPEK
ANEWLTTFLMSFFQSQPIKVFVLVAFIACILKLGEQRQLRMMELKMEEVIREIGIYALFVLILFFLSYQTRD-
EKISTVTDYWNWLENKLIPNLYATHYADMESRRVGLPRLRQLRCRDEYYLPGWYQNSVKLKSAPYMGTISTYGGGYVALFLKKDVWLDVYTRGVFLEFAVYNPNINLFSSAILLTEFVSSGGRLLNYFQVIYACFYFGRF
WNLLEYVNFQSMAMYDFLATVQFLKLLQFNKKMNMLGDTIKVATKDLKVFSIAFALYFFAFTACGHLLFGLGPLFFLMSMFFTIIGDAFTEVEIVGFMWKKIKGVLG
>762107655
YSTQTFVPSCWKGDGCLCTHLVPPNTIDFYIMLLIWSRYMDKRDKWGATPLEDNLPSDNYHYQMTVQTGVRKNASTTSKVSFIVSGSGVRRLLESGSILNYICSVEKPLGSLTFMRIWHDNSGKSWFLDQVQITDLQTGD
RYIFLCDRWLALGMVDRILPVAGLEDLVAFKQLFSANVRKKLSNEHLWFSVFSRPSSFTRAQRISCCMSLLFMTMTNAMWFTVQQFFVSFASTMLVFPNFILITFFKLPHWCIYIAWLCFFSITSAGFFTILYSLEWGKEKAN
EWLTTFLLSFFQSQPIKVMFLVAFIACILKLEEARQQRLIELKMEEVLKELSSYIFFVIILFFLSYQQRD-
ESITTVDQYWVWLENKLIPSLYATHYHDMASRRVGIPRLRQLRCRDEYYLPRWYQSSVKLANAPYMATMSTYGGGYVYMFLKNQIWMDAHTRAIFLEFTVYNPNINLFGSMIMVIEFLHSGGRLLSYAEVIYTIFYFKGF
WNKLEYVNFQSLALYDFLATVQFLKLLQFNKKMGMLGETVQYATKDLKVFSIAFLLYFWTFTATGFLLFSLGPIFFLMSIFLTIIADAFSTVEFLDYFTKRLKQVMG
>Pdum688

LPRKCLPVGWISVLAILTSGFFVILYSMDWGAEKANAWLVSYTTSFFQSDPIKVIAIAAFTALLVWIEKQRKRAVLTQAAKDMVKSLVAHALLLWVLVSLSIMCSS-
SMATSKGGFFDWLHNEFIPNYFVEKDYDFQSHRVGPARLRQVTCYGTYFGIALYNTSLELDSLPIWGHLSVYGGGFAASLLNEYRWLDRQTRAVFLEFSLFNANVGMLTYITLLAEMPNSGAKVYSACAAIFLIYHFTKV
MTWFNFVSFAYVAALDSLAIIKLIFLLKLYRRISQLMLTLRKSIKPLASVLVALSIMVLSYAFYAYLYLSFTWFFFMLSILITVIMDSHYEFEVIDYFVKSVFR--G
>919052864

MVISLPDNDPLATYTYTVSVFTGSRRGSGTSANITITLIGSDPHLLGQRGAVDFFIIKTAASLGEVKKIRLWHDNSGPAWFLSRVVIHDLKADQKFYFICHSWLAVGQVDQVFSNATDDDIRNFTHLFISKTKIGLTDGHL
WFSIFARPSHFTCVQRLSCCLCLLMTTMLTNIMFYTWSQFIIGVESALFMFPSILIVELFRLPWWFIYIGWFVGLTTLSCGFFTILYGLSYGFRTSISWVVSMLTSFFTSQPVKVIGLAVFFSLVFKLKSAKQLRYQQKRMYAIIRE
VIGYIIYCWLLMLMAYGQRDFMQVKTFDDFWVYVKDTAITNLYDQ---
TKAAYLVGSPRLLQIRCASTYYNQSWYQTPAELQYMWQSGILGEYWGGYVADLLQDATWLNEGTRMVVLEFTAYNANVNQFTMISTLVEVSVGGGRLYRYCEIIYCIYYFISGWTWISFVNFEQTVLLDMLSSLKFISILR
FNPVVHLLTQTFAKAAGQILWAVALVVAYFIAYAFLSNLIFGVAPFILFLALFIAILTSEIGVEDLIRLGIEKVRELLG
>585652993
FDVYTYTTTECWLAHGCLCNHLVMPNTVDFYLFIAVWAYRKDRHDQAGVTVLEDNDPYALYRYDITVHTGFRRGAGTTANVTVTLYGSEPHVLLQRGTTDSFLMTTNESLGSLKSIRLWHDNVGPEWFVGHIMVHDL
ETDERWYFLCNTWLAVGLVDKFFNTATDHDLKQFNHLFVAKTAKDLRDRHLWFSIIGCPSHFTRLQRLSCCLSMLLCTMLANIMFYSWQQLIIGVQSGIIVLPNIIIVQTFR----------
AKKNLWTMRNMIAIWNSLSMLVLFVAVFVALFLKVHQSKEAKIRVKKMYALMREIFGYIFFIWLLLIAHAQRDSNNLHDYSTFFNWTSDRLMPGLYGD---DDVTRLVGGARIRQIRCNSSYYSESWFQES---
VSNPYWGQAATYGDGYVCVLVEDALWLDSHTRALFVEFTLYNANSNLFSIITLLIERPDIGGSLFRYCEIAYIGAYFMCAWNWLEFVSFHECALLVVLGTVKFLHLLRINPKVYLLTAVLASSGREMAIFTLMLMVIFTGFA
CMAYIVVGFAAFLSVLNIFISLLITTLKRVVIVEMMISKIMSICM
>1207182030
VYITTIAADCWSGEGCLCTHLVMPNVVDVYLLVVIWARRKDIQDKVKVTVLEDNDPLAEYRYLLSINTGQRRGASTSSQVTVTLQGSEPHHLFERGGLDLFLLTTPFSLGELQSIRLWHDNSGPAWYVNKVTVQDLENG
QKWHFLCNSWLSICTLDKVFPVATEMDLKRFSNLFFAKTAKDFRDGHIWFSVISRPSNFTRVQRVSCCFSLLLCTMLTSIMFWTWQQVMIGIQSSIIMFPNLLIVSIFRLPWWFVFVGWLVLATSGVSAYFTMLYGLTYGKER
SISWLISMVVSFFESQPVKVLGFAVFFALVLKIEKMRNNMIKEQKVFALIREILIYVGFLWMLLLVAYGQRDISETLNHKDIFIWANSSLLRNLYGQ---
DGNSKLVGNARIRQVRCHAPYYSPGWYQTQAKLRANPVWGGLALYGGGFVVDLLFDNTWLDVYTRAVFVEFTVYNANVNLFCIVTLIFETTGVGTHLYQSSEAIYYIFYFKSKWNLLEFVSFNETATADLLATIKLWH
LLRLNPKLHMITSTLQRAWTDISGFIVVMTIMFLAYSIACNLMYGLGAFLIVLNLFISVILVAFNEEEIVELMLMKVCSLLG
>116006951
VGITTFLSHCWDDSGCLCNHLVMSNAINIYVLVVIWARRKDAQDKVKVTVLEDNDPFAQYHYLVTVYTGHRRGAATSSKVTVTLYGREPHHLFERGAVDAFLLSTLFPLGELRSLRLWHDNSGPSWYVSRVLVYDLVM
DRKWYFLCNSWLSICVLDKVFPVATEQDRKQFSHLFFMKTSAGFQDGHIWYSIFSRCSSFTRVQRVSCCFSLLLCTMLTSIMFWTWQEVMIGLESSILMFPNLLIVQIFQLPWWCILVGWLVAATSGVAAFFTMLYGLHYGR
ASSLRWLISMAVSFVESQPLKVLGFAAFFALVLKIEKMKTTHLKEQKAFALIREILAYLGFLWMLLLVAYGQRDFSGVLGFREFFKWANTTLVSNLYGH---
DGNSKLVGSAQIRQVRCRAPYYGEGWYQSQDQRQGYPIWGKLTVYGGGYVVPLLFDNTWLDALTRAVFVESTVYNANVNLFCIVTLTLETSALGTRLYPFAELIYFLFYFCSKWNLLEGISFSETAAADLLSTVKLWHLL
RLNPKMNMITAALRRAWGDISGFMIVILTMLLAYSIASNLIFGLGSFLIVLNLFISVILVAFSEEEIVDLLLMKILSFLG
>942189627
VPVTTFTSHCWSGNGCLCNHLVMPNVIDVYILAVIWARRKDIQDKVKITVLEDNDPFAEYRYLITVFTGHRRGAATTSKVTITLYCSDPHHLFERGGSDMFLLTTLFPLGEMQSIRLWHDNSGPSWYVNRVIIHDMEVDQ
KWYFLCNSWLSICVLDRVFPVATEAELKQFSNLFFMKTVRDFRDGHLWYSVLSRPSSFTRVQRVSCCFSLLLCVMLTSIMFWTWQEVMIGFESSLLMFPNLLIVQIFRLPWWFVFIGWLVAATSGVAAFFTMLYGLHYGKDS
SIKWLISMAISFFESQPLKVLGFAAFFALVLKIERMKKNYIKEQKVFELIREILAYLGFLWMLLLVAYGQRDFHSIHSYRDFFTWTNTTLISNLYGY---
DGNSKLVGSARIRQVRCRATYYGERWYQSQAKLRGHAIWGKLAMYGGGYVADLLFNTIWLDVYSRAIFVEFTVYNANVNLFCIVTLMLETNALGARLYPYAEVIYFLFYFHSKWNLLEFASFYETATADLLATVKLWHL
LRLNPKLNMITSTLRRAWGDISGFITVIVIMFLAYSVASNIMFGLGSFLIVLNLFISVILVAFSEEEIVDLLMLMKLCSFFG
>918996724
LLVTVHSSACWSSEGCMCHHMILPNAIDIYLAVVIWARKKDREDKAGVTVLEDNDPENKYGYLVTVITGWRTGAGTTANVGCFIKGSDRHVLFESGGVDSFLLTSAQCLGDLLSVTIWHDNSGPAWFLEKLMVRDLQR
NEVFEFLCNRWLALFCIQRTLRPASVEEMKEFSYLFSSNSGFSMREKHLWVSVFSCPCPFTRTQRLTCALTLLLTTMLTNIMFYSLSDIIIGIESGMLMFPNLVIVQLFRLPWWCIYIAWVAVLTCLVCSYYTMLYGLKYGHKK
SVQWVVAFFTGFFQSQPIKVILIALLVSLILKVEHMRGIRMKERRMYQILVELALYVTFTLALVFVTYGHRDFEEVSDVDSMWEYLEETFVPSLFTSTWYDMVSLHVGAARLRQLRCTVRYYNTSWYQSAFKLRSLPFAGK
HGTYGGGYVKLLLQEFKWIDQRTRAVFLEFTLYNPQVNLFSVAQLLFEFHPTGAKFYQYCEVMFILFYFSDPWSYIEFVNFFPAAFWDVLTVLKFFKLLRFNVRMNMLSATLKKSWKPTLHFYLVLAFLMMAYAFFGQC
AFGIGIVFLLINMFIAIIDESFRYVEMVDFMWTRLLVILG
>148539638
YRIRPFTSKCWGTDGCFCTHLVPMNSIDLYLTVVLWAYKKDQRDKAGATPLIDNDPRDHYSYEITVFTGMRSGASTTAKVSLIITGSRPRLLFQRGGIDSFIMSAPRHLGNLLHLRIWHDNSGPSWALSRLVIKDLETSKMY
FFMCDRWLAVGQIERVIPIAGKSELTGFGYLFYSKTRRNLTDGHLWFSVFARPSSFTRMMRCTCCLSLLFSTMMSNIFFYSVGEICVGVISSLMVMPNLIIVQMFRLPWGCQFIGWVAVGTVGVSFWLTIEVAGQFGPE
KAAEWLSSMACSLVQDQPIKVLCLATFYALVIKLNEARETRFKEIKMQAILKEIAVYMFYLYILLLIGYGNKDLSSVSSRQLFWKHMREVFVPSLYAGNHYDRAHVLLGTARLRQVRCIDEYYDEGWYRDWLELDSYPAF
AHQNSYGGGFTVEFLQQNLWVDQQTRAIFLEYVVYNPVTNTHVVSLNVLEFPPTGGTLDHYGAEVIFALLYFKDLWNIYEFHNFQFLAYWDFMAVIKFTRLLRFNIKILMLSETVNEFAYELSLFMVMFAVVYMGYASFC
FLVFYMGPIFFLINMLLAIICEAFAKVEIVDFMVTRFKAWSG
>1005483664
YSVLVKAIGCWKRDGCQCNHLVPPNSLNLYLLVLVWARREDRKDKVGASVSPSNKAEDTYAYEVIISTGMRRHAGTTANIAMTMTGSHPFLLLARGSVNSFLVTTSESLGEISYVRVWHDNFGPAWYLKQICIREIAVNR
VWHFICERWLAIGLIDRILTLTSHEEMKEFRHLFLTKTYKDLTDSHLWFSVVWRPSPFTRVQRVSCCFSLLLCTMMANAMWYSWEQVFIGVCSSLIVFPNLILVQIFRLPWWFVYVGWLVIVTSLVAGSVTLLYGIEFGLQKS
SRWLLSMFFSVTQDQPLKVVGLAIFFALIFKLQLAREKAAKERAMHAILLDFGTYILFTLIVLLIAYGIRDFSEILRPRDMWDWTEQALLPSLYNLPSYGEIGLVVGVARLRQHRCVDDYYTEGWYSTADELKGYAYKGEV
KTYGGGYVLELARSKLWMDRFTRSVFTEFNLYNPNINLFSAVTLLLERPPTGGRLYRYCEVLFLVFYIMESSNLLEFISFQYLSAWQFVTCLKFIKLFRFNRRIFLLSCTLRQAGGELLPYFIVFMINFLAFAQVVYHILLALTAV
IFFLNVLIVIIIDSFAVVELLEYVTERIRGVLG
>999981301
YNITLNVVSCWSTYGCLCNHLVPPNKLDIYILVVVWARREDKKDKVGVCVSPSNKPENNYIYEVTVTTGIRRGAGTTAXTAIIMYGSDPYLLLQRGVSHTFLLTTKEFLGQLVHIRIWHDNRGPALFLKQISVRDVQSKK
VFHFLCNRWFAVGLIDRVLPVTSDEEMVQFTQMFLSKTKKDLTDSHLWFSVVWRPSVFTRVQRVSCCLSLLLCTMMANAMWYSWEQVSIGIMSSLIVFPNVALVQIFRLPYWFRYIGWGVGTTSLAATVVTLLYGIQFGK
RTSEKWLLSMFVSVTQDQPVKVIAIAVAFAFIIKLEMARKRLQQEKQMGVILKDIISYLTFMLIMLLIAYGQRDFDKAHKIADMWKWLENLVLPVLYNVPWYGDSGMVVGVARLRQNRCRAPYYDPGWYNSARKLTGF
PFKGNVEXYGGGYVLTMAKAELWVDVYSRAIFVEFNVYNPSSNLFTACTVLMEVLPTGGRLYRYMEVFFLMFYFKEMWNWLEFTNFHYISAWQFAGXLKMVKLLSFNKRIYLFAHTLRHALRDIINFCFIFIIVFIAFSQL
YFSMLGLGAISFLLNILIVLIIASFQQVELIDFVVEQFKSLTG
>156395095
YNISLQVVGCWSTRGCLCNHLVPPNTLDLYAIGVVWARRQDKNDKVGVHFSPSNRPGSRYLYEIVVCVGTRRRAGTTANVAIKLYGSQPFLLLSRGSNNSFIVTTSEALGQLSYIRVWHDNTGPALFVKKVIVRDHVTGD
TWHFLCNRWFAVGMIDRVVPMTSQEELKHFKTIFMDKTKKDLTDSHLWFSIVWRPSNFTRVERVSCCLSFLLCTMMANAMWYSWEQVSIGIMSSMVVFPNLVIVQIFRLPWWMTYIGWLVTLTVLASTVVTLMYGIQFG
KHTSEKWLLSMFISVTHDQPLKVILLALFFAYVLKLEKARDRNDKERKMSVILTDIGTYLTFLWVLLVIAHSRVDFQKISNDSEIWLWTNHVLLPALYNLDWYGGAGLVVGFARWRQLRCTDPYYNSRWFNSASKLNGY
PYFGEISTYGGGYIVPFARNNSWIDRYTAGLFTEFTLYNPYSNLFTASTVLVEVLPTGGRLYSLCDVVFLLFYFRDYWNWAEYTNFQYISAWQFISTIKFIKLLRFNKKMFLLAYTLSRSAKDLAHYGIMFLIVFLSFAQFYYL
LLGMGAFGFLLNMFIAIIVESFTEVEIMEFTVERVREVLG
>1229159547

YALTSSLYSCWSTDGCFCDHLIPINHIDLYLLVIIWARKADKRDKAGIAPLPDNDPHDRFLYEVTVYTGLKRNAGTTANVTLLMSGSVPRVLFQAGGVDSFLLATPVSLGNLMHVRIWHDNKGPAWYLARIMIKDLCLDR
TFYFMANRWLAVGMIERVIPVAGRSELTSFGHLFFSKTRRNLSDAHIWFSVFARPSKFTRVQRATCCLSLLFCTMMANIFFYSFGQVVTGVISSLMVLPNLMVVQLFRLPWQFCCLAWLVLTSVAVAFWLTIEVAGQFGPDK
AMEWLTSMCFSLIQDQPFKVLLLALFFSLVIRLAEARDIRVKEIKMHKIIREIVFYLIFLHLTMLICFGDRDFTKIRGRDHFWNYTENVFFPSLYSYDWYDNSMHLVGRVRFRQLRCRAPYYDAGWYNHWYDVSSFPYYGK
HAFYGGGYIVEFLQQDMWVDQHTRAVFLEFVVYNANVNMYSTSFLLVEFLPQGGRLDRYGQIVYTLFYFDSVWNLCEFLNFHFVAYWDFLVNVKFIRLLRFNRRLLLFSMTLRQCGKDLAYYGVVFCLVFGAYALFAY
AIFHLGPTFFMINMFLSIINETFGKVEIVDFMLQRFKRWSS
>1005475249
YTFDVRSFACWSSKGCQCTHLIPPNSLNIYILGLLWARRKDRKDKVGLVPVLDNDPSDRYRYEVTVYTARGRNTGTTANVSILIAGSAVHMLFLQGSVDSFLITTQECLGGLTFIRVWHDNSGPSWFLSQVVIRDIETDDK
FFFLCKSWLACGQVDRLFPAAGKEELRDFVHLFKTKTSRDFSDGHLWFSIAFRPSHFTCVQRVSCCLSLLMCSMLTNVLWHTWHEIIIGLQSSLIVFPNLLIVQLFR----------
ARQDSWDDDFFASFLETEEPDFLSSDIYYYLYQTLHAARVRLFKERKMHLVIREIVLYLLFVWMVLLIAYGHRDLVEVTNTVDFWDWMIQTLVPGLYNDAWYDKMSYVVGIARIRQLRCNEPYYAPGWYKTTADLKTR
PIMGHLALYGGGYNTDLLKSHQWVDKYTRAVFVEFTVYNVHSNLYCVANLLLEFTAAGGRIDRYCEVTFVLFYLAEFWSWIEFVDFHHLALVDFLTSVKFLRLFRFNRRMSLLGSTISASARELFHFGIIFGLVFVGFSHLC
YLVFSLGPIMFLVNMFLSIIIENFKRVEIIDFMIDQFSNWLG
>156394387
YTFGAYSSCCWSSQGCKCSHLIPPNAIDLYLVALIWARRRDLKDKVGLVAVPDSDISDLYRYEVTVYTGRGKAAATTSNVTLIMCGSQAVTLFQAGGIDSFFVTTPNPLGALCYIRIWHDNTGPSWYLNQVAVRNIDSNER
FYFLCYRWFACGEVDRVLPVAGKEELAQFMYLFRSKATRDFSDAHLWFSIALRPSLFTRTQRVTCCLSLLLCTMLANVLWYTWHELLIGIQSSLLVFPNLLIIQIFRLPWWFVYVGWLCLLTACTAAFFTMLYGLTFDKPKQ
HAWLISMFFSLVQDQPLKVLGFALVFALVLKLQRARLQRFKERKMFEVIQEILFYALFVWVVLVVAYGHRDLFQVSDITSFWNWTEAVLIPGLYSESWYDHTSWMLGTARMRQLRCNGPYFRPGWYRDSSELRSLPYLG
KLALYGGGYAFDLLKRDKWVDRSTRAIFVEFTVYSPYTNLHCVGYLLLEVLPSGGRIDRYCEILFVLFHFHEYWSWVEFVNFHALAASDFFVSLKFLRLFRFNKRMSLLATTLKCSAKELAHFSIVFGLVFMAFVHFCYLIF
SLAPFMFLINMFISIIIENFQRVEIVDFMTEQFRDWLG
>734563164
YSIRVFDKGCFNNYGCYTNHLNPHISEFL----------
ENPGRIRYMKDNMYHADYKYLVAVETGYRMFATTDSKVLMNVIGEVARELFRWGTTTDRFLMTTVWPLGALQTIRIWTDESGQSWYCSRIIFKDLQTNEIFEFRVNKWFGSGKTQRLVHVTKKRN---
CNVLVDALSFHSLADHITYWNMFTGGLRFDRVERVLNVYLAVLLAMLTNILVASWKDVVLGFVFSWVCIPSILIPYIHSLRWWHRL-
AHIELAIAIAIAICIVQLAFGYYVEQTMAFSRRFIITLLFWEPLKAFICSLFCTVACKVSTTRDRRMRDEQLFVTTREVVFVFSVAICLGFAFSSQD----------
NGVSRMMGYGIVRQVRCDGYTYNVGWYHSEEALQGRPFYGDHDSYGGGYIQTLLENESWINEGTRAVFIEFSVYNAQTNHFSVIQLVLEIPPFGTRLLKNFEVLYILLYFSSFWNYVDYIRLDRQRDLQFLTCMKMINIL
RFNRRIGLFTQTLSHSAPTMVIFGAVFCIVSLAFDSALFVILAFGAAVYLLNFFVMLVMTEFETLEALDHIASKALKVCK
>392891400
YQVAAVSKGCFNSEGCSTDHLNPTIDADFYGCLTINAIICQRKDRGRLRFLKDNEPHDGYMYVIAVETGYRMFATTDSTICFNLSGQIFRSFFSWGTTDRFVMTTAFPLGELEYMRLWLDDAGESWYCNRIIVKDLQTQ
DIYYFPFNNWLGTGETERLARVEYKRR-----FLDESMSMHMLAQTISWFAMFTGGDRVSRQDYSVSIIFSLVVVSMISITILTIKDIAFGVGFGV------
LITFLNSSGSWPMFMAGIIVFPVLMGLIYISGAGMSLMDDLANSFYIRFLISLILWEPIKGLIWAFLILKTRKLRDTENRKMRDEQLFTTIRDMLCFFASLYIMVMLTYYCKDFMSIQHADDFWDWARESLATALL-
ASWYDKVSRSMGIGTIRQVRCGEELMQAGWYKTSEELSTETVSGLLYSYGGGYTISMLDSERWIDDHTRAVIIEFSAYNAQINYFSVVQLLVEIPKSGIRLIKSYEMLYIFFGNFSPWNFMDYINLTEQRNWEFFTSCKMIRI
LRFNRRIGVLAATLDNALGAIVSFGIAFLFFSMTFNSVLYAVLGFAFVVIVLQLYVTIIMFEFEEIEIIDHIIKYKTKRRLG
>676439278
YNLVAGTDGCWDDTSCDCEDPSPPNTIDFYILLLLWARRADKRDKWQVGYLKDNHKDDEYFYLLTVETGLRKHSGTRSNVYFILSSSGVRVLLETGSVMTFIMTSDIQLDELQFLRIWHDNSGADWLFNKMIVEDLET
RKRSIFVCNRWLSVGLIDRIVPISAKKD-
LTFTHLFSEHARNNATDSHILWLSVVFRPSTFTRAQRVSCCICILFLTMITSAMFFSAYTIWVSFVSILITVPSMVIVYLFRLPFWCLYISWLLTVTILVCGFLLILYSMQWGKSVSEEWVTSFFLSFFENDPLKIAVFAIVFSVLMKL
VRARRRQEQETHAYSVLIEIVIYTGFIIALCVIAYGTRDFDKIQNRTHFMAWTEKNLFPFAFATYDFDLVSFRVGPVRIRQIRLSPEYYYCLYWYTSSNQDWGLPLIGKISLYASGYFFDMLYNNLWIDRQTRAVSVEFSLFNA
DANLLMVVQLIAEFPEYGSRIFQHCEIIVAIVYFQDFWHIMDFVNFEHIVWFDFFTTIGMMRVLGYNKKITIIALVLKRSAKNLGSFGVLFILMFFGFISSGYILFHLGALFFLLTMFQAILNSGITTVDVLMNMYRGILVDF-
>1207165955
YTVRIESVECWKPDGFRYNHLSITNNYSSHIMVSVICRKADVSLTTGSFLLPENNPSDRFLYAVTIDTGFRSRTRMTAKVYIVLHGSQTREVFTNNSRNTFILSSPESLGQVWKVHLWHDNGGPSWYLSHVVVKDLVQGSC
WFFLGQCWLAVGRVERNLVASEG-G-
LTFKQLLYLKLTEYLEDFHPWFSVYSRPSSHTHTQRLSVCLLLLQGYMCVNTLLIPAVSMLTGVCSALAVMPGGLVSLVFRLGAWCRYLGWLCVSLCLTCVTGLNFIHSRFNKTETQLWAHSVFFSLICGHPAMVNHLDPSLR
R----LKLVRGQIKKQTILLYKTTRDLMLHLIMLWMIIFVAYGKSDFHSIQKHDDWWNWTTTSLLEKLYNNADNGAAFSLIGSPVITKIKCFDNQ------SATKK--
LAPFCGKLGCYGERVYVSLLLETGWMTPFTQIVILQFVLYNGPSNLFTSVTILVEKTSTGARLYYFCEVTVIICFLKKEDFKLLLPICQLPIYHLNLLLLKCLHLLRLNKTVAAAVVTWKLIFSNLIWLLAPGVIIALALSNLCN
LILHVPPTFFVKAVLIATLAIFVKQAELTSYVKA--LF--T
>19923084
YTVHFQWIRCWKSERCSYHRLKLKASFEVYGFLVAKSRQVDHHEKAGYIFLQEASLPGHQLYAVVIDTGFRAPARLTSKVYIVLCGSETKELFERNSRHTFILSAPAQLGLLRKIRLWHDSRGPGWFISHVMVKELHTGQ
GWFFPAQCWLSAGRVERELTCLQG-G-LGFRKLFYCKFTEYLEDFHVWLSVYSRPSRYLHTPRLTVSFSLLCVYACLTALVA-
LGSFRVGLLCTLLASPAQLLSLLFRLQPWWSSAVWICGTASLACSLGTGFLAYRFGQEQCVQWLHLLSLSVVCCQPLMVCLMALGFAWKRRLRGTRQRMRRESRTRAALRDISMDILMLLLLLCVIYGRFSLGGLRNIADW
WDWSLTTLLDGLYPGGTPGGKCYLIGSSVIRQLKCSPEV-----IDPENQ--NVTLNGPGGCGED-
CVLSLLRASMWIDRSTRAVSVHFTLYNPPTQLFTSVSLRVEILPTGSSIFRSPQLVFLALYWRKPRNWLEFMDLTLMASWNFLFTLKCVYLPGIQNTMASCSSMMRHSLPSIFVAGLVGALMLAALSHLHRFLLSAMACYFCFG
MLRGFLM--TLPQDVTAYMWEKVLTFLR
>1039738876
YTVHFQWIQCWRSEGCSYHHLKLNATLEVYGILVSKSRYVDCHENPGFIFLEEDTLPGYQLYAVVIDTGFRSPVRFTSKVFIVLCGSETKELFGRNSRHTFILSIPNQLGPLQKIRLWHDSSGPCWFISHVMVKELCSGQAW
FFSAQCWLAVGHVLREFFCLSH-G-
LGFWKLFYSKFTEYLEDFHIWLSLYSQPRSYLHTQRLAVSFCLLCVYSCLTALVTTLEPFCMALLCTLLACPAQLLSLLFRCLLWSSSVAWISGSASLACGLGTGFLGYWFVPAQCMWWLYLLLLSLVCCQPLMICLAALVFA
WKRKLRVTGERLRRESIMQAALRDMTTHSIMLLLLLFIAYGRFCLGDLSSTEDWWDWTLSTLLDELYPERT-GGQCHLIGPPVVKLLKHSND------LDLENQ--
NVSPGGPETCGKESYMHSLLRASKWIDHSTRAMSVHFTLYNPPTQLFTSVILGTECLPSGGRIFYSSELLFLVLYWRKPRHWLELVDIGLMVSWHFLWMLKYVHLLSSLSTMTPFSAVT---
CFPLFRVLLVGALLLAAHYHSRWFLLFVMRCYYGFRMLRATFL--TVFQDIAVYTWHKVLTLLG
>1238871302
YTLDLSWQQCWSFKTCRCSLIVPTSVVMHFVVLMVILLTVCNPSGSHFFILNKNDSPTSQQYLVTIATGSFWEAGTTSQVCLVLHGSETKQLFLRGSLVSFVVSTAKPLGQLWKVQVWHNNSGPSWYLKDITVTHLNTNE
TFYFTCEQWLSVGQVEREIPVCEQ-S-
PSFWEKFYQQLMENLNNHHMWNSIFYHSSTYTSVERLACCLSEVLTGAFLTALWYSPEQVFSGFIIAVAMVIHQCVAVLFHLPSWVHYVGWIILVSVFSSLYYTPVYTLWLNPLQVLCWLQTYVISLLCAYPVWVCLLSLLKS
IVEILIESRLQAQKARKFCFAFRGILGFVTILLILVVVAYGRDLLQHVKTWANWWDWIQHDFLSSFYQNVYKPTGSVLVEQPRLRLFQ-G--------
VNHNSSWGTRKLWGQHSTYGGSDIANLLRNSSWLNRNTRAVSVEFLLYNPVFSLFTSVVLLTEFPPSGGQLHNYCEFLLFLVYWLDPWNVTKLPTFIQLAEWNFHLIASISLLELDTTFKKSLFALRTVSKQLFLLGGLYL
FLAIGYSLLSPLVSGIMLL-LTFSLIRVILIAAKRTTEVPQLFKNLMWSLLG
>919066838
YALRMWWGDCWSSEGCSCNHLPVSSDLTVYIMLLILCCRFDTG-
RREVICLEDNSPTDKQRYEIMIETGSLKGAGTTAKVSIILHGSETRELFKPNSRHTFLATLPEGLGKLWKVQLWHNNAGPSWYLSTVIVKDLNTGQKYYFICERWFAVGRIEREIPALEG-K-
PGFQRVFWTKGWQYFCDYHIWSSIITRPSRFSRAERLTCCFTMLMAYMCLDCLWFSWKAIAVGAICSAIVGPYVLLAFLYRCPLCFRYIVWLCLVIASGCGAVTVMYGLNFGHVKSVMWLQSFYFSLMTSQPVVILLAVLITA
CKYRLIKARRRNMKEKKAYAILWDILSFAFVFLLLIFMAFGKFSLEKVKHLSDWWQWGKTDMLDSLYGTSSNPGNSYIIGYAILRQLRCRHDY-----YSSL---
QYSSIMGQHAWYGNGFIVRLLQTAGWTSRQTRALILEFTTFHAPTSFFSSVKVICEFPPTGQHLYKYCELLFIIVFFCHPWNMVEFVDLTFVAFWDFLMLTQFLRLLRYSRLYAKIGNVYKRARWDLLTFFLAFVIWVMAYS
STGVALFNLAAIFLLTAYVIAILTVHYRAYEWFGFFWEKILSCTG
>443717936
YRLQLWQGQCWSSQGCSCNHLLSPNTLDFYGILMVIAVQKDNQDQPPIVELMDNSPSDKQKYLVSLETGFRKGAATTAKVSLVLHGSETRELFQRHSRQSFILTLPENLGRLFKCQIWHNNAGPSWYLSRVIVQDLNTG
YFYYFMCERWFAAGKVEREIMASEG-G-LGFSKVLWAKGTQYFADYHLWFSTFSRPSRFTRAQRLTCCLSLLMSYMALNCVWYSWRAVVVGAIISLIALPNLIIVMLFR----------LRGRTLPWMTA----------
AKVLLLSVIVAMRNRLLKSRNDILKQKGASAFSSETLLFTVQFVLLVLIAFGKNPLAKVHSISQMWKWMQGHFTDTMYDDPTRLSNSYLIGQVQLRQHRCRGDM-----
YHSAELTGAGRVQGQHDDYGDGYLVKLLEEDGWVNSMTSAVIAEFTTFHVPSNIFNTVRILFEIPAFGGYLYKYCELLLIVCYFHNVWNALEFMDVSDVHEWDFLVTLQLVKALRHQPPFSVFAKVYHRAAMDILVFA
LMFFIWLASYTCLTMALYSVWRLVILTAYMLAILNYHYKRHESVRFCRNRLSNVFR
>PdumPKD1L1
YRLWLWWGDCWSPDGCNCNHLLAPNDLSFYFILLAVSIKLDRHDKGSIVHLVDNRNVDKQTYMVTLETGFRKGAGTTAKVSLVLHGSEMRELFERNSRDTFIVTLKESLGKVWKVQLWHNNAGPSWYLSRVTVKDL
TAGHTYYFSCEKWFAVGKIEREIMLQDG-G-
LGFGKVLWAKGTQYFADFHLWLSVFSRPSRFTRAQRLTCALSLLLSYMAVNAAWYSWRSVFVGIITALLVLPNFIIVLIFRLPSWCIYIAWLSVVIILGCATITILYGFRFGRARSIMWLQSLFFSFIQCHPFIILIISVVTAIRYHLI
EARKTTLKQRKTGAVFSETLSFSILLFLLLVVAFGKDSFKDIQTVDQWWQWSHNDLLDLVYADGLGVGNSQLVGQVSIRQKRCHGGY-----
YKEKDQTGASCVVGRFAYYSAGHQVFLLYNYEWLGHQTRAVIVEFVLFHAPSNLFSSVHLVLEIPPLGRYVYKYCELLLIIFYIKSPWNCVQFVDTSFIAYWDFLLLVMSLRLLRNQPLFSKYGAIYKKTVHEIITFGCLFLV
WVCAYTSLGVALYS-SRLLVLTGYIVAVLIYHLKRHQALRYLWNRLVMFAQ

>871204620
YTLSSWWGTCWRSNGCRCNHLLIPNDLSFYVMLGSMLRHSDLHDKGSCVYLQDNTLTHRQKYEVVMETGFSRNAGTTSKISIILHGSETRELFERNSRDRFIVTLPDSLGRIFKVQLWHNNSGPGWYLRQVVVRDLSAG
QAFYFLCQRWLAVGKVEREFTALDSKN-
ISFVMVFWCKGSQYVSDFHTWLSVWTCPSRFTRVQRLTVCLTLLLSYMCLNAMWFSWHNIIVGFICCLVVVPNLLLSFLFRLPFWCRYVAWLCFLISGTCAAVTVMYGFRFGHTKSAMWVQSVYFSFMICQPLMVLLSVLYT
ALTHQLKSSRKVLIKQKRALIFLRDCVSFVFMFVLLCLITFGKDGFERLGSPAHWYTWSQTTLLDVLYKNLPGNKFGSLIGQVQLRQLRCIYAY-----
VDSRSWEVDPGLWGQAGWYSSGLVLSLLESQHWLSVQTRAVFVEMTLYNAPSNLFSSVTMLLEVTPAGKHVFRYCELLYIVIYFTHFWNFIEFINLALLTFWDFCAMLKLLQVLRYHQLFWRLQTVYRRCRKEILTAGAL
YLSLVMAFSSLSCGLFGWLPVVFVFTYIFAALSHRLRQTQLFAFFWDQVLRWSG
>908481541
YTLNVWWGSCWGSTGCKCSHLLIPNELFFYGLLMVVLKQADLHDKGNCVYLQDNTSSHQQKYEIIMETGLCRSAGTTSRISMILHGSETRELFERNSRDRFIITLPSSLGKIFKIQLWHNNTGPGWFLRQVIVRDLTAGQ
AFYFLCQRWLAVGKVEREFTAIDGKN-
ISFSLIFWCKGLQYLSDFHIWMSIWTCPSRFTRVQRLTVCLTLSLSYMCLNALWFSSHNIIVGAACCAIMLPNLLLSFLFRLPFWCRYVAWLCFLICCACATLTILYGFRFGHTKSAMWIQSVYFSFMICQPFLILLIVLYTALCHR
LRSSRKKLIKQRRAMTLIRGLGTFIFMVVLLMIIAFGKDTFLRLKSQLSWYKWSQTVLLESLYTHFSVEKCAHIIGHVQLRQKKCSPV------SNTSSQTWNPKIWGE---
FMSGQVIQLLESDNWIDLYTQAVMVEMTLYNVVTNLFSSVTLLLEVGPAGNYVFRYCEMLYIVVYFYHFWNYMEFVNLQFLTLWDFCVVLKALQVLRYHQTFWRFQITYRRSKNELLLAGSLYFGFVMSFASLSTCVFG
HLDLVLIVTYVFVVLSHKFQKTQLFSFYWEQL-KWFG
>918292252
----------CDLRE------L--
YSVLLFLCHKADLHDKNGIIYLEDNSDTDRQKYEIIVETGFRKGAGTTSKISIILHGSETRELFERNSRDKFVMTLPDSIGKLQKIQLWHNNSGPSWFLSRVILDDLNSGQTYYFISERWFSVGRVERELWAAKD-N-
VGFNTIFFAKGAQYIADYHLWTSIITKPSRFTRAQRVSVCLTLLMSYLSLNALWYSWQSIAVGFCTCIITLPNWILVFLFRLPVYCRWIDYLCLLLTLLCAAVVMFFGLRFGHTKCMMWLQSVYFSLTTCHPLCVFILTVYNAY
NYRLLSSSKKSIKERKSYSFFRDAAMLGLMIGVLCVMAFGKSSFLGIKNVDQWWEWLQTSFLDQMYWDPWYDGSSYLLGKVHIRNVQCEDS----------
DGATIRNTLTKLMKKDEIINEQTQQIIVEFTLYNPPTNLFSSVSLNLRILETGVHLFRYCELLFVILFLGQAWKLIDYVDISFITIWDFLSVVKLLSLSRYKQLFSMTESVYSRASTEILIFAVTLFVVIIAYTSLGCLAFGGGKLF
VFTAYIFSILSSNHGSYESLIFIWKKLLVWTG
>1139740091
YTLTVWQTQCWSRDGCRCNHLEVFPVLSLYVTLMVLCHRADRHDKGGILYLRDNALTDKQKFEVIVETGFRRGAGTTAKISIILHGSGTRELFERNSRDKFILTLPESIGKIWKVQIWHNNVGPSWYLSRVIVKDLNNER
EFYFISEKWLAVGKVEREFLALDG-N-
LGFKRVFLTKGTQYLADYHLWLSPFTCPSQFTRTQRLTCCLTLFSVYMCLNTLVYSWNSVIVGVVCCVVVIPNFILSFLFRLPTWMQYVTWLCALIIVCSTVLTIWIGFRFDEVRSILWLQSLYISVLFCHPILIFIAVLYTTVRY
RLIESRKKVMKEKKALTILLDIVSLVATFLIVLVMAYGKDIFTSIKFRQEWYDWASTTFMESIFKENPAVLQGYIIGHVQLRKVKCRNI------
STFQSKSSPYLLFGRHGIYNSGHVIKLLASSHWMDDFTKAVFVEFTLFHPSTGLFTSVILYLEIPDTGGYLFKYCELLFILQFLHYGWKVLQFVDVSFVTFWDFVVILQSLRLLRYERLFLLFGKIYSRAHRELKVLAILLFVLL
LGFAALGHALFGATRMFISTIYVVSFLTFRFHVTETLTYYWRMLHQLAG
>260823633
YTFRLWWGECWSGEGCSCNHLASMSANDFYLILLFCCYRADKHDKEGPVQLQDNSLSDQQWYEVTVETGFRPGASTSAKVCLVIYGSETRELFERNSRDIFIVSVPDNLGPLWKVQVWHDNSGPSWYLSRIIVRDLHSN
YAWYFPCEQWLAVGQVQRELTPLAD-G-
AGFKKLAYCKGCQYMTDYHLWLSIYTRPIKFSRVQRLTVCLSLLLSYCCINAAWYSLRSIIVGALTALVVLPHLGIVFLFRLPPWCRYVAWACVVIVILLSGMVTILYGFRFGKTQSIRWLQSLFFSLMQSQPLLVAVVTLYISW
RERLRPAQEKSKQEKLLSEGIRQGWLYGTLFVLMMLIAYSRVSFGSIKSVDEWWRWSQGHLLDHMYWTSWYDGGSVLLGEVTLRQLRCRPDY-----YTAARS---
SSHSGHFSYYGGGYVVALLQSQQWIDEKTRGVAVEFSYYNPPTNLFGTVSLLLELPSLGGTLYRYCELVMLAVCYCNPWHWLQFVDLSHIADMDILLMLKMLHLLRFYSVFTTFSMVFNRASWEIFTFMLWLLSLLVAFS
HLGTLLFGQLGLYLLLGLIIGILTSAYRVVTMATYVWEKFLMWAG
>585678603
YTLKLWWGECWLNDGCRCSHLQTEQELSFYLILLIVCYLGDVHDKRGVVYLQDNQQSDQQCYEITVETGFRSGAGTTAKVSLVLHGSETRELFERNSRDKFIVSVPETLGTIYKLQVWHNNAGPSWFLSRISVRDLING
QKWYFICEKWFAVGKVERDIYPQDY-G-
VGFQKVFYAKGSQYFADYHLWSSVFTKPSRFTRVQRLTCCLAILLTFMCIIGTWYSIESLVVGIVTALMALPNGPIVMLFRVPHWCRYIAWLVSQVIFWAAFFVIFFGIRFCKSKSVMWIQSLYFSILQCQPLVILTVTLIRSFQY
RLQKAKERMMKEKTMYSILKGCLFYCILFTCVMLMCYGKFHFTSIQSIPDFWAWCENDLMDGLYWDSWYGGSHILLGKVSVRQLRCRGTF-----
YQDTSHIVRYCKWGRFGYYGEGFVTNLLKEDNWIDVHTRLVIVEFTLYNPPTNLFTSVALVLELPGTGGMLYRYLELMFLALHFLRIWAYIEFVDMSFLALWD----------
YSGVGVIVFGAVPVFYLLGM----------
>780160480
YLLKMWWGECWSPEGCLCTHLPILNIVSFYALLFIYCYYADRHDKQGLIYLKDNQPGDKQYYDITVETGFRSGAGTTARISIVLHGSETKELFENNSRDRFIISVPESLGNIQRLQLWHNNAGPGWYLSRISIRDLLTGQK
WFFICERWFAVGKIERELLVVDD-G-
VGFTKAFFAKGAQYFADYYLWSSVFTRPSIFTRVQRLTCCLCISMGFMCVNTLWYVWGAVLVAIEQALVAIPNFPLIHAFFLPHFFYFVAWGSIFIVIASSVITIVLGLKFSYAQAIYWMQSMYFSFLQCQPTVMAMLTIGKA
WKHSLQKAKEKMLKEKHMYAILRDVILFLILIFCVLWMAYGKETFTNVTTVDDLWHWLQEDMVDTLYWQRWYTGESILLGKPSLRQLRCQGGY-----
FRTAHQLLRYCKWGMFTTYGDGYAVDFLRNNSWIDAQTRAVILEFTLYNAPYNLFSAVAFVAEIPGSGAQLYRYCELLFVLLYFRSAWSWVEFVDLSFLAWWDFIVTLKSLHLLRFNRMFAMFGAVFEAALGHLLLHML
CYTILETAFACLGNLLFGLWPMFYMAGLMKAILCHNYRKVEVISFFWHRFLCWLG
>1229127695
YLMRLWWGECWVGDGCKCNHLPISVILDVYIVLLVYCYLADKHDKLGLIYLSDNQPADQQYYEITVETGMRAGAGTTARVSMVLHGSETKELFEKNSRDRFIVSVPESLGNLRRLHVWHNNAGPSWFLSRVAVRDLLT
GRKWYFICERWFAVGRIERDLEVLDK-G-
VGFQKAFFAKLFQYFADYYLWSSIFTRPSVFTRIQRLTCSLSISMGFLCINTIWYSGEAMLVGIVTAILAIPNFPIIMVFRLPNWCGKLAWICGFITVFSAVVTIYYGYRFSDVKSVYWMQSVYFSFLQCQPAVMLMFTILKSVR
HRLQKAKEKMLKEKNMWAILKDVIMYSTLFACLMWMAYGREVFTAVSSVDELWSWMNGDLLDNLYWEKWYNGESVLLGKASLRQVRCRAGY-----
HAGKAALKRYCKWGYFGSYGDGFVISLLQNNSWIDEQTRAVIVEFTLYNAPTNLYTSVAFLAEMPGTGALLYRYCELLFLIFFKSIWSWMEFVDFSLLAYWDFLASVKFLRLLRFNQTVAMFGAVIDGAIKEILTFALYM
LVLEAAFACLGVLLFSLWHLFYMAGLIKAILCHSYKQVEVVRFFWNKFLVRVG
>443732271
VTLMSSVGRCWKGDGCKSNHLPVVTSLSFSLVLMILVKYLDVLDQMSMIPLCGP--
NGIYKFEVTVVTGQRPGAGTTAKVGLRLCSCFSRNLFQRSSVDVFLVATDEDLDEVDVLRVWHDNTGPSWYLNYIIVKNTQNGQKAFFPANCWLSLQRLQKGIQRADKKQLSSFALQFSNEVSRCFQEHHAWFSIFSCP
SRMTRVPRLLLCTTMCLSIMSIASIWH-LDSAILGVLLPIFIYPGILVYVLCSLPHWTVYVAYVCILPAFCSLIVVLVLASRMQE--
PLLWLLSVSMSVLVSEPAKIIHAGINSLHGNLQQAKEESRKYVQLKSLIQHFVVNLILLWLLLSVNYSAYFLRDTLNYTSFWDWVEGTASC-VVHLESSTTHSYTLGEARLVQLR-RE--
KTLEFYKSSAELNSIPYHGHLAEYAGGYVKLLLRQMNWIDRSTRLVFLEYTKYNVNKDAFAAVRWAFEFTPSGGRFLTQCQVLLVPIFFCSFWHWIRHISLLPLISARFLMMLILARQLRFLRPLSVFGRTLGHAAPALCGC
AFMFTFVHIAFSSLANQWFGFTSVFLATGLVIAVFSNAYKFVEVVEFMIRRLKTWMG
>PdumPC1
VTMTLFYSQCWEDNGCRCNHLHAPINIHFYILALLFGKYQDNKTLTSVIPLCGK--
DGVFKYEINITTGRYPSAGTSAHVGIIHGSSARHLFKRNGDDTFMVAMDHCLGHLTKVKVWHDNSGPSWYVHRIIRDLQSGQRYLFHANKWLSLPVIEYEIPVTVDGELVKFETYFKESTAVGFADSHMWLSLINRPS
RFTNVQRLTCCVTLLYCFMMTNVIWYDWQDLVIGALAGLLAFPIYVLTVIFKLPHYFIYVAYLCFALSVFSIIVTLLYGNSFGRHLAQKWLIAFFISLIGSEPLKVIIVALVMSTFKPLMQAKEETKKVQRMRLILRQMTAYCL
FFWIMVVLCYTRHLNEDALNQKS---WYENVFFDVLQKLYRPNFQMYMVGIPRLRQIRITG--
YLEVWHTDAEKALGQTYRGKTAEYGGGYIINLLKMQNWIDRSTIAVFLEATFYYPPKNTYTSLILAAEFPLSGGELKPW---
LLAATYLKEFWNCYEFISLAKNLELDFMLTIWIVKQIRFIKSWSLHGNTLRVSAPALLGISFILTLVILAYAQLGYLWYGFTQLYVIMGLVIGILNSTYRDLEMIEFTIRRFKRWLG
>1087812391
VGVGVYYSSCWSISGCLSAHLSPPNIVDIYILGAMWARKLDLKDNASLIPLCGH--
NGSHKYEITIKTSPGIGSGTTANVGICLYGTGSKHLFERGSVDVFQVASDTSLGTLYKIKIWHDNTGPGWNLSRVMVKDLQNGEKYYFLVDSWLQVTCLARTIRVATEEEMEDFSYLFGNQMSYGWEDAHLWLSLIERP
SRFTLVQRLTCCVTLLYSFMVASAFWFRWSDIAVGAIVSLMIFPNHILVTLFKLPHWTVYFTYLCFSIWVVSIFFHLYVHAFGVTVAQRWLLTMFIALLESEPTKVLILAAYIAMFKRLEHCKLNAKKSYRSRKLLVMLTLHGLF
LWAVMAFCVWDRDFDNIRTLSEVWGWMEQVMSPLLFEDTLDINQNHLLGNVRIRQQRCLP--YYPGWFEVPC-
FTGIGYWGQEQIYDSGYSKYIMQEDSWIDRYTRVCMVEFTMLNPASQWFTVTTVMFEIGSTGARLFTPI-
SFIGVLYAKKPTNWMNFTNFYKMAFLSTILLIKLTRLIRFIKRWSVFGDTIYCACKHVLGTAIMLLVLMVGYALFGYLIFGFSICFFFCRLFSAIYIQEYKDAEMIPYMFQRLKLRLG
>871263454
MSVGVYTSTCWSNEGCQCNHLVSITDIDFYVIVAIICRYLDRLDRISRIPLCGR--
DGAFKYEITVVTGRHYGSGTTANIGIKLYGGEARHLFQRNCRDTFYIAHDASIGEVNKVLVWHDNGGPSWFLSHIVVKDVQTGEKFTFLVNSWLSLGVVQKTVKSADKEELRKFRLKFSTVFHGTMSDVHAWLSVSH
RPSRFTRVQRATCVFTTVLLYMCINAMWYGWEEVIIAIISTLMVLPLIGLSFAFKLPPWCVYVTYVCVVLCVLSVILMLLYGYNFGTAISLKWLVSLFISVFLSEPLKVLLISVIVALITKLLQAKEEGKRKQRLHKMLREVFMF
TLMLALFMAIVFVNYSFTNMRSTDFWQWSETVLARSLHLQDLSESFGYLLGPAWMRQIRCTD--
YNIGWYSSCGQLDLYSKFGWSLFYGGGYVQTLLHSEGWLDLLTRAVMVEFSLYYPSADLTSYTILIIEFPISGGKFLRFCQVFFSLFFFSYFWNWMDPSRLGQVAVLHFLFVLKIVKQVQFSRPLFKFYTTLSTASGPLLGVFI
IFVILLLTHAQFGYLIYGLTHFFFTMALSVSVLMWSYKWSEMIDFMIKQFKLMTG

>908506947
LSVATYYSFCWAVDGCQCNHLAPIVDFDFYLVIAFICGYFDRMDRISRIPLCGR--
DGGYKYKITIVTGRSFGAGTSANVGLKLYGGQTRHLFQRNCHDSFLVAFDTSLGEITKMMIWHDNSGPSWYLSHVIVQDVQTNTKYTFIVNSWLSLGTIQKTVMVADDEQLKSFKHQFSTVFFATLAELHAWLSVSNR
PSRFTRVQRATCCLTCTYLYMGVNAIWYGWEEIIIALVSTIMVLPLFGLSYLFKLPPWCLYVSYICMKLCLLSIILVLWYGYRFGASVAVKWLVSLCFSVFASEPVKVVFITVIFKVLKKLLHAKEEGKKSQRLHTMLREAVMF
MLMFALFLSVVYSNSNLSKIKSVPHFWQWSQSVMARVIHFRELDETFGILMGAAWLRQIRCVD--
YGVRWYSSSSELNMYSKFGKSLYYGSGYIQKLLETDQWIDNKTKAVFVEFMLYSAGVDLTTSGSVLIEFPITGGKLLRYCQVIFLMFFFLGLWNWIDSAHLAQVVLMHFLFCLKVVRQIRYIRPLYKFYMTLSVSSCPLMGV
FFIFIILLLSYAQLGYLIFGLTLIFFVTALSISILGWSYKWSEMIDFMLHKFKLITG
>919023136
VAIGLYVTSCWDTGGCLCNHLVTPNAIKFYILCVILARRWDLADKVSVVPLCGR--
DGVYRYEIVVITGRQRGAGTSAHVGVRLYGSEPQHLFQRNSRDVFIIASDINLGDIYKLRIWHDNTSPAWFVSRVVVRDLQTNQKYQFILDCWLSLPTIEKEVLVANAEELNNFSTIFGSEASKGWSQRHLWLSTFERPTS
FTRVQRVTSCVTFLYTFMCVNAMWYSWEEIVVGVVSSVIVFPNLVLIQMFALPHCCIYLAYLCFALSGLSILVILLYGHKFGRDVALRWLISLAFSFVQSEPLKVVFIALFLAAISKLIQAQEEAQKKKRIHVLIRQTVVYFIFIW
LVLVINYANHDFSNITSVETFWQWADTVLAPSLQGINLYEFEGHLIGNVQMVQVRCYR--
YREEWYFLADEIDSLPLMGRLQWYGGGYVQTLLRDHSWLDERTRAVVVYFTVYNANLNLYSSATLLVEFPLPGKKLQVFIQCLMGAVYFTWFWNLYENQKFDVGLKIHFLLCLKLTRQFRFVKSMAVYGKTLNEAAT
YIMGCFILLVIFLISYAALGYLFFGFTAIFFIGGLVMATVLNIFKGVELVDFMLKRFKLLIG
>405970250
VSAAIYMSSCWSRSGCTCNHLVPPDAISFYIVAVIVCRRMDKYDRIATVPLCGK--
DGSFKYQITVVTGKSRGSGTTAHVGIKLYGSKERQLFQRNCVDSFFIANDSNLGNVHKIMIWHDNYGPSWHLIQVIIRDMQTEKKYYFFGNCWMTLGYIQKEIKAAGSSEIRRFSRMFSHAVSNSMADRHLWVSVVDRP
SRFTRVQRATVCVTLLYVFMGVNAMWYGWEEVVLAFLCNIMVFPSLLLIFIFKLPEKCIYLGHLCFLLSGTSCILVILYGQKFGADIALQWLLALLFSIIFSEPLKALIIALYLAAWEKLLKAKEDGIKIQRLRLMFRQFVAYM
FYLLLVMALCYTNFTFRNISTIGEIWKWTQSVLISSLHYDEFSTELNTLLGVARLRQVRCEG--
YGYMWFHSDSQTGSKRREGHVFSYGGGYVQSLLQSKDWIDLRTRALFIDFTLYNAAEDLTTFTTLMIEFPLTGGHLLRFCEVILFAVYFCDFWNWLEFTNFERVADIHFFLMFKVVKQIRFIKFLHKFEKTLNIAFPKLM
GVMLIFGILYLNHGMIAYLFLGFTHCLFTIALIIAVLNDSYKTVEMIDFMMKRFKLWAG
>1205910066
VSTAMYTVSCWSPSGCECNHLVPIDAISFYLVATIISRRLDLRDRISSVPLCGK--
DGSFKYEITVVTGKQRNAGTSAHVGIKLYGSNRKHLFQRNCQDSFIIAHDTNLGDISKVFVWHDNHGPSWYMVQVIVRDLQTDQKYYFFGNIWLTLGFIQKEIYAAGSAEIQKFTRIFGHAWSSSVSDRHLWMSVMDR
PSRFTRVQRATCIVTILYVFMCINVIWYGWQEVVVALATNAVVFPSILLVSIFKLPSWCVYVGYLCFCLSVLSVIVVMLYGYQLGAEVALQWVLALLFSVTFSESFKALFMALYIAGIEKLVKAKEEGQKVQRMRLMLRQFV
AYMFYLWLLMVLCYINFNFQNMTSIQDFWVWSQDVLASALHWEEINTEYGTLLGTARLRQVRCVG--
YGETWYYTSDLTGSKWRQGNVFSYGGGYVQFLLQAKDWINLHTRAVFVEFTLYNPATDLTTHVNLLVELPLTGSKLLRYCEAILFVVFFTSAWNIYEFVNFETTVDVHFLLMFKIVKQLRFIKILYIYERTLSESVGKLSG
VALIFAILFLNHGMVSYLWFGFTHFFFTVCLVVAILSDTYKTLEMIDFTMKRFKMWAG
>205360954
VSVGLYTSLCWRTEGCLTRHLVPPSHVRFYMVMAAILHKLDQLDRGRAIPFCGQ--
RGRFKYEILVKTGWGRGSGTTAHVGIMLYGSGHRHLFHRNSLDIFRIATPHSLGSVWKIRVWHDNKGPAWFLQHVIVRDLQTARSAFFLVNDWLSVGLVEKEVLAASDAALLRFRRLLVAELQRGFFDKHIWLSIWDRP
SRFTRIQRATCCVLLICLFLGANAVWYSVDTVAVGLVSSVVVYPYLAILFLFRLPAWCASLAHLSLLLVAVAVAVSGWVGASFPPGVSVAWLLSSSASFLASEPLKVLLEALYFSLVAKLFLAKEEARKVKRLHGMLRSLLVYM
LFLLVTLLASYGDASFLATRSEELWPWMAHVLLPYVHGNQSS----PELGPPRLRQVR-----YDVGWYSAPD-
LLGAWSWGSCAVYSGGYVQELLQLHNWLDNRSRAVFLELTRYSPAVGLHAAVTLRLEFPAAGRALRRLTSVCLLLFRVLRLGAWARFTSFDQVAQLSFLLVKAAQQLRFVRQWSVFGKTLCRALPELLGVTLGLVVLGV
AYAQLAILLVSLSPLLCALRLGAVILRWRYHALEMVELFLRRLRLWMG
>1207107604
LEVAVFTSLCWRTDGCRTHHLVPPDAVSFYAVAGAVLRKLDQMDRASVVPLCGK--
DGLYKYEIQLKTGWSPGAGTTAHVGISLHGSGHRHLFARNSLDIFHIASDSSLGSVLKIRLWHDNKGPAWLLQYVLVKDLQTGSSFFFLIEEWLSVGVREIEVEASEELYLRDWPRLLRWELQRAVCESHVWLSLWERPSP
FTRLQRLTCCSLLLHLQMLAAAVWYNVETLWAGMVSCLLVYPYLLVFCLFRLPPLCGLAALASWAGLILSSSLSMWLGRGFSEGVALMWLISCIVSFLSSEPLKVLLEGVYFSLVVRLSQARFEEARKVRLLHTMLKGFLLYML
FLMVVLLLNYSDSEFNNISSREDVWAWLSNSLLPRLLDEQVLKTGSLLIGRARLRQTR-----VQAGWVSTAQ-
SSGVWHWGQVSVYSSGFVQDLLNTHQWLDPLTRALFLEFSLYNINTDLLSVFSFLLEFPVSERSLHTLLTLILLVLYFLRVWNLASFTDFFPLAQQNFLLVLKASHQLRFLREWGVFGRTLRKSFWELLMVAVTLLVFLLAYS
HTGHLLFH--SFPCL----MSALLRNYRRAEMVELFLRRLKMWMG
>260801585
VSVGLYCSSCWDDAGCYTSHLLPPNAITLFLVAGFVVRKLDLRDKVAVVPLCGR--
DGPYKYEISVMTGMTRNSGTTAHVGVNIFGSGRRVLFLRNSTDVFQIAVDGSLGDIIKIKIWHDNTGPSWNLKSLVVHDIQTDKRYHFLANSWFSLGFVEKDVYVASPDALKKFSTVFKNETARGLSEQHLWLSVLERP
NRFTRVQRVTCLATLVYSFMCVNAMWYSWEELAVGIISNLMVFPNLLIVTLFRLPYWFVYIAYLCAIIIAGCMVVIIMYGQAFGRDQATKWVLSMFLSLIQSEPIRVILIALFLASISKLLQARFEEARKVRVMYTIHQCIFYLLF
ILTILIIAFNDNDFLTITQSEDWWEYMDSVLLPTLLDTSQQSLEGVVLGGVRLRQIRYYN--
YGETWYSETNSYNGLMYYGRDGSYDNGYILDLLNSSDWIDRGTRAVFVEFSLYNPNVNLFSVVTLLLELPVTGVRFVVYMQGLFALFYLKRFQTYYEFISFFHSAYLGFLLMIKVARQMRFRRQFSIFGKALRLVTWELA
GCFLIFFILILTYAQLGYLLFSLAFLFFVGRLFGAILIHGFTAVEMVEFMIKRFKMWTG
>585696579
VSVGVYFTTCWESTGCQCNHLS----SAVYIVAVFIARWRDSIDRIGHPLCGK--
EASFKYEISIKTGMKRSAGTTAHVGVNIYGSGTRHLFMRNSLDVFQIAADINLGDIKKIRIWHDNTGPCWYLSRVVIRDMQTNKKYYFLCEQWLTLGVVKKELTLASDKQLSKFSRIFWAETSHGFSERHLWFSVYERPS
RFTRVQRVTSCFVLAFLLLALNAMYYSLDELVVGVISSLLAFVILVVVQMFQLPSWFLYINYVCLLLVGTSLAVVLYYSRLFTVNQSLQWIVAVAIAFAASEPIKVLFISVFRALLRKLIQAKEEAKKINMMYTLLRQFIVWVLF
LSLILLINFTANNFLSIRVAEDFWIWSKNILARGLHSVELDYTYGFPIGVARLRQLR-----YNEGWYQTEEQLDGMSFKGYTNTYGGGYVQHLLLANKWIDRWTSAVFVEFTVYNSNVNLYSVLTLLVEFPASSGNLT-
YFMMLLGIFYFASLWHWIDFTNFYRAVYYILLVTLKLLRLLRFNRMMSTFSRTMSEAGRYLFCCTMIFVIVLWAYVQLTYLVYGLTPFIFIAALFAAIMINSYKVSEMIDFMMKRFKQMVG
>565304042
LRIAVFGAGCWKEETCVCKAKLFPNQVDMYISLAVWVRRKDRTDEEQIIDLPDNDPFHTMRYLVTIYTGSRPSAGTKADVFIELIGSDVHVLLQNTCIDTFMLTTKEDLGDLFSIHIWHNYCGPNWYLSRIKVLNVDTN
KSWLFLCRKWLALYQIEWTFEVTDPEVPISRTDYVLIMFSYHWVQDHLWFSIFVRSSSWNRFQRLSCCLAIFLTSLLVNTIIFYLKSISAGVLSALISIPEWIAVSLFKWDRCLSCSAWFVFLVSGVASFFITFIGLSFDNKTSTEW
LLAALISFCLHETLQIVIFSGWNICLAKVTALQKKQMIQYLAYKYVKDMICHFIFLALIFTISHPINILPNLNVAQDVYRWVNNNFLPLIHNSDHPNTWSKILGLPRMRQIRCKNKYYMGSWYEPSSIHWDYITEGQLNEYS
RGYTFYFLETKSWIDNKTWAVILFLTTFNADEHMFCSISVIFELSLLGPKLPFF--
CLTIYFYFTLAANLINFIPFHAASQFDFFLVLKTYRYVRFLFDLRLAQKSLRTALPGIINIVVMGGIFFSAYMALGYVGFGLGGLFLFINLSQVIIMSAHVDMEVMYFILQKVYRMWY
>5174633
VRISIFSVQCWREGYCICKNVVIPNPVDLYVGLAFWALYRDEMDRGHVIVLPDNDPYDNLCYLVTIFTGSRWGSGTRANVFVQLRGSDVHCLYR-
GSINTFLLTTKSDLGDIHSIRVWHNNEGPSWYLSRIKVENLFSRHIWLFICQKWLSV-
TLDRTFHVTHPDERLTRKDFFFIDVSSNLRKNHMWFSIFASVKTFNRLQRLSCCLAMLLSSLLCNIMFFYMRSMMIGIESVLITIPQLLITFLFTLPWWCVYVAWLVFATSSISSFFIVFYGLTYGYDKSIEWLFASFCSFCQSQPS
KIILLSGFRTNKPKIRIFKRKKRIKRRALLFLSYILTHFIFLALLLILIVLLRHLATVTKLEDIYRWLNSVLLPLLHNDLNPESSSKILGLPLMRQVRCHPKYYSGFWYKPQGTQWLYYSYGLLHTYSGGYALYFLQESNWLDEK
TWAVVLELTTFNPDINLFCSISVIFEVSQLGVSLADF--
VAILIFYVRSVYNLLNFIPFHAVSQVDFLTILKTLRYSRFFYDVRLAQRAIQAALPGICHMAFVVSVYFFVYMAFGYLVFGLGVLFLLINLFQAVILSAYEEMEAMTYLCRKLRTMFS
>115583675
VRIALFMVHCWQQSTCVCRNGVVPNPVDLYAILAFWALHRDVIDRDNVVILTDNDPFDTLCYLVTIFTGSRWGSGTRANVFIQLMGSDVHCLYR-
GSINTFLLTTKNDLGDIHSIRVWHDNSGPSWYLSRIKVENLFNKRIWLFVCRRWLSV-
TLDATFSVTNPDEPLKRTDFFMIDVRDKLRKNHMWISIFTEIKPFNRLQRLSCCLAMLLSSLVCNIMFFIIRSMLIGIESVVITIPQLLITFFFTLPRWCVYIAWLVFATSGISSFFIVFYGVTYGYAKSIEWLFASFCSFCQSQPCNI
LLRSGTRSYKPKIQILRRENRIRRRSFLFLSYLVTHFIFLTLLLLLIFSLRHLATVMKLGDIYTWLHGVFLPLLHNDPNPDSSSKILGLPLVRQVRCHPQYYSSVWYKPPGKRWVYHSYGVLNTYSGGYVFYFLEGKNWLDEL
TWAVIVELTTLNPDTSLMCSISVVFEVSPLGVSLADF--
AAILIFYIRSVYNLLNFIPFHAVSRVDFLTILKTLRYSRFFYNVRLAQKAIQAALPGICHTALVVSIYSFMYVAFGYLVFGLGVLFLFINLFQAVILSAYDEMEAVTYLCNRLKSGFD
>1101394602
YNFTAFAVACWSTQGCDCDHTVPPNKLEPYFVTLIYARRLDRKEKRGVHVLLDNDPNDTYFYVLTVSTGARPKAGTSAKVGIVLKGTEPHILFETGGEDSFLLATSDYLGELQTISIWHNGEGPSWFVDEVLVRDVQT
NEVFFFLPDRWLSAYQLSAEIPVSTHDDLTSTYHRFTTFTTSGIRDGHLWISIFFRPSPFTRAQRATCVLSLLLTTMLANIMFYNWNQIMLGIRSALVVLPNLGVIYLFQLPFWTKLVAWVGLGTAIVSTILVVWFGLHFGRTK
SQEWLASFFSALAMDQPLKVLLVSFLLAAVLRVAAVRKMTFLTERMRTWLTDIGWFVLFMFSVFVFTYGNTPFSKVKTVEMLWLYLESTFLVSYSPEKTADEANLCLGPARMRQVRCLPEYFLASWYRNATELGIGHFMA
NLETYGGGYVAVILRNSVWIDDLTRAIFLQMNTYNPSTNVFSIMFLVFEFHQSGLRLMYTCAVIFLCFYFRDPWNWLEFVSFYFAVLTEFVCSVKLLKIFRINRKVNSLFYVMNDARVPLVEYGISLLLAMTPFHAGHVIFG
LGPVFFYTNAFVSLIMIVYEKTELADFVVERLKRLLG
>48094207
IALYAYGAQCWDTTGCQCNHLVLPNTVHFYVLVAVWANRHDRRAQARVIFLGDNNPEAQYNYHVTIFTGLRRVAGTTAVATITLYGSEPHVLLQRGGVDSFVLSTHGSLEELTAVRVWHDNSGPEWHLNRILVHDLQ
TKQLWYFLCHTWLAVCTVDKTFPVATTEDLQQFQHLFITGLLRDMKDHHLWVSIFTRPSNFTRLQRITCCFTLLMMSMLASIMFFTWFDILVGIESGLLVLPNLLILEIFRLPWWFVYIGWFVVVTNLVCGYLVILYGLTYG

LQTSMDWLVSMMVSLFQSQPGKVILVAAFSAVIWKVNQVRERREAEAKMCKMLKELLGQILLVVLALEVAHFQRDFEDVSSINDVYGYLEESFVPELY-
GENRDESTYRVGRARLRQVRCTGILYNKSWFTPSRSSHGSYHMGTVTLYQGGYIQVLLEDRKWLDELTRALYIEWTVFSADAGLMCVCTVLFEATTMGNKIFSLLQSVYLLVYF-
QFWNILNFISFFYPVIMDFINTLRCAKILRFNSSIQTIHHTFSNITSEFFIFTLMVLIFLFAFALLFCVTFGLGPLYFLFNFYILLILYAFRDADVLSAIVDWLFSVCS
>1229186594
YHVYAFSTQCWLGDGCLCNHLVLPNPVDLFGLCAVWAHRKDRIEKARLIILADNDQEDTYRYHVTVFTGVRRGAGTTAVVTLTLQGSEPHVLLQRSSVDSFLLTTRESLGILRSVRLWHDNGGPAWYLSRILVHDLETD
RWWYFLCHTWIAVGKIDKTFPVANTEDLKDFHHLFFTRSVRDMKDNHLWLSVLSRPSSFTRLQRVACCFSLLLSSMLMNIMFYSVKTIVIATEASLLALPNLAIVQIFRLPWWWVYIGWLAISSSLFAAYLIILYGLTFGKQKS
LDWLVSMIVSLFQSQPVKVIAIAAFFAIVLKVEQAREKRHAERKMYSLIKAFLGHLVFIWLVLQVSIAQRDLDQVKTMPDVYTWIKGEFAPSLY-EAEPQE-SLLVGSPRLRQQRCYA--
YNASWYQSHWQAKGSYFLTDPGRYRGGFLLDLLQDLGWVDQYTQAIILEWVVYNANTNLFCACTVVFRAEGTTTCIYRYLEVLYVMVYW-
AFWNVLNHLSFADIVVLDFCSLLKCLKILHYNDSMYLLHSTFRRLTKTLLIYLSMLTVTFVAFAYYGYLSFGQAAYFFFIVNLFIAMLLHCYSAADLVSMLTEWFLSAFS
>1126169400

LPVPDCRCDHLVAPNPLNIYFMGVLWARKQDRRDKVGVGILPGHNPRKDCQYVITLYTGFKGNAGTTAEVTIVLYSSPPFILFEKGSVDSFLVSTSEPLGDLTYLRVWHNNAGPAWFLSQIVVTNRENNKSQFFLCNRW
FATGKVYRLLPESGPDEMKEFRNVFLAKSSRDMNDGHLWFSVAGRPSPFTRVQRLSCCLTLLYSTMVTNIMFFSLPQIMIGLQSAAIILPNLLIVFLFRIPWWAVYFGWLVWSASFVAAFFTVLYSLSFGRAKAEAWLFTFLTS
FLTDQPFKLLLVAMLFALLVKLAEQRTESKQKRRRRKQIVEVLMFGLFVTVIMLTSYGERSFSEIQDIPSFWTWVITGLIPATHVSRWYDMLTHPLGPVQLRQVRCVPVYYTERWYTFAAVTDAFPYFGEHGTYAGGYVT
SLLQRQDWLDDKTRAVFIELTLYNPHVNLFSVVSMAVEFTNLGARLIQRLRGLFALFYLSEFWSWVEFEGYKNAVSWYCCSTFKFVRILRFNSHVYALSMTLRRSLKPVTQFMFTAGIVIMAFTQTASLIFGLGPLMFLLSM
FMAIIMDVYADQVIKESATKTFRKVRG
>260785879
YTTITFSSTCWKGDGCLCDHLVAPNSLDIYVLVVIWARRKDRTDKVGVTVLPGMT-
EYRHRYIVTLYTGFKRFAGTTAKVAIVVTGSEPHLLFETAGVDSFLLTTAQSMGPLSSVHVWHDNSGPSWFLDQITVFDLQQEQKYVFMCNRWLAEGLIERHVPVASEIDLKDFYYVFSNRMSKDLRDGHLWFSVAGRP
SLFTRVQRVSCCLALLLCTMVVNIAFFTWTELMISIQSAL1VLPNLHVQIFRLPWWFLYIGWLVLAVCTVSAFFTLLYGNSYGKRKAEAWLLTFFTSFLLDQPVKVLVVAVALAAIVKLAAARELRSKQRQMYSVMVEVIFYL
LFVWVVVSVANGHRSFEEIGELDDFWTWLEEGLIPGLT-ETWYGHVHYLVGVTRLRQLR-----VEYSWYQTS----------
SALGGGYVAELLKNNMWIDDNTRTVLVEFTIYNAYVNLFSVVTLALELPVSGGRLYNYCEILYVIMYFGEFWNWVEFVNYSEAAFWDTLVIAKFVKLLRFNRRMSLIGDTIKHAAPKVLGFIVIYIVVIMAFAQALFFVLG
LGPAIWLVFTFVAILMDSFESVEIVDFMVKQFKDYIG
>641798600
YTITTVATQCWRQDGCLCNHLTVPRTVDIHGVVGVWAWRRDQVDKVKVTILADSDPGAHVRYVVQVFTGYRRGADTTAKVVLTLYGSEPHCLFERGGLDVFLVTTRSSLGELHNIRLWHDNSGPSWYVRQVVVSDV
TARRKWHFLCSTWLAACQLDQVFPAASRSELLSSRHLFCSALVEKVTHDHLWLSVLTHCNPFTRLQRLACCLTLLLCTMVINIMFWTWQELVTSVQTALILLPNLVVGHIFQPPREAAFICWTVAAASLASAYFTVLLSLKL
DRQRATSWVISMLLSLLQNQPLKVAGLTLLFSLVLKAPPRRGRTWKKKKLYTQLGEIVAQLVLLASLMVVCYTERSFYHRGLGDFYRWAHGTLLPSLD-GPRPNA--LLLGSVRLRQLRCRQ--
HGPGWNHGAESLGEHPVWGQLSLYGGGYLAHLLERSRWLDGCSCAIFVEFTVYNANVNLLCVVTLLLETPGTGARLYS-
AQVTYLLLYFRSQRNVLDFVSFYEAAQVAALATGKLWNLLRLSPRMQLISGTLRRAWQEVLGFLLVLLLLLVGYSMACNLLFGLGSLLIVINLFVSGLLTTFSEESMMEGMLLKLWGLFG
>115583681
VSVVTAVTQCWKSDGCLCDHLIVPRTVDVYILLAMWASRKDREDKVKVTVLADNDPSSASHYLIQVYTGYRRRAATTAKVVITLYGSEPHHLFERGALDVFLLSTGSWLGDLHGLRLWHDNSGPSWYVSQVIVSDMTTR
KKWHFQCNCWLAVCERDRVFTPASRSELSSFRHLFSSTIVEKFTQDYLWLSVATRHNQFTRVQRLSCCMALLLCDMVINIMFWTLSELLVSIQTSHLFPHLIFGRLFQLFRWLKCSCWLLGVISLASAFFITLYSLELDKDQAT
SWVISMMLSVLQDQPIKVIFLTLLFSLMANLAK-PTRKAWKKQLSKLTGGTLVQILFLTLLMTTVYSAKDFSEIKTVEDFYPWANGTLLPNLY-GDYRNS--FLLGNVLIRQTRQHQ--
YGAGWYQNQESLGGYPIQGELATYGGGYVVRLLEQRRWLDHCTKALFVEFTVFNANVNLLCAVTLILESSGVGTTSLQSSQVVYYLLFFTRKRNLLDFISFYEALRVNLFATVRVWDLLRHHAQLQVINKTLSKAWDEV
LGFILIIVVLLSSYAMTFNLLFGLGSLLVIINLFVSAILIAFGKETLTDMLLQKLSSLLG
>31559825
VSVITAVTQCWSSAGCLCNHLVVPRTVNVYVITVVWARKKDQADKVKVTVLADNDPSAQFHYLIQVYTGYRRSAATTAKVVITLYGSEPHHLFERGGLDVFLLTTWTSLGNLHSLRLWHDNSGPSWYVSQVIVCDMAV
KRKWHFLCNCWLAVCELDRVFIPVSKRELFSFRHLFSSMIVEKFTQDYLWLSIATRHNQFTRVQRLSCCMTLLLCNMVINVMFWAWSELLVSIHTAVILFPNLVIGRLFPLSKWLTSVCWLLGFTSLASAFFTALYSLELSKD
QATSWMISIILSVLQNQPVKVVFFTFLYSLMMSPTKHPERTLKKKKLFKLTGDILVQILFLTLLMTAIYSAKNFSEIKLLQDFYPWANHILLPSLY-GDYREPSHCKCGVQLIF--Q----------EGQLAFCNG------
LHLRRFSYICSPRPMVLIPTDELHERLTSKNENGFSYIMRGATSLQMSQVIYYLLFFTGKRNILDFISFYEAVKVNLLATVQLWNLLRHSPRLRVISRTLSRAWDEVVGFLLILILLLTGYAIAFNLLFGLGTFLIVINLFVSAILM
AFGKEALIDTLLQKLSNLLG

Tree file (Newick format)

(585652993:0.58041683,((641798600:0.33155222,(31559825:0.26364266,115583681:0.16380056)1.000000:0.21364537)1.000000:0.38968327,(1207182030:0.15449853,(116006951:0.17385572,942189627:0.09045394)0.989000:0.06975504
)0.997000:0.14110540)0.998000:0.19614304,(919052864:0.6631392,((260827108:0.67559152,((1229186594:0.41383741,48094207:0.54512083)0.986000:0.16892815,(565304042:0.68354580,(115583675:0.19335668,5174633:0.08310723)
0.997000:0.29931234)1.000000:0.83749236)0.173000:0.09558130)0.961000:0.13487314,((((1005483664:0.38826214,(999981301:0.31580961,156395095:0.33136318)0.491000:0.07447336)1.000000:0.30761538,(196000368:0.67539498,((((
260823633:0.46375612,(1238871302:1.40479287,(1207165955:0.98776963,(19923084:0.18910464,1039738876:0.34478073)1.000000:0.54515019)0.993000:0.30609620)0.948000:0.17664864)0.828000:0.09819614,((585678603:0.34981084
,(780160480:0.31638090,1229127695:0.23056500)1.000000:0.18128852)0.997000:0.16278096,((919066838:0.49103533,(443717936:0.50262089,PdumPKD1L1:0.33833189)1.000000:0.19246624)0.746000:0.06515648,(918292252:0.58747
346,(1139740091:0.53516905,(871204620:0.21079664,908481541:0.25679466)1.000000:0.35522132)0.895000:0.11148088)0.740000:0.07171903)0.998000:0.15436091)0.899000:0.06091093)1.000000:0.47635298,(585696579:0.56226467,((
(1205910066:0.22784608,405970250:0.25896910)1.000000:0.22796956,(908506947:0.31853160,871263454:0.23660034)1.000000:0.37338223)0.998000:0.19636717,(919023136:0.50111370,(PdumPC1:0.74397043,443732271:1.04807898
)0.721000:0.14684112)0.219000:0.03102147)0.972000:0.08602450,((205360954:0.44669062,1207107604:0.54010941)1.000000:0.43669245,(1087812391:0.88976947,260801585:0.51796746)0.791000:0.03887543)0.948000:0.07221924)0.7
54000:0.05570260)1.000000:0.31091598)0.999000:0.20583014,((((((585651646:0.45148390,((148539556:0.55710345,1139813591:0.55648789)0.803000:0.03510197,((((Pdum9951:0.64029596,676479791:0.52365407)0.960000:0.12658446,(
58564412:0.55743164,585688456:0.49745853)0.983000:0.19068339)0.806000:0.05485511,((1087817335:0.73214192,(((Pdum12681:0.58828930,Pdum688:0.89574639)0.000000:0.05827439,((1139824403:0.49911829,(871245685:0.47615
237,676439278:0.50311690)0.996000:0.17242095)0.613000:0.07833442,(961088331:0.91204394,(762108557:0.43345023,(405968482:0.32835386,762157119:0.46848710)0.975000:0.12328428)0.984000:0.14296030)0.220000:0.04732619)
1.000000:0.17557770)0.986000:0.15526635,(919100788:0.88256346,(Pdum14759:0.40439390,Pdum15290:0.51286858)1.000000:0.39177594)0.991000:0.23243132)0.154000:0.04251623)0.190000:0.03795675,(260824333:0.66759814,(Pdu
mPKD11:0.36337631,(919018042:0.21090669,(871242653:0.28526900,(1205898338:0.12662226,762107655:0.18878254)0.353000:0.05333355)0.996000:0.10290432)0.878000:0.03735103)1.000000:0.19530139)0.841000:0.07234452)0.29
3000:0.03054696)0.793000:0.03146914,(1126215987:0.50000507,1087817330:0.56681608)0.998000:0.15799469)0.748000:0.01221693)0.939000:0.05388943)0.975000:0.07510168,(((957833931:0.54948667,((734563164:0.62996362,39289
1400:0.65352859)1.000000:1.03905783,(1009565777:0.43051164,(1067095036:1.12306064,(675368552:0.14598423,(1078802969:0.20801164,1238854278:0.15078289)0.886000:0.06716885)0.916000:0.10151354)0.569000:0.05483302)0.8
76000:0.08127272)0.202000:0.03380496)0.940000:0.08991643,(291224264:0.24841061,(780039266:0.31500732,1229191925:0.23057128)0.968000:0.08225276)0.999000:0.16281133)0.620000:0.04435388,((((684373764:0.41695035,(Ptor35
851:0.14314728,847028803:0.25144789)0.999000:0.16224372)1.000000:0.29623136,(443713927:0.30040300,PdumPKD12:0.32679539)0.877000:0.07372408)0.925000:0.04450313,(919074653:0.29187507,(918302224:0.33005068,(((67647
4191:0.19649822,(871215788:0.26131608,(871254812:0.08180970,908400046:0.16891654)0.974000:0.07186542)0.997000:0.09958001)0.672000:0.03012716,(1207917521:0.15585185,762156788:0.17559308)0.994000:0.07828050)0.98400
0:0.06458104)0.924000:0.05465347)0.739000:0.04528112)0.967000:0.06840323)1.000000:0.17054690)0.903000:0.03864735,((1229159547:0.38081932,(148539638:0.17633904,1229145338:0.17803396)1.000000:0.24062450)1.000000:0.19
722320,(919098890:0.42477205,(1139824617:0.47001959,(PdumPKD916:0.42003373,919082289:0.38801457)0.782000:0.06051749)0.157000:0.03080113)0.977000:0.08091023)0.993000:0.10452085)0.863000:0.03101696,(828228329:0.5
0464164,(1005450057:0.12902735,1191059136:0.23897853)0.999000:0.15976865)1.000000:0.27609605)0.539000:0.04355548,((156406861:0.61525353,((999987444:0.33699027,1005458745:0.38492367)1.000000:0.22004887,(999981318:
0.40709024,1005492071:0.36713491)0.951000:0.09687902)0.938000:0.07841521)0.959000:0.07540587,(828232853:0.80354299,1005480473:0.37193708)0.987000:0.15589845)0.511000:0.02789542)0.814000:0.04305091)0.577000:0.02996
031)0.689000:0.02974031)0.041000:0.02512348,(156394387:0.29181953,1005475249:0.35836211)1.000000:0.26434384)0.933000:0.04419205,((260785879:0.41322250,(260798008:0.47523085,1126169400:0.72541960)0.933000:0.124393
03)0.967000:0.08847455,(1101394602:1.01320033,918996724:0.49838543)0.978000:0.15010416)0.952000:0.05743276)1.000000:0.17901912)0.175000:0.05496840)0.483000:0.05835449);

Code

Arduino macros

1. Switch_motor.ino

This script was written by David A. Mellis and adapted by the author of this study to control the shaft-less motor. The Arduino circuit used was assembled as shown in **Figure 0-2**A. Only the relevant snippet of code is shown.

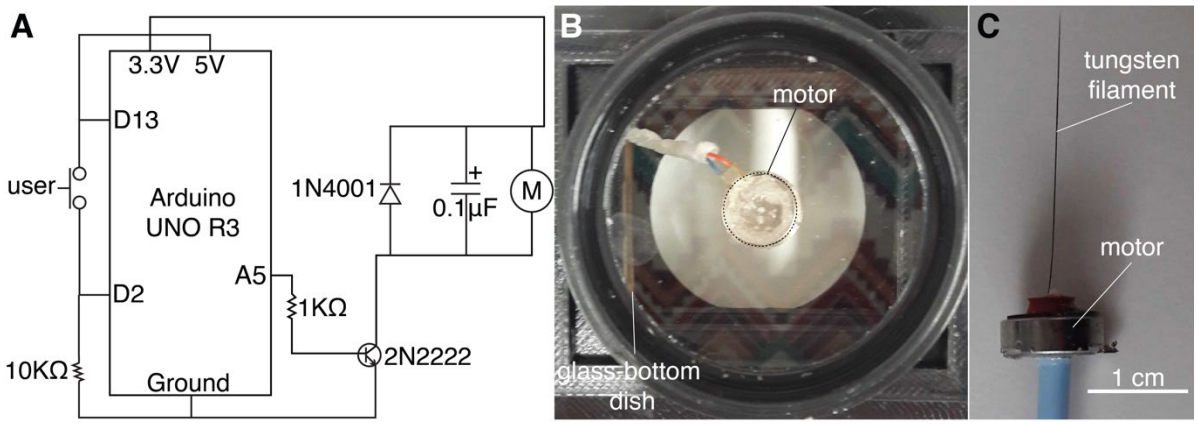


Figure 0-2 Circuit diagram and motor-dish setup. (**A**) Circuit diagram showing motor (M) and button switch connection, redrawn from original source[25]. (**B**) Snapshot showing position of motor glued to the glass-bottom dish.(**C**) Tungsten filament attached to a shaftless motor used for stimulation.

A switch button was connected to the Arduino UNO that controlled whether the motor attached to a glass-bottom dish (**Figure 0-2**B) was in the `HIGH` state or `LOW`. Whenever the button was pushed, it turned the "state" variable to `HIGH`.

This allows to set the motor to the maximum analog value (255) for 100ms, after which it is again turned off:

```
digitalWrite(outPin, state);
  if(state == HIGH){
     analogWrite(motorPin,255);
     delay(100);
     analogWrite(motorPin,0);
  }
  previous = reading;
```

2. Motor-control-script.ino

[25]Original circuit obtained from this source:
http://www.learningaboutelectronics.com/Articles/Vibration-motor-circuit.php

This script is used to control the activation of a motor-filament tool at different speeds (**Figure 0-2**C). The Arduino circuit used is shown in **Figure 0-2**A. Only the relevant snippets of code are shown. The full script is available at GitHub[26].

This code is based on `Switch_motor.ino`, but it also includes the possibility to modify the time delay during which the motor is vibrating. This is how the filament speed is controlled. A list of delay values is stored in a vector (`myDelayVals`) and used to loop through the values to stimulate a larva with a varied range of intensities. However, for most experiments only one value per recording was used. A 6 sec resting period is included between each delay value (i.e. each stimulation trial). The following snippet encodes the steps mentioned above:

```
  if(state == HIGH && LOOP == 0){
      Size=(sizeof(myDelayVals)/sizeof(int)); // this variable store the
number of delay values to be used.
      delay(4000);  //An initial delay before starting the stimulation
     for(int i=0; i<Size; i++ ){
       analogWrite(motorPin,255);  //0-255 is the dynamic range of the
motor. The highest value is used (255).
       delay(myDelayVals[i]);  //Leave the motor on at the defined value
for the time specified in the myDelayVals vector.
       analogWrite(motorPin,0); //Switch off the motor
       delay(6000); //Rest time before the next stimulation value.
     }
     delay(1000);
     LOOP=1;   // Flag variable needed to stop the program.
   }
```

FiJi/ImageJ macros

3. FigS1_StartleFreelySwimming.ijm

This macro was written to extract the speed and other parameters from freely swimming larvae (recorded from above) startled with a motor attached to the container dish. The input data are time-lapse recordings of freely swimming larvae startled at a defined point in time. The larva density should not be too high to be able to isolate instances of the startle response in individual larvae. The number of pixels per larvae has to be high enough to detect differences in area due to parapodial elevation. The frame of stimulus onset for each video has to be registered in a separate

[26] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Motor-control-script.ino

table for further processing (available at GitHub[27]).Only relevant snippets of code are shown here. The full script is available at GitHub[28].

After assembling the list of files per directory and setting up the output directories, the videos are processed to remove non-moving objects and background features of the arena:

```
//{Removing non-moving elements from stack
run("Invert", "stack");
run("Z Project...","projection=[Average Intensity]");
selectWindow(img_title);
selectWindow("AVG_"+img_title);
imageCalculator("Subtract stack", img_title,"AVG_"+img_title);
run("Smooth", "stack");
selectWindow("AVG_"+img_title);
close();
```

Then, the user has to specify rectangular ROIs that include one and only one larva across and interval of the video that includes the point of stimulation:

```
do{
//Finding the ROI where a single larva displays the startle response, as
well as prior and post stimulus behavior.
     selectWindow(img_title);
     run("Duplicate...", "title=1 duplicate");
     setTool("rectangle");
     waitForUser("Set rectangle in region with startled larva");   //The
user has to select the ROI to measure. Only choose rectangle ROIs that
during the whole behavior only include one larva.

     Dialog.create("ROIOfInterest");
     ROI="untitled";
     Dialog.addString("Title:",ROI );
     Dialog.show();
     ROI = Dialog.getString();
     Roi.setName(ROI);
     roiManager("Add");
     run("Crop");
```

The larva in the ROI is thresholded. The user is prompted again to select the interval where the measurements will be performed:

---

[27]https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip (File: File_info15fps.txt)
[28]https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/FigS1_StartleFreelySwimming.ijm

```
    //Determine where to cut the stack so that a single track is left and
that it includes also the startle part.
    waitForUser("Scroll through video to detect beginning and end of
behavior. Only one larva visible");
    Dialog.create("Beginning and End of Startle behavior");
    Dialog.addNumber("Beginning:", 512);
    Dialog.addNumber("End:", 512);
    Dialog.show();
    begin = Dialog.getNumber();
    end = Dialog.getNumber();
    run("Slice Keeper", "first="+begin+" last="+end+" increment=1");
    rename("video");
    //tracking
    mintracklength=(end-begin-((end-begin)/2));
    print(mintracklength);
```

The following snippet implements the object tracker Mtrack2[29]. It extracts the XY coordinates of the thresholded across the selected frame interval:

```
run("Set Measurements...","redirect=None decimal=3");
run("Set Scale...", "distance=0 known=0 pixel=1 unit=pixel");
run("MTrack2 ", "minimum=200 maximum=999999 maximum_=20
minimum_="+mintracklength+" display show save save=" +fullpathresults);
//The Mtrack parameters have to be adjusted depending on the source video.
```

Finally, the particle analysis tool is used to measure the area of the thresholded area across the frame interval:

```
run("Set Measurements...", "bounding fit shape feret's redirect=None
decimal=3");
run("Analyze Particles...", "size=30000-100000 show=[Bare Outlines] exclude
clear include summarize stack");
selectWindow("Summary of video");
```

[29] Nico Stuurman & Johannes Schindelin: https://imagej.net/MTrack2

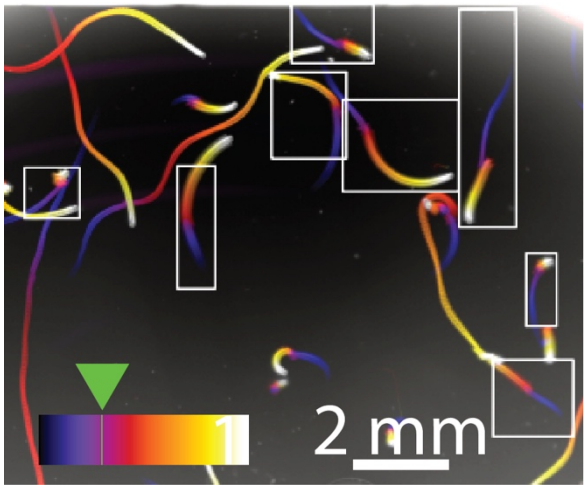


Figure 0-3 Example of video analyzed with the FigS1_StartleFreelySwimming.ijm macro. Maximum intensity projection of stack projection color-coded for time (fire LUT, from black to white). Green arrowhead and line show the stimulus onset mark (in the purple). The 8 ROIs analyzed in this video are marked with white rectangles.

4. CiliaLengthMeasurement.ijm

This macro was written to measure the length of selected cilia in the episphere of wildtype and mutant larvae in a batch of videos taken with DIC microscopy. The larvae can move slightly during the recording, but the measurement has to be only done based on references seen in a single frame.- The identity of the cilia has to be known beforehand.

The macro requires a folder path that contains all the videos to be analyzed. The scale of each video has to be already in µm. The ROI manager has to be empty before the macro is run. If the script is interrupted before processing all folders, the user has to make sure to first empty the ROI manager before running the macro again. The output file is a table with the list of measurements for each cilium and for each video. The user is advised to save the ROI table for future reference. Only relevant snippets of the code are shown here. The full code is available upon request.

As a first step the macro requires the user to specify the input and output folders. If the randomized version of the script is used, the list of files is sent to the function `shuffle`[30]. The name of the files should not reveal the genotype, or the randomization will not be useful.

The script then will remove any undesired measurements settings before opening the first video in the input folder. After this is done, the first step of the macro is to set the visualization conditions that will be used for measuring the length:

```
run("Brightness/Contrast...");
run("Enhance Contrast", "saturated=0.35");
run("In [+]");//Set the same zoom for all videos.
run("In [+]");
run("In [+]");
```

[30] Obtained from this source: https://imagej.nih.gov/ij/macros/examples/RandomizeArray.txt

Next, the polyline tool is set and the user is asked to measure the cilium of interest and give the name of the cilium measured:

```
do{
     setTool("polyline");
     waitForUser("Draw a line on the cilium to be measured, from outer
cuticle to tip");
     Dialog.create("CiliumMeasured");
     ROI="untitled";
     Dialog.addString("Title:",ROI );
     Dialog.show();
     ROI = Dialog.getString();
```

The name given will be appended to the name of the video file and this will be used as the whole name of the ROI. The ROI will be added to the manager:

```
     ROIname=ROI+"_"+img_titleEd;
     Roi.setName(ROIname);
     roiManager("Add");
     roiManager("Select", ind);
     roiManager("Measure");
```

Finally, the stack will be scrolled to the end to give chance to the user to check if more cilia need to be measured in the video currently opened. Afterwards, the user will decide if this is the case, in which case the Macro will go back to the measurement section. Otherwise, it will close the video and open the next one in line:

```
     Dialog.create("More Cilia?");
     selectWindow(img_title);
     for(i=0;i<nSlices;i++) {
     run("Next Slice [>]");
          wait(100);
    }
     Dialog.addCheckbox("Check box if more cilia are to be measured in the
present stack", false);
     Dialog.show();
     MoreROIs = Dialog.getCheckbox();
     print(MoreROIs);
     ind++;
     }while(MoreROIs);
run("Close All");
```

To properly name the measurements in the "Results" window the following plugin has to be run:

```
runMacro("/ebio/ag-jekely/share/Luis/programs/Fiji/Fiji.app/plugins/Analyze/Rename_Labels_in_Results_Table.txt");
```

The macro has to be included in the plugins list as a text file that includes the following code[31]:

```
macro "Rename Labels in Results Table" {
  for(i=0; i<nResults; i++) {
    oldLabel = getResultLabel(i);
    delimiter = indexOf(oldLabel, ":");
    newLabel = substring(oldLabel, delimiter+1);
    setResult("Label", i, newLabel);
  }
}
```

5. Fig3_FigS3_Measure-intensityvaluechanges.ijm

This macro was written to batch-process the videos obtained from calcium imaging experiments described in Chapter 2. It is used to extract pixel values within user-defined ROIs from the green and red channels. Each green/red channel stack has to be previously aligned using any competent registration method. The red and green channel stacks have to be stored in separate folders. The output files are text files with the raw intensity values.

Only relevant snippets of the code are shown here. The full macro was deposited in GitHub[32].

After opening both channel stacks and adjusting the LUTs, the ROIs (in this case the cell shapes) have to be defined using the red channel (with the tdTomato signal). The polygon tool is used for that purpose and a text box appears where the user can enter the name of the ROI (if the same cell will be analyzed in different videos, the name of the cell has to be written in the same way in all videos):

```
do{
     selectWindow("video");
     setTool("polygon");
     waitForUser("Set ellipse in region for measuring intensity");
     Cell="untitled";
     Dialog.create("CellOfInterest");
     Dialog.addString("Title:",Cell );
     Dialog.show();
     Cell=Dialog.getString();
```

---

[31] The code was written by Tiago Ferreira as referenced from this forum: https://stackoverflow.com/questions/20800207/imagej-how-to-put-label-names-into-results-table-generated-by-roi-manager

[32] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig3_FigS3_Measure-intensityvaluechanges.ijm

```
ROIname=img_title+"_"+Cell;
Roi.setName(ROIname);
roiManager("Add");
```

The file name will be added as a prefix to the name of the ROI, after which it will be added to the ROI manager. The pixel values will be measured using the Z-axis intensity profile tool:

```
run("Plot Z-axis Profile");
```

Next, the pixel values are extracted from the green channel stack. The same ROI for measuring pixel values in the red channel is used in this case. For finding the corresponding ROI a function called `findRoiWithName` is used:

```
selectWindow("videoGC");
index=findRoiWithName(ROIname);
roiManager("select",index);
run("Clear Results");
run("Plot Z-axis Profile");

/*Function written by oburri:
 http://forum.imagej.net/t/selecting-roi-based-on-name/3809/2
*/
function findRoiWithName(roiName) {
    nR = roiManager("Count");

    for(i=0; i<nR; i++) {
        roiManager("Select", i);
        rName = Roi.getName();
        if(matches(rName, roiName)) {
            return i;
        }
    }
    return -1;
}
```

Each cell has to be analyzed one at a time. Therefore, the macro may need to be run on the same video more than once if there are more than one cell of interest. Moreover, a measurement of the background has to be extracted using the same ROI defined for the cell. In this case, the ROI has to be called with the name of the corresponding cell and with the word 'bckg'.

The final snippet in the code has the purpose to scroll through the green channel stack, so the user can assess if there is any additional cell that needs to be analyzed:

```
for(i=0;i<nSlices;i++) {
```

```
                run("Next Slice [>]");
                wait(100);
              }
      Dialog.create("More ROIs");
          Dialog.addCheckbox("Check box if more rois in the present video",
false);
          Dialog.show();
          Satis = Dialog.getCheckbox();
          print(Satis);
      }while(Satis);
```

R scripts

1. FigS1_StartledataFreelySwimming.R

This script reads in the output from `FigS1_StartleFreelySwimming.ijm` to calculate the change in speed and area of individual larva relative to an average prior to stimulus start. A separate input file storing the beginning and end frames for each ROI measured, as well as the stimulus start is needed to run the script. The dataset used as input for this script is available at GitHub[33].

Table 0-8 Example input table needed to run script `FigS1_StartledataFreelySwimming.R`

Experiment_ID	ROI_ID	Start_frame	End_Frame	Stimulus_Start_ROI-dependent	Track_length
Experiment-75	ROI1	16	123	28	108
Experiment-75	ROI2	24	140	20	117
Experiment-77	ROI1	19	116	29	98
Experiment-77	ROI2	38	121	10	84
Experiment-79	ROI1	41	136	10	96
Experiment-80	ROI1	37	117	12	81
Experiment-81	ROI1	38	117	10	80

Only relevant snippets of the code are shown below. The full script is available at GitHub[34].

The code can be divided into two sections. The first part processes the output from Mtrack2. For that, a matrix is created (`FinalSpeed`) to store the final speed values for each ROIs to be processed. Then, the files storing the XY coordinates are read using a set of nested `for` loops:

```
for(j in 1:length(FilePropList$Experiment_ID)){
  for(k in 1:length(MtrackFile)){
```

[33] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip (Folder: Freely_swimming_analysis)

[34] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/FigS1_StartledataFreelySwimming.R

```
    MFullpath<-paste(MtrackPath,MtrackFile[k],sep="");
if(grepl(FilePropList[j,"Experiment_ID"],MtrackFile[k])&&grepl(FilePropList
[j,"ROI_ID"],MtrackFile[k])){
      numbRows<-FilePropList[j,"Track_length"]
      Tracks<-read.table(MFullpath,skip=2, nrows = numbRows,header = F)
```

Once the coordinates are read, the pitagoraean distance is calculated. This value is used to measure the speed (defined as the change in position from one frame to the next).

```
      Speed<-list();
      OfSet<-1 #Important value to decide how rough or smooth is the speed
calculation.
for(i in 1:(length(Tracks$V1)-OfSet)){
        Vx=Tracks$V2[i+OfSet]-Tracks$V2[i];
        VxP=Vx/0.1077  #0.1077 this value is the distance in pixels.
        Vy=Tracks$V3[i+OfSet]-Tracks$V3[i];
        VyP=Vy/0.1077
        Speed[i]<-(sqrt(('^' (VxP,2))+('^' (VyP,2))))/(1000*OfSet*0.06)
```

The speed thus calculated is stored in the matrix `FinalSpeed`. The frames at which stimulus starts is set as frame 0 (all tracks have to be cropped to the same number of frames for this to work):

```
      RelativeFrame<-Tracks$V1-
FilePropList[j,"Stimulus_Start_ROI.dependent"]
      Speed<-c(rep('NA',OfSet),Speed);
      Tracks$Speed<-Speed;
      Tracks$RelativeFrame<-RelativeFrame;
      FinalSpeed[,j]<-Tracks[which(Tracks$RelativeFrame > -10 &
Tracks$RelativeFrame < 70),"Speed"]  # This has to be adjusted depending on
the length of the video analyzed.
```

The stored speed values are normalized relative to the mean speed prior to stimulus:

```
NormFinal<-FinalSpeed
  for(k in 1:(ncol(FinalSpeed)-1)){
    PriorAvg<-which(FinalSpeed$time < 0)
    MeanVal<-mean(FinalSpeed[PriorAvg,k],na.rm=T) #Calculate the arithmetic
mean of the speed values prior to stimulation.
    NormFinal[,k]<-FinalSpeed[,k]-MeanVal
  }
```

The second section of this script measures the change in area. As in the previous section, a matrix is created to store the final area values (`SumFinal`). The area values calculated in FiJi are extracted in a similar manner as the XY coordinates, and then stored in `SumFinal`:

```
if(grepl(FilePropList[j,"Experiment_ID"],SummFile[k])&&grepl(FilePropList[j
,"ROI_ID"],SummFile[k])){
      Area<-read.table(SFullpath,sep="¥t",header=F,colClasses =
c(rep("NULL",3),rep("numeric",1),rep("NULL",11)), nrows = numbRows,skip =
1)
      Beg<-1-(FilePropList[j,"Stimulus_Start_ROI.dependent"])
      End<-length(Area$V4)-(FilePropList[j,"Stimulus_Start_ROI.dependent"])
      ReFrame<-seq(Beg,End)
      Area$ReFrame<-ReFrame
      SumFinal[,j]<-Area[which(Area$ReFrame > -10 & Area$ReFrame <
70),"V4"]
```

The areas read in are standardized to the mean area prior to stimulation:

```
NormSumFinal<-SumFinal
for(k in 1:(ncol(SumFinal)-1)){
  PriorAvg<-which(SumFinal$time < 0)
  MeanVal<-mean(SumFinal[PriorAvg,k],na.rm=T) #Calculate the arithmetic
mean of the speed values prior to stimulation.
  print(MeanVal);
  NormSumFinal[,k]<-SumFinal[,k]-MeanVal
}
```

2. FigStartleKynetics.R

This script was written to process and plot the data collected from the startle response recordings in tethered animals. It takes as input the table that registered the relevant event and measurements for each video. The response type (see Results Chapter 1) was precomputed in Excel and also included in the input table (**Table 0-9**).

Table 0-9 Fragment of table file used as input for FigStartleKynetics.R. Data are shown for four data entries. The table is split into three parts to ease readability. Variable names were modified from the ones used in the script for clarity.

File name	Larva ID	Stimulation side	Maximal filament displacement (before prototroch closure) (µm)	Start of maximal filament displacement prior to prototroch closure (frame no.)	Stimulus start (frame no.)	Onset prototroch closure (frame no.)	Onset bodytroch closure (frame no.)
Experiment-77.czi	L030317	Anterior	88.3	642	632	646	646
Experiment-78.czi	L030317	Anterior	38.2	721	712	730	731
Experiment-79.czi	L030317	Anterior	64.7	674	666	678	679
Experiment-80.czi	L030317	Anterior	114.3	631	624	633	635
Maximal filament displacement (before maximal parapodial elevation) (µm)	**Start of maximal filament displacement prior to maximal elevation(frame no.)**	**Onset Parapodium 1st segment (main body side)**	**Onset Parapodium 1st segment (opposite body side)**	**Onset Parapodium 2nd segment (main body side)**	**Onset Parapodium 3rd segment (main body side)**	**Maximal parapodial elevation angle**	**Relative elevation to maximum angle observed**
84.8	649	654	656	654	658	64.6	0.651866801
38.2	721	756	736	NA	NA	5.3	0.053481332
110.9	679	689	683	690	NA	17.9	0.180625631
213.8	636	638	638	638	639	86.7	0.874873865
Time of maximal elevation (frame no.)	**Distance probe (µm)**	**Total frame no.**	**Glued side**	**Frame rate zeiss from metadata (fps)**	**Distance probe from anterior (µm)**	**Time per Frame**	**Response Type**
685	119	1960	Ventral	295	119	3.38983051	ClosureWideE
814	120.1	1960	Ventral	295	120.1	3.38983051	ClosureLowE
738	120.1	1891	Ventral	281.9326128	120.1	3.54694688	ClosureLowE
659	120.1	1874	Ventral	281.9326128	120.1	3.54694688	ClosureWideE

Only relevant snippets of the code are shown. An abridged version of the code is available at GitHub[35]. Additions to it are detailed below.

The script reads in the table and calculates latency values for ciliary band closures and parapodial elevation, and adds them to the main table as new columns:

```
#Calculating Latency of max.parapodia Elevation and transforming to ms.

DurationRest2Elevation<-(TableResults$Beginning.Max.Elev1stpara-
TableResults$Beginning.1stPElevMAIN)*TableResults$TimeperFrame
#adding variable as a new column.
TableResults$DurationRest2Elevation<-DurationRest2Elevation

#Calculating Latency of closure start and transforming to ms.

LatencyClosure<-(TableResults$Beginning.Closure.PrototrochMAIN-
TableResults$Beginning.stimul)*TableResults$TimeperFrame
TableResults$LatencyClosure<-LatencyClosure
#Bodytroch
LCloBody<-(TableResults$Beginning.Closure.BodytrochsMAIN-
TableResults$Beginning.stimul)*TableResults$TimeperFrame
TableResults$LCloBody<-LCloBody

#Calculating Latency of Elevation start and transforming to ms.

LatencyElev<-(TableResults$Beginning.1stPElevMAIN-
TableResults$Beginning.stimul)*TableResults$TimeperFrame
```

[35] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig1-S1_Analysis_startledataWT.R

```
TableResults$LatencyElev<-LatencyElev
```

Probe speed prior to parapodial elevation is also calculated in this section:

```
ProbespeedP<-
(TableResults$Max..frame2frame.Displacement.tip..before.PARAPODIAmaxElev..d
iagonal...um.)/TableResults$TimeperFrame
TableResults$ProbespeedP<-ProbespeedP
```

Next, the variables storing the delays between various events are calculated. These include the delay between parapodial elevation and ciliary band closures, or the delay in frames between elevation of parapodia on different segments and on different sides:

```
#Calculating delay of Elevation to prototroch closure

Elev2Proto<-(TableResults$Beginning.1stPElevMAIN-
TableResults$Beginning.Closure.PrototrochMAIN)
TableResults$Elev2Proto<-Elev2Proto

#Calculating delay of Elevation to bodytroch closure

Elev2Body<-(TableResults$Beginning.1stPElevMAIN-
TableResults$Beginning.Closure.BodytrochsMAIN)
TableResults$Elev2Body<-Elev2Body

#Calculating absolute delay Left-Right Elevation 1st parapodia Elevation

LRElev1stPa<-abs(TableResults$Beginning.1stPElevMAIN-
TableResults$Beginning.1stPElev_Oppo)
TableResults$LRElev1stPa<-LRElev1stPa

#Calculating delay segment coordination between segment 1 and segment 3.

Coord12<-(TableResults$Beginning.2ndPElev-
TableResults$Beginning.1stPElevMAIN)
Coord13<-(TableResults$Beginning.3rdPElev-
TableResults$Beginning.1stPElevMAIN)
TableResults$Coord12<-Coord12
TableResults$Coord13<-Coord13

#Calculating delay of closures of bodytrochs relative to prototroch

Beginnbody2Proto<-(TableResults$Beginning.Closure.BodytrochsMAIN-
TableResults$Beginning.Closure.PrototrochMAIN)#*TableResults$TimeperFrame
```

```
#Calculating delay of LR bodytroch closure

ProtoTrochLR<-(TableResults$Beginning.Closure.PrototrochMAIN-
TableResults$Beginning.Closure.PrototrochOPPOSITE2stimulation)
TableResults$ProtoTrochLR<-ProtoTrochLR

#Calculating delay of LR bodytroch closure

BodyTrochLR<-(TableResults$Beginning.Closure.BodytrochsMAIN-
TableResults$Beginning.Closure.BodytrochOPPOSITE2stimulation)
TableResults$BodyTrochLR<-BodyTrochLR

#Calculating delay of Elevation to prototroch closure

Elev2Proto<-(TableResults$Beginning.1stPElevMAIN-
TableResults$Beginning.Closure.PrototrochMAIN)
TableResults$Elev2Proto<-Elev2Proto

#Calculating delay of prototroch to bodytroch closure

Body2Proto<-(TableResults$Beginning.Closure.BodytrochsMAIN-
TableResults$Beginning.Closure.PrototrochMAIN)
TableResults$Body2Proto<-Body2Proto
```

To sort values according to the stimulation distance and to probe speed (for stacked bar plots), the raw values were rounded to the closest hundred:

```
RoundDist<-
paste(round(TableResults$Distance.probe..um./100)*100,"um",sep="")
TableResults$RoundDist<-RoundDist
#Rounding probe speed.
RoundSpeedP<-round(TableResults$ProbespeedP/20)*20
TableResults$RoundSpeedP<-RoundSpeedP
```

Duration of stimulation, of closures, of the elevation response, etc, are also calculated with the input data:

```
#Duration stimulation

Durastim<-(TableResults$End.stimul-
TableResults$Beginning.stimul)*TableResults$TimeperFrame
TableResults$Durastim<-Durastim

###Duration Prototroch closures##
```

```
DurationClosuresP<-(TableResults$End.Closure.Prototroch-
TableResults$Beginning.Closure.PrototrochMAIN)*TableResults$TimeperFrame
TableResults$DurationClosuresP<-DurationClosuresP

###Duration Bodytroch closures##

DurationClosuresB<-
(TableResults$End.Closure.Bodytrochs..until.Telotroch.is.bearing.again.-
TableResults$Beginning.Closure.BodytrochsMAIN)*TableResults$TimeperFrame
TableResults$DurationClosuresB<-DurationClosuresB

###Duration prototroch closure to duration of stimulation

RelativeDuraClosP<-(TableResults$End.Closure.Prototroch-
TableResults$End.stimul)
TableResults$RelativeDuraClosP<-RelativeDuraClosP

#Overall duration of elevation response (summing up delay and time to max
elevation)

OveraElev<-(TableResults$DurationRest2Elevation+TableResults$LatencyElev)
TableResults$OveraElev<-OveraElev
#Full duration Elevation
FullElevdura<-(TableResults$EndMax.Elev1stpara-
TableResults$Beginning.1stPElevMAIN)*TableResults$TimeperFrame
TableResults$FullElevdura<-FullElevdura

#Duration max elevation

DuraMaxElev<-(TableResults$EndMax.Elev1stpara-
TableResults$Beginning.Max.Elev1stpara)*TableResults$TimeperFrame
TableResults$DuraMaxElev<-DuraMaxElev
```

Subgroups were assembled for plotting purposes:

```
#Grouping all experiments stimulated 100µm from anterior

Head100<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head")

#Grouping all experiments stimulated 100µm from anterior classified as low
elevation.

Head100LowAlev<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head" &
```

```
TableResults$Relative2max_Elevation < cut_off[1] &
TableResults$Relative2max_Elevation > 0.001)

#Grouping all experiments stimulated 100µm from anterior classified as wide
elevation.

Head100HighElev<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Relative2max_Elevation > cut_off[1])

#Grouping all experiments stimulated 100µm from posterior classified as
wide elevation.

Pygid100HighElev<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Pygid" &
TableResults$Relative2max_Elevation > cut_off[1])

#Grouping all experiments stimulated from the side classified as wide
elevation.

SideHigh<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Side" &
TableResults$Relative2max_Elevation > cut_off[1])

#Grouping all experiments based on stimulation side.

Head<-which(TableResults$Closest.body.part.to.probe == "Head")
Pygid<-which(TableResults$Closest.body.part.to.probe == "Pygid")
Side<-which(TableResults$Closest.body.part.to.probe == "Side")
HP<-which(TableResults$Closest.body.part.to.probe == "Head" |
TableResults$Closest.body.part.to.probe == "Pygid")
```

Another set of groupings was made to plot the titration of the response dependent on stimulation distance:

```
#Grouping all experiments stimulated from the posterior end more than 200um
from anterior end.

TailRound<-which(TableResults$Distance.probe_from_head > 200 &
TableResults$Closest.body.part.to.probe =="Pygid")

TableResults[TailRound,"RoundDist"]<-"350um"
TableResults[PygTit,"RoundDist"]<-"350umRangePyg"
```

```
##This group is for selecting a subsample of the anterior 100 category that
can be comparable in number of observations and probespeeds to those of the
200-500 categories (the 100R category in  ).

Head100Tit<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$ProbespeedP > 19.5 & TableResults$ProbespeedP < 86.7)
Head100TitR<-sample(Head100Tit,length(TableResults[Head500,"ProbespeedP"]))
TableResults[Head100TitR,"RoundDist"]<-"100umR"

Head100TitRHEle <- which(TableResults$RoundDist == "100umR" &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Relative2max_Elevation > cut_off[1] &
TableResults$Distance.probe..um. < 150)

PygTit<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Pygid" &
TableResults$ProbespeedP > 19.5 & TableResults$ProbespeedP < 86.7)

#Different groupings based on the stimulation distance.

Head200<-which(TableResults$Distance.probe..um. < 250 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Distance.probe..um. > 150)

Head200High<-which(TableResults$Distance.probe..um. < 250 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Distance.probe..um. > 150 &
TableResults$Relative2max_Elevation > cut_off[1])

Head300<-which(TableResults$Distance.probe..um. < 350 &
TableResults$Closest.body.part.to.probe == "Head"&
TableResults$Distance.probe..um. > 250)

Head300High<-which(TableResults$Distance.probe..um. < 350 &
TableResults$Closest.body.part.to.probe == "Head"&
TableResults$Distance.probe..um. > 250 &
TableResults$Relative2max_Elevation > cut_off[1])

Head400<-which(TableResults$Distance.probe..um. < 450 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Distance.probe..um. > 350)

Head400High<-which(TableResults$Distance.probe..um. < 450 &
TableResults$Closest.body.part.to.probe == "Head" &
```

```
TableResults$Distance.probe..um. > 350 &
TableResults$Relative2max_Elevation > cut_off[1])

Head500<-which(TableResults$Distance.probe..um. < 550 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Distance.probe..um. > 450)
```

Additional groupings were created after a first analysis, including groups based on response type such as those including stimulation groups 100 µm from the anterior that showed no elevation but prototroch closures (Head100NoElevClosP):

```
#Group anterior stimulation samples showing no elevation + prototroch
closures

Head100NoElevClosP<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Relative2max_Elevation == 0 & TableResults$Closure_Prototroch
== "Closure")

#Group anterior stimulation samples showing low-angle elevation +
prototroch closures

Head100LowAlevClosP<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head"  &
TableResults$Closure_Prototroch == "Closure" &
TableResults$Relative2max_Elevation < cut_off[1] &
TableResults$Relative2max_Elevation > 0.001)

#Group anterior stimulation samples showing wide-angle elevation +
prototroch closures

Head100HighElevClosP<-which(TableResults$Distance.probe..um. < 150 &
TableResults$Closest.body.part.to.probe == "Head" &
TableResults$Relative2max_Elevation > cut_off[1] &
TableResults$Closure_Prototroch == "Closure")
```

The stacked bar plots such as that shown in **Figure 1-3**E are generated by creating a table with only a subset of the categories from the full input table (speedvsResptypeH100). This is used as the input to a function that creates proportions of defined categories (prop.table). The output value is transformed into percentages:

```
bins=seq(0,80,by=5)
```

```
PerPlot<-
100*prop.table(table(speedvsResptypeH100$ElevationParapodia,cut(speedvsResp
typeH100$ProbespeedP,breaks=bins,labels=as.character(seq(2.5,77.5,by=5)))),
2)
```

Boxplots, scatterplots and histograms shown across Chapter 1 are generated using the groupings created before. Statistical tests between distributions were run in the script using the built-in Kolmogorov Smirnov function `ks.test`. In most cases a two-sided alternative was used except in cases where noted. This is an example comparing the distribution of the latency of prototroch closures (`LatencyClosure`) when there is not concomitant elevation (`Head100NoElevClosP`) against the distribution when there is both wide elevation and closures (`Head100HighElevClosP`):

```
ks.test(TableResults[Head100NoElevClosP,"LatencyClosure"],TableResults[Head
100HighElevClosP,"LatencyClosure"],alternative = "two.sided")
```

To evaluate the correlation of two variables, the function `cor.test` was implemented using the Spearman's method. The following example shows the test performed to test the correlation between the maximal probe speed (`ProbespeedP`) and the duration of the stimulus(`Durastim`):

```
cor.test(TableResults$ProbespeedP,TableResults$Durastim,method =
"spearman",alternative = "greater")
```

3. Startlethresholdestimation.R

This script uses a Finite mixture model to compute the threshold separating low from wide angle parapodial elevation. It only works after loading the data into the script `FigStartleKynetics.R`. The code is taken with modifications from Marc Choisy (Trang et al. 2015)[36]. Only relevant snippets are shown here.

The full code is available at GitHub[37].

First, a group is defined to include only experiments of stimulation 100 μm from anterior (only those experiments where there was an actual elevation response):

```
AbHeCa1<-which(abridgedtableHead100$Relative2max_extension >0 )

AbHistVals<-abridgedtableHead100[AbHeCa1,"Relative2max_extension"]
```

[36] An explanation on how to implement the model can be accessed here: http://marcchoisy.free.fr/fmm/index.html.
[37] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig1G_Startlethresholdestimation_paper.R.

Then, the parameters (μ σ) of the Finite Mixture Model are estimated using an expectation maximization algorithm (`em` function) assuming normality of the data:

```
(HiVal_out <- em(AbHistVals,"normal","normal"))
```

The confidence interval of the parameters is the computed using the `confint` function:

```
###Estimating the 95% confidence interval####
confint(HiVal_out,nb=10,level=.95)
```

A cut-off value is estimated with the modeled distribution using the `cutoff` function:

```
(cut_off<-cutoff(HiVal_out))
```

This value is used in `FigStartleKynetics.R` to separate events of low angle elevation from wide angle elevation.

4. CiliaStats.R

This R script extracts the cilia length measurements obtained with the Macro `CiliaLengthMeasurement.ijm` from the output files and groups them by cell and by phenotype. The script needs as input a folder with all the text files to be analyzed in a tabulated format. It is important that the name of each measurement contains the names of the cells and genotypes in the same format. Otherwise, they will be skipped by the script. Only relevant snippets of the code are shown. The full code is available upon request.

First, the list of files, cells and genotypes to analyze are defined:

```
Listfiles<-readLines("./filelist.txt")
CList<-c("MS1","MS2","hCR1l","hCR1r","hCR2l","hCR2r","hCR1","hCR2")
GList<-c("WT","P2T2")
```

The core of the script is based on three nested `for` loops, each scrolling through the different lists:

```
for(j in 1:length(GList)){
  for(i in 1:length(CList)){
    LabelsList<-list() # Two new lists are defined that will store the
labels…
    MesList<-list()  #…and the length measurements.
     Cell<-CList[i]
    for(k in 1:length(Listfiles)){
```

Inside the nested loops the current table file is opened, and a query string is made from the current cell and genotype in the loop. This query is searched at once in the whole table using a logic grep `grepl` function:

```
        ListMeasure<-read.table(Fullpath,header=
TRUE,sep="¥t",stringsAsFactors=F) #Make sure labels ae not imported as
factors.
    Genotype<-GList[j]
        Query<-paste(Cell,Genotype,sep = ".*")
        CellGen<-paste(Cell,Genotype,sep = "_") #This variable will be
used for naming the dataframe
        Logic<-grepl(Query,ListMeasure$Label)
```

The matching rows of labels and measurements are added to the end of the lists storing these values:

```
        NewLabels<-ListMeasure[Logic,"Label"]
        LabelsList<-append(LabelsList,NewLabels)
        NewMeasures<-ListMeasure[Logic,"Length"]
        MesList<-append(MesList,NewMeasures)
```

Finally, the list of labels and measurements are merged into a data frame with the name of the corresponding cell and genotype:

```
assign(CellGen, (do.call(rbind,Map(data.frame,Labels=LabelsList,Length=MesLi
st))))
```

An additional snippet of code has to be introduced for dealing with randomized measurements. Specifically, two lists with the names given to the files of each genotype have to be created. In this case, the files were given numbers:

```
WTlist<-c(seq(36,67,by = 1))
P2T2list<-c(seq(1,35,by = 1))
```

An additional series of lines are added in the last nested for loop to identify and assign genotype to each measurement:

```
ListMeasure$Genotype<-""
        WTindx<-grepl(paste(paste("_",WTlist,".zvi",sep =''),collapse =
'|'),ListMeasure$Label)
        P2T2indx<-grepl(paste(paste("_",P2T2list,".zvi",sep =''),collapse =
'|'),ListMeasure$Label)
        ListMeasure[WTindx,"Genotype"]<-"WT"
        ListMeasure[P2T2indx,"Genotype"]<-"P2T2"
```

In this case, the search for the cell and for the genotype is performed in two lines, and the logical intersection combined into a new variable that is used to retrieve the matching rows in the data frame:

```
Logic1<-grepl(Cell,ListMeasure$Label)
Logic2<-grepl(Genotype,ListMeasure$Genotype)
Lresult<-Logic1&Logic2
NewLabels<-ListMeasure[Lresult,"Label"]
```

5. Fig2_Analysis_CaimagingCR.R

This R script calculates the metric $\Delta R/R$ from the raw pixel values obtained with the macro `Fig3_FigS3_Measure-intensityvaluechanges.ijm` from the videos recorded in the calcium imaging experiments described in Chapter 2. This metric is used to evaluate changes in fluorescence in GCaMP6s relative to a reference fluorophore (in this study tdTomato):

$$\Delta R/R = \frac{F(t)_{GCaMP6s} x F0_{tdTom}}{F0_{GCaMP6s} x F(t)_{tdTom}}$$

Where $F0$ is the average fluorescence of either of the fluorophores prior to stimulation, and $F(t)$ is the fluorescence at time t. The metric was taken from (Böhm et al. 2016). The script also aligns all the measurements to a common stimulus start and plots the mean and individual values.

The inputs of this script are as follows: 1) a single folder containing all the text files with the raw pixel values per ROI for both the red and green channels. The background text files have to be included in the same folder and named with the name of the cell and the word "bckg". The red channel files have to have the word "Tom". This naming is added during the obtention of the pixel values using the macro mentioned above. 2) A table with the name of each video from which ROI measurements were taken, and the frame at which stimulus started. 3) A file with the cells to be analyzed. The files used as input for this script are available at GitHub[38].

Only relevant snippets of the code are shown and explained here. The full script is available in GitHub[39].

Four files are needed to calculate $\Delta R/R$ for a given cell in a given recording: two files correspond to the green and red signal inside the cell, and two to the background signal collected with the same ROI shape. To find these files in the folder a set of nested `if` functions are used, which in turn are located inside two nested `for` functions that loop through the cell list and the different video files.

[38] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip (Folder: Calcium_imaging)
[39] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig2_Analysis_CaimagingCR.R

The conditional functions use regular expression pattern matching to locate the background files and to distinguish the two channel files. Once inside the `if`, the files found are read in, specifically the pixel values. These set of steps are detailed in the following snippet:

```
Listsnames<-list();
ListOC<-list();
o=1;
 for (i in 1:length(OFTable$File_name))
  {
   for(j in 1:length(CellList))
   {
     FOUND=0
     CombFC<-paste(OFTable$File_name[i],CellList[j],sep="_")
     for(k in 1:length(ResFList))
      {
       if(grepl(CombFC,ResFList[k]))
       {
           Fullpath<-paste(Resultspath,ResFList[k],sep="");
           if(grepl("Tom",ResFList[k]))
             {
             if(grepl("bckg",ResFList[k]))
             {
               Tom_bckg<-read.table(Fullpath,sep =
"¥t",header=TRUE,colClasses =
c(rep("NULL",2),rep("numeric",1),rep("NULL",1)))
             }else{
               Tom_signal<-read.table(Fullpath,sep =
"¥t",header=TRUE,colClasses =
c(rep("NULL",2),rep("numeric",1),rep("NULL",1)))

             }
            }else
            {
              if(grepl("bckg",ResFList[k]))
             {
               GC_bckg<-read.table(Fullpath,sep =
"¥t",header=TRUE,colClasses =
c(rep("NULL",2),rep("numeric",1),rep("NULL",1)))
             }else{
               GC_signal<-read.table(Fullpath,sep =
"¥t",header=TRUE,colClasses =
c(rep("NULL",2),rep("numeric",1),rep("NULL",1)))
               FOUND=1
             }
           }
```

```
        }
      }
```

Once the four list of pixel values are collected the FOUND flag variable is activated. Only then, the background pixel values are subtracted from the tdTomato and the GCaMP signal to get corrected values:

```
if(FOUND==1)
      {
        Correc_Tom=Tom_signal-Tom_bckg
        Correc_GC=GC_signal-GC_bckg
```

To calculate $F0$, first an interval of time prior to stimulation had to be defined. The lower limit of this interval was set at the half of the total pre-stimulation period. The upper limit is set at one frame prior to stimulus start. Once set, the mean pixel value in the interval is calculated:

```
noframes<-read.table(Fullpath, sep = "¥t", header=TRUE, colClasses =
c(rep("numeric", 1), rep("NULL", 3)))
LowMeanlimit=OFTable$Stimulation_Start[i]-
(trunc(OFTable$Stimulation_Start[i]/2))
        MaxMeanlimit=OFTable$Stimulation_Start[i]-1
        meanrange<-seq(LowMeanlimit, MaxMeanlimit)
        MeanTom=mean(Correc_Tom[meanrange,])
        MeanGC=mean(Correc_GC[meanrange,])
```

With these values in hand the $\Delta R/R$ is computed and stored in the dR variable (1 is substracted from the metric to bring the base line to the origin):

```
dR<-(Correc_GC*MeanTom)/(Correc_Tom*MeanGC)-1
```

The frame number is also shifted relative to the stimulus start. All the values are stored into a data frame with the name of the file and the cell using the assign function:

```
CorrFrame<-noframes-OFTable$Stimulation_Start[i]

assign(CombFC, setNames(data.frame(noframes, Tom_signal, Tom_bckg, GC_signal, GC
_bckg, dR, CorrFrame), columns))
        Listsnames[o]<-CombFC;
        o=o+1;
       }
   }
 }
```

In this manner a set of data frames per file per cell is created. To be able to merge all recordings for a given cell into a single matrix, the data frames have to be trimmed to the same size. Thus, the shortest recording dictates the length. This assessment has to be done by the user to define the number of rows (nrow) in the Final matrix:

```
Final<-data.frame(matrix(ncol=length(Listsnames),nrow=21))
#nrow:hCR1/MS1:28,hCR2:21
names(Final)<-Listsnames
```

In each column of this matrix the dR for each file is stored:

```
for (j in 1:length(CellList))
{
  for(i in 1:length(Listsnames)){
    Final[,i]<-subset(get(Listsnames[[i]])[6],get(Listsnames[[i]])[7] > -6
& get(Listsnames[[i]])[7] < 16)
  }
}
```

Next, the time units are transformed from frames to seconds using the known recording rate:

```
IntervalFrame<-0.24
CorrFrame<-seq(-5,)*IntervalFrame
```

A new matrix for each cell analyzed is created, only selecting those recordings with a defined set of stimulation levels. With this matrix per cell a mean $\Delta R/R$ can be calculated:

```
for (j in 1:length(CellList))
{
  FinXCell<-Final[,grep(CellList[j],names(Final))]
  y<-c("L21","L25","L30","L20") #Select the desired stimulation levels
(only for recordings at 0.24 time interval).
  CellMean<-rowMeans(subset(FinXCell,select=grep(paste(y,collapse =
"|"),names(FinXCell)))) #Calculate Mean across all rows (per timepoint)

assign(CellList[j],data.frame(Final$CorrFrame,subset(Final,select=grep(Cell
List[j],names(Final))),CellMean))
}
```

The individual $\Delta R/R$ values and the newly calculated mean are plotted using ggplot.

6. Fig3M_BoxplotsPredRates.R

This short script was written to perform the Wilcoxon-Pratt signed rank test on the data collected from predation experiments. This is a non-parametric statistical test used for assessing whether a common median in matched-pairs data. The script required to pre-format de predation rates in a table as the example shown here:

Predation rate	Pair_ID	Genotype
11.52	1	PKD2
12	2	PKD2
12.2666667	3	PKD2
19.5652174	4	PKD2
25.5072464	5	PKD2
11.04	1	WT
11.04	2	WT
7.46666667	3	WT
14.3478261	4	WT
12.7536232	5	WT

The predation rates were calculated as described in Chapter 4. The 'pair_ID' column associates predation rates of WT or *PKD2-1* mutant larvae from the same experiment (i.e. paired values).
The full dataset is available at GitHub[40].

The main snippet of code implements the statistical test on the input data using the `wilcoxsign_test` function. The null hypothesis is a mean difference between WT and *PKD2-1*$^{mut/mut}$ around 0. The alternative hypothesis considered is a difference actually significantly greater in the mutants compared to the wildtype larvae:

```
wilcoxsign_test(PredRates$Predation_Rate_Mutant~PredRates$Predation_Rate_WT
,alternative="greater",distribution="exact",zero.method="Pratt")
```

To correct for zero-differences present in the input data, the method proposed by Pratt (Pratt 1959) was implemented in the same function. Exact *p-values* were calculated in this case.
The full script was uploaded to GitHub[41].

[40] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/SourceDataforR.zip
(Folder: Predator_assay)
[41] https://github.com/JekelyLab/Bezares_et_al_2018/blob/master/Fig3M_BoxplotsPredRates.R

List of acronyms and abbreviations

ADPKD: Autosomal Dominant Polycystic Kidney Disease

ASW: Artificial Sea Water

bp: base pairs

CR: Collar Receptor

EM: Electron Microscopy

hr: hours

LowE: Low-angle parapodial Elevation

min: minutes

MSW: Mixed Sea Water

MS: MechanoSensor neuron

NSW: Natural Sea Water

nt: nucleotide

PBS: Phosphate Saline Buffer

PB: Penetrating Biciliated neuron

PC: PolyCystin

PU: Penetrating Uniciliated neuron

PM: Penetrating Multiciliated neuron

PTW: PBS+ Tween

PKD: Polycystic Kidney Disease

RT: Room Temperature

s: seconds

SEM: Scanning Electron Microscopy

SNR: Signal-to-Noise Ratio

sTEM: serial Transmission Electron Microscopy

WT: Wild Type

WideE: Wide-angle parapodial Elevation

WMISH: Whole Mount *In Situ* Hybridization

VNC: Ventral Nerve Cord

List of Figures

"If I have seen further it is by standing on the shoulders of Giants"

-Isaac Newton

Bibliography

Achim, K., Eling, N., Vergara, H.M., Bertucci, P.Y., Musser, J., Vopalensky, P., Brunet, T., Collier, P., Benes, V., Marioni, J.C. and Arendt, D. 2018. Whole-Body Single-Cell Sequencing Reveals Transcriptional Domains in the Annelid Larval Body. *Molecular Biology and Evolution* 35(5), pp. 1047–1062.

Achim, K., Pettit, J.-B., Saraiva, L.R., Gavriouchkina, D., Larsson, T., Arendt, D. and Marioni, J.C. 2015. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature Biotechnology* 33(5), pp. 503–509.

Ackermann, C.F. 2003. Markierung der Zellinien im Embryo von Platynereis - Deutsche Digitale Bibliothek. Doctoral dissertation.

Akerboom, J., Rivera, J.D.V., Guilbe, M.M.R., Malavé, E.C.A., Hernandez, H.H., Tian, L., Hires, S.A., Marvin, J.S., Looger, L.L. and Schreiter, E.R. 2009. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *The Journal of Biological Chemistry* 284(10), pp. 6455–6464.

Alkema, M.J., Hunter-Ensor, M., Ringstad, N. and Horvitz, H.R. 2005. Tyramine Functions independently of octopamine in the Caenorhabditis elegans nervous system. *Neuron* 46(2), pp. 247–260.

Almeda, R., van Someren Gréve, H. and Kiørboe, T. 2017. Behavior is a major determinant of predation risk in zooplankton. *Ecosphere* 8(2), p. e01668.

Anisimova, M. and Gascuel, O. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology* 55(4), pp. 539–552.

Anon 1900a. 28. Bohumil Němec: Ueber die Art der Wahrnehmung des Schwerkraftreizes bei den Pflanzen. *Berichte der Deutschen Botanischen Gesellschaft*.

Anon 1900b. 31. G. Haberlandt: Ueber die Perception des geotropischen Reizes. *Berichte der Deutschen Botanischen Gesellschaft*.

Arendt, D. 2005. Genes and homology in nervous system evolution: comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory in Biosciences = Theorie in Den Biowissenschaften* 124(2), pp. 185–197.

Arendt, D., Benito-Gutierrez, E., Brunet, T. and Marlow, H. 2015. Gastric pouches and the mucociliary sole: setting the stage for nervous system evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370(1684).

Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M.D. and Wagner, G.P. 2016. The origin and evolution of cell types. *Nature Reviews. Genetics* 17(12), pp. 744–757.

Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A.W. and Wittbrodt, J. 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science* 306(5697), pp. 869–871.

Arkett, S.A., Mackie, G.O. and Meech, R.W. 1988. Hair Cell Mechanoreception in the Jellyfish Aglantha Digitale. *Journal of Experimental Biology*.

Arkett, S.A., Mackie, G.O. and Singla, C.L. 1987. Neuronal control of ciliary locomotion in a gastropod veliger (calliostoma). *The Biological Bulletin* 173(3), pp. 513–526.

Asadulina, A., Panzera, A., Verasztó, C., Liebig, C. and Jékely, G. 2012. Whole-body gene expression pattern registration in Platynereis larvae. *EvoDevo* 3(1), p. 27.

Ascoli, G.A., Donohue, D.E. and Halavi, M. 2007. NeuroMorpho.Org: a central resource for neuronal morphologies. *The Journal of Neuroscience* 27(35), pp. 9247–9251.

Ayers, T., Tsukamoto, H., Gühmann, M., Veedin Rajan, V.B. and Tessmar-Raible, K. 2018. A Go-type opsin mediates the shadow reflex in the annelid Platynereis dumerilii. *BMC Biology* 16(1), p. 41.

Backfisch, B., Veedin Rajan, V.B., Fischer, R.M., Lohs, C., Arboleda, E., Tessmar-Raible, K. and Raible, F. 2013. Stable transgenesis in the marine annelid Platynereis dumerilii sheds new light on photoreceptor evolution. *Proceedings of the National Academy of Sciences of the United States of America* 110(1), pp. 193–198.

Bacon, J.P. and Strausfeld, N.J. 1986. The dipteran ?Giant fibre? pathway: neurons and signals. *Journal of Comparative Physiology ? A* 158(4), pp. 529–548.

Baldwin, K.L., Strohm, A.K. and Masson, P.H. 2013. Gravity sensing and signal transduction in vascular plant primary roots. *American Journal of Botany* 100(1), pp. 126–142.

Bang, H., Kim, Y. and Kim, D. 2000. TREK-2, a new member of the mechanosensitive tandem-pore K+ channel family. *The Journal of Biological Chemistry* 275(23), pp. 17412–17419.

Bannister, S., Antonova, O., Polo, A., Lohs, C., Hallay, N., Valinciute, A., Raible, F. and Tessmar-Raible, K. 2014. TALENs mediate efficient and heritable mutation of endogenous genes in the marine annelid Platynereis dumerilii. *Genetics* 197(1), pp. 77–89.

Barr, M.M., DeModena, J., Braun, D., Nguyen, C.Q., Hall, D.H. and Sternberg, P.W. 2001. The Caenorhabditis elegans autosomal dominant polycystic kidney disease gene homologs lov-1 and pkd-2 act in the same pathway. *Current Biology* 11(17), pp. 1341–1346.

Barr, M.M. and Sternberg, P.W. 1999. A polycystic kidney-disease gene homologue required for male mating behavior in C. elegans. *Nature* 401(6751), pp. 386–389.

Bauknecht, P. and Jékely, G. 2017. Ancient coexistence of norepinephrine, tyramine, and octopamine signaling in bilaterians. *BMC Biology* 15(1), p. 6.

Bedini, C., Ferrero, E. and Lanfranchi, A. 1973. The ultrastructure of ciliary sensory cells in two turbellaria acoela. *Tissue and Cell* 5(3), pp. 359–372.

Berg, H.C. 1993. *Random walks in biology*. Expanded ed. Princeton, N.J: Princeton University Press.

Besschetnova, T.Y., Kolpakova-Hart, E., Guan, Y., Zhou, J., Olsen, B.R. and Shah, J.V. 2010. Identification of signaling pathways regulating primary cilium length and flow-mediated adaptation. *Current Biology* 20(2), pp. 182–187.

Beurg, M., Fettiplace, R., Nam, J.-H. and Ricci, A.J. 2009. Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging. *Nature Neuroscience* 12(5), pp. 553–558.

Bezares-Calderón, L.A., Berger, J., Jasek, S., Verasztó, C., Mendes, S., Gühmann, M., Almeda, R., Shahidi, R. and Jékely, G. 2018. Neural circuitry of a polycystin-mediated hydrodynamic startle response for predator avoidance. *eLife* 7.

Bhattacharyya, K., McLean, D.L. and MacIver, M.A. 2017. Visual threat assessment and reticulospinal encoding of calibrated responses in larval zebrafish. *Current Biology* 27(18), p. 2751–2762.e6.

Bickell-Page, L.R. 1991. Repugnatorial glands with associated striated muscle and sensory cells inMelibe leonina (Mollusca, Nudibranchia). *Zoomorphology* 110(5), pp. 281–291.

Bierman, H.S., Schriefer, J.E., Zottoli, S.J. and Hale, M.E. 2004. The effects of head and tail stimulation on the withdrawal startle response of the rope fish (Erpetoichthys calabaricus). *The Journal of Experimental Biology* 207(Pt 22), pp. 3985–3997.

Blaxter, J.H.S. 1978. Baroreception. In: Ali, M. A. ed. *Sensory Ecology*. Boston, MA: Springer US, pp. 375–409.

Bleckmann, H. 1994. *Reception of Hydrodynamic Stimuli in Aquatic and Semiaquatic Animals*. illustrated. G. Fischer Verlag.

Bleckmann, H., Budelmann, B.U. and Bullock, T.H. 1991. Peripheral and central nervous responses evoked by small water movements in a cephalopod. *Journal of comparative physiology. A, Sensory, neural, and behavioral physiology* 168(2), pp. 247–257.

Böhm, U.L., Prendergast, A., Djenoune, L., Nunes Figueiredo, S., Gomez, J., Stokes, C., Kaiser, S., Suster, M., Kawakami, K., Charpentier, M., Concordet, J.-P., Rio, J.-P., Del Bene, F. and Wyart, C. 2016. CSF-contacting neurons regulate locomotion by relaying mechanical stimuli to spinal circuits. *Nature Communications* 7, p. 10866.

Bollens, S.M. and Frost, B.W. 1989. Predator-induced diet vertical migration in a planktonic copepod. *Journal of plankton research* 11(5), pp. 1047–1065.

Bone, Q. and Ryan, K.P. 1978. Cupular sense organs in Ciona (Tunicata: Ascidiacea). *Journal of zoology* 186(3), pp. 417–429.

Booth, I.R., Edwards, M.D., Black, S., Schumann, U. and Miller, S. 2007. Mechanosensitive channels in bacteria: signs of closure? *Nature Reviews. Microbiology* 5(6), pp. 431–440.

Boulter, C., Mulroy, S., Webb, S., Fleming, S., Brindle, K. and Sandford, R. 2001. Cardiovascular, skeletal, and renal defects in mice with a targeted disruption of the Pkd1 gene. *Proceedings of the National Academy of Sciences of the United States of America* 98(21), pp. 12174–12179.

Bradley, C.J., Strickler, J.R., Buskey, E.J. and Lenz, P.H. 2013. Swimming and escape behavior in two species of calanoid copepods from nauplius to adult. *Journal of Plankton Research* 35(1), pp. 49–65.

Brandl, H., Moon, H., Vila-Farré, M., Liu, S.-Y., Henry, I. and Rink, J.C. 2016. PlanMine--a mineable resource of planarian biology and biodiversity. *Nucleic Acids Research* 44(D1), pp. D764-73.

Brierley, A.S. 2014. Diel vertical migration. *Current Biology* 24(22), pp. R1074-6.

Brierley, S.M., Castro, J., Harrington, A.M., Hughes, P.A., Page, A.J., Rychkov, G.Y. and Blackshaw, L.A. 2011. TRPA1 contributes to specific mechanically activated currents and sensory neuron mechanical hypersensitivity. *The Journal of Physiology* 589(Pt 14), pp. 3575–3593.

Broglio, E., Johansson, M. and Jonsson, P.R. 2001. Trophic interaction between copepods and ciliates: effects of prey swimming behavior on predation risk. *Marine Ecology Progress Series* 220, pp. 179–186.

Brohawn, S.G., Su, Z. and MacKinnon, R. 2014. Mechanosensitivity is mediated directly by the lipid membrane in TRAAK and TREK1 K+ channels. *Proceedings of the National Academy of Sciences of the United States of America* 111(9), pp. 3614–3619.

Browman, H.I., Kruse, S. and O'Brien, W.J. 1989. Foraging behavior of the predaceous cladoceran,*Leptodora kindti* , and escape responses of their prey. *Journal of plankton research* 11(5), pp. 1075–1088.

Browman, H.I., Yen, J., Fields, D.M., St-Pierre, J.-F. and Skiftesvik, A.B. 2011. Fine-scale observations of the predatory behavior of the carnivorous copepod Paraeuchaeta norvegica and the escape responses of their ichthyoplankton prey, Atlantic cod (Gadus morhua). *Marine biology* 158(12), pp. 2653–2660.

Brunet, T., Fischer, A.H., Steinmetz, P.R., Lauri, A., Bertucci, P. and Arendt, D. 2016. The evolutionary origin of bilaterian smooth and striated myocytes. *eLife* 5.

Budelmann, B.-U. 1989. Hydrodynamic Receptor Systems in Invertebrates. In: Coombs, S., Görner, P., and Münz, H. eds. *The Mechanosensory Lateral Line*. New York, NY: Springer New York, pp. 607–631.

Budelmann, B.U. and Bleckmann, H. 1988. A lateral line analogue in cephalopods: water waves generate microphonic potentials in the epidermal head lines ofSepia andLolliguncula. *Journal of Comparative Physiology ? A* 164(1), pp. 1–5.

Bullock, T.H. 1984. Comparative Neuroethology of Startle, Rapid Escape, and Giant Fiber-Mediated Responses. In: Eaton, R. C. ed. *Neural Mechanisms of Startle Behavior*. Boston, MA: Springer US, pp. 1–13.

Bullock, Theodore Holmes 1945. FUNCTIONAL ORGANIZATION OF THE GIANT FIBER SYSTEM

of*lumbricus*. *Journal of neurophysiology* 8(1), pp. 55–71.

Bullock, T.H. 1997. Neuroethology of Zooplankton. In: Lenz, P., Hartline, D. K., Purcell, J., and Macmillan, D. eds. *Zooplankton: Sensory ecology and physiology*. Amsterdam, The Netherlands: CRC Press.

Bullock, T H 1945. ORGANIZATION OF THE GIANT NERVE FIBER SYSTEM IN NEANTHES VIRENS. *The Biological Bulletin* 89(2), pp. 185–186.

Bullock, T.H. 1948. Physiological mapping of giant nerve fiber systems in polychaete annelids. *Physiologia comparata et oecologia; an international journal of comparative physiology and ecology* 1(1).

Burdick, D.S., Hartline, D.K. and Lenz, P.H. 2007. Escape strategies in co-occurring calanoid copepods. *Limnology and Oceanography Letters* 52(6), pp. 2373–2385.

Burgos, A., Honjo, K., Ohyama, T., Qian, C.S., Shin, G.J.-E., Gohl, D.M., Silies, M., Tracey, W.D., Zlatic, M., Cardona, A. and Grueber, W.B. 2018. Nociceptive interneurons control modular motor pathways to promote escape behavior in Drosophila. *eLife* 7.

Burighel, P., Lane, N.J., Fabio, G., Stefano, T., Zaniolo, G., Carnevali, M.D.C. and Manni, L. 2003. Novel, secondary sensory cell organ in ascidians: in search of the ancestor of the vertebrate lateral line. *The Journal of Comparative Neurology* 461(2), pp. 236–249.

Buskey, E.J. 1984. Swimming pattern as an indicator of the roles of copepod sensory systems in the recognition of food. *Marine biology* 79(2), pp. 165–175.

Buskey, E.J. and Hartline, D.K. 2003. High-speed video analysis of the escape responses of the copepod Acartia tonsa to shadows. *The Biological Bulletin* 204(1), pp. 28–37.

Buskey, E.J., Lenz, P.H. and Hartline, D.K. 2002. Escape behavior of planktonic copepods in response to hydrodynamic disturbances: high speed video analysis. *Marine Ecology Progress Series* 235, pp. 135–146.

Buskey, E.J., Strickler, J.R., Bradley, C.J., Hartline, D.K. and Lenz, P.H. 2017. Escapes in copepods: comparison between myelinate and amyelinate species. *The Journal of Experimental Biology* 220(Pt 5), pp. 754–758.

Byrne, J.H. and Hawkins, R.D. 2015. Nonassociative learning in invertebrates. *Cold Spring Harbor Perspectives in Biology* 7(5).

Cai, Y., Maeda, Y., Cedzich, A., Torres, V.E., Wu, G., Hayashi, T., Mochizuki, T., Park, J.H., Witzgall, R. and Somlo, S. 1999. Identification and characterization of polycystin-2, the PKD2 gene product. *The Journal of Biological Chemistry* 274(40), pp. 28557–28565.

Calbet, A., Carlotti, F. and Gaudy, R. 2007. The feeding ecology of the copepod Centropages typicus (Kröyer). *Progress In Oceanography* 72(2–3), pp. 137–150.

Cantell, C.-E., Franz[illegible]n, [illegible]ke and Sensenbaugh, T. 1982. Ultrastructure of multiciliated collar cells in the pilidium larva of Lineus bilineatus (Nemertini). *Zoomorphology* 101(1), pp. 1–15.

Card, G. and Dickinson, M.H. 2008. Visually mediated motor planning in the escape response of Drosophila. *Current Biology* 18(17), pp. 1300–1307.

Celić, A., Petri, E.T., Demeler, B., Ehrlich, B.E. and Boggon, T.J. 2008. Domain mapping of the polycystin-2 C-terminal tail using de novo molecular modeling and biophysical analysis. *The Journal of Biological Chemistry* 283(42), pp. 28305–28312.

Chalfie, M. 2009. Neurosensory mechanotransduction. *Nature Reviews. Molecular Cell Biology* 10(1), pp. 44–52.

Chalfie, M. and Au, M. 1989. Genetic control of differentiation of the Caenorhabditis elegans touch receptor neurons. *Science* 243(4894 Pt 1), pp. 1027–1033.

Chalfie, M., Hart, A.C., Rankin, C.H. and Goodman, M.B. 2014. Assaying mechanosensation. *Wormbook: the Online Review of C. Elegans Biology*.

Chalfie, M., Sulston, J.E., White, J.G., Southgate, E., Thomson, J.N. and Brenner, S. 1985. The neural circuit for touch sensitivity in Caenorhabditis elegans. *The Journal of Neuroscience* 5(4), pp. 956–964.

Chartier, T. 2017. Chemosensation in the marine annelid Platynereis dumerilii : anatomy, physiology , behavior. Doctoral dissertation. Heidelberg University.

Chartier, T.F., Deschamps, J., Dürichen, W., Jékely, G. and Arendt, D. 2018. Whole-head recording of chemosensory activity in the marine annelid Platynereis dumerilii. *Open biology* 8(10).

Chatzigeorgiou, M., Bang, S., Hwang, S.W. and Schafer, W.R. 2013. tmc-1 encodes a sodium-sensitive channel required for salt chemosensation in C. elegans. *Nature* 494(7435), pp. 95–99.

Chatzigeorgiou, M. and Schafer, W.R. 2011. Lateral facilitation between primary mechanosensory neurons controls nose touch perception in C. elegans. *Neuron* 70(2), pp. 299–309.

Chatzigeorgiou, M., Yoo, S., Watson, J.D., Lee, W.-H., Spencer, W.C., Kindt, K.S., Hwang, S.W., Miller, D.M., Treinin, M., Driscoll, M. and Schafer, W.R. 2010. Specific roles for DEG/ENaC and TRP channels in touch and thermosensation in C. elegans nociceptors. *Nature Neuroscience* 13(7), pp. 861–868.

Chen, C.S. 2008. Mechanotransduction - a field pulling together? *Journal of Cell Science* 121(Pt 20), pp. 3285–3292.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K. and Kim, D.S. 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499(7458), pp. 295–300.

Chen, X. and Chalfie, M. 2014. Modulation of C. elegans touch sensitivity is integrated at multiple levels. *The Journal of Neuroscience* 34(19), pp. 6522–6536.

Chen, Y., Bharill, S., Isacoff, E.Y. and Chalfie, M. 2015. Subunit composition of a DEG/ENaC mechanosensory channel of Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* 112(37), pp. 11690–11695.

Cheng, L.E., Song, W., Looger, L.L., Jan, L.Y. and Jan, Y.N. 2010. The role of the TRP channel NompC in Drosophila larval and adult locomotion. *Neuron* 67(3), pp. 373–380.

Chia, F.-S., Buckland-Nicks, J. and Young, C.M. 1984. Locomotion of marine invertebrate larvae: a review. *Canadian journal of zoology* 62(7), pp. 1205–1222.

Chia, F.-S. and Koss, R. 1979. Fine structural studies of the nervous system and the apical organ in the planula larva of the sea anemoneAnthopleura elegantissima. *Journal of Morphology* 160(3), pp. 275–297.

Chou, S.-W., Chen, Z., Zhu, S., Davis, R.W., Hu, J., Liu, L., Fernando, C.A., Kindig, K., Brown, W.C., Stepanyan, R. and McDermott, B.M. 2017. A molecular basis for water motion detection by the mechanosensory lateral line of zebrafish. *Nature Communications* 8(1), p. 2234.

Christensen, A.P. and Corey, D.P. 2007. TRP channels in mechanosensation: direct or indirect activation? *Nature Reviews. Neuroscience* 8(7), pp. 510–521.

Christodoulou, F., Raible, F., Tomer, R., Simakov, O., Trachana, K., Klaus, S., Snyman, H., Hannon, G.J., Bork, P. and Arendt, D. 2010. Ancient animal microRNAs and the evolution of tissue identity. *Nature* 463(7284), pp. 1084–1088.

Cleves, P.A., Strader, M.E., Bay, L.K., Pringle, J.R. and Matz, M.V. 2018. CRISPR/Cas9-mediated genome editing in a reef-building coral. *Proceedings of the National Academy of Sciences of the United States of America* 115(20), pp. 5235–5240.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. and Zhang, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121), pp. 819–823.

Conzelmann, M., Offenburger, S.-L., Asadulina, A., Keller, T., Münch, T.A. and Jékely, G. 2011. Neuropeptides regulate swimming depth of Platynereis larvae. *Proceedings of the National Academy of Sciences of the United States of America*

108(46), pp. E1174-83.

Conzelmann, M., Williams, E.A., Krug, K., Franz-Wachtel, M., Macek, B. and Jékely, G. 2013. The neuropeptide complement of the marine annelid Platynereis dumerilii. *BMC Genomics* 14, p. 906.

Conzelmann, M., Williams, E.A., Tunaru, S., Randel, N., Shahidi, R., Asadulina, A., Berger, J., Offermanns, S. and Jékely, G. 2013. Conserved MIP receptor-ligand pair regulates Platynereis larval settlement. *Proceedings of the National Academy of Sciences of the United States of America* 110(20), pp. 8224–8229.

Cooke, I.R.C. and Macmillan, D.L. 1985. Further Studies of Crayfish Escape Behavior: I. The Role of the Appendages and the Stereotyped Nature of Non-Giant Escape Swimming | Journal of Experimental Biology. *Journal of Experimental Biology*.

Corey, D.P. and Hudspeth, A.J. 1983. Kinetics of the receptor current in bullfrog saccular hair cells. *The Journal of Neuroscience* 3(5), pp. 962–976.

Coste, B., Mathur, J., Schmidt, M., Earley, T.J., Ranade, S., Petrus, M.J., Dubin, A.E. and Patapoutian, A. 2010. Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* 330(6000), pp. 55–60.

Coste, B., Xiao, B., Santos, J.S., Syeda, R., Grandl, J., Spencer, K.S., Kim, S.E., Schmidt, M., Mathur, J., Dubin, A.E., Montal, M. and Patapoutian, A. 2012. Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* 483(7388), pp. 176–181.

Costello, J.H., Loftus, R. and Waggett, R. 1999. Influence of prey detection on capture success for the ctenophore Mnemiopsis leidyi feeding upon adult Acartia tonsa and Oithona colcarva copepods. *Marine Ecology Progress Series* 191, pp. 207–216.

Crisp, M. 1981. Epithelial sensory structures of trochids. *Journal of the Marine Biological Association of the UK* 61(01), p. 95.

Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J.P., Tsegaye, G., Holt, G.T., Hu, A., Walpita, D., Patel, R., Macklin, J.J., Bargmann, C.I., Ahrens, M.B., Schreiter, E.R., Jayaraman, V., Looger, L.L., Svoboda, K. and Kim, D.S. 2016. Sensitive red protein calcium indicators for imaging neural activity. *eLife* 5.

von Dassow, G., Emlet, R.B. and Maslakova, S.A. 2013. How the pilidium larva feeds. *Frontiers in zoology* 10(1), p. 47.

DeCaen, P.G., Delling, M., Vien, T.N. and Clapham, D.E. 2013. Direct recording and molecular identification of the calcium channel of primary cilia. *Nature* 504(7479), pp. 315–318.

DeCaen, P.G., Liu, X., Abiria, S. and Clapham, D.E. 2016. Atypical calcium regulation of the PKD2-L1 polycystin ion channel. *eLife* 5.

Delling, M., DeCaen, P.G., Doerner, J.F., Febvay, S. and Clapham, D.E. 2013. Primary cilia are specialized calcium signalling organelles. *Nature* 504(7479), pp. 311–314.

Delling, M., Indzhykulian, A.A., Liu, X., Li, Y., Xie, T., Corey, D.P. and Clapham, D.E. 2016. Primary cilia are not calcium-responsive mechanosensors. *Nature* 531(7596), pp. 656–660.

Delmas, P. 2004. Polycystins: from mechanosensation to gene regulation. *Cell* 118(2), pp. 145–148.

Delmas, P. and Coste, B. 2013. Mechano-gated ion channels in sensory systems. *Cell* 155(2), pp. 278–284.

Denes, A.S., Jékely, G., Steinmetz, P.R.H., Raible, F., Snyman, H., Prud'homme, B., Ferrier, D.E.K., Balavoine, G. and Arendt, D. 2007. Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell* 129(2), pp. 277–288.

Denis, V. and Cyert, M.S. 2002. Internal Ca(2+) release in yeast is triggered by hypertonic shock and mediated by a TRP channel homologue. *The Journal of Cell Biology* 156(1), pp. 29–34.

Denton, E.J. and Gray, J. 1985. Lateral-Line-Like Antennae of Certain of the Penaeidea (Crustacea, Decapoda, Natantia). *Proceedings of the Royal Society B: Biological Sciences* 226(1244), pp. 249–261.

Digby, P.S.B. 1961. Mechanism of sensitivity to hydrostatic pressure in the prawn, palaemonetes varians leach. *Nature*

191(4786), pp. 366–368.

Dodson, S. 1988. The ecological role of chemical stimuli for the zooplankton: Predator-avoidance behavior in*Daphnia*. *Limnology and Oceanography Letters* 33(6part2), pp. 1431–1439.

Domenici, P. and Blake, R.W. 1991. The Kinematics and Performance of the Escape Response in the Angelfish (Pterophyllum Eimekei). *Journal of Experimental Biology*.

Donaldson, S., Mackie, G.O. and Roberts, A. 1980. Preliminary observations on escape swimming and giant neurons in*Aglantha digitale* (Hydromedusae: Trachylina). *Canadian journal of zoology* 58(4), pp. 549–552.

Doran, S.A., Koss, R., Tran, C.H., Christopher, K.J., Gallin, W.J. and Goldberg, J.I. 2004. Effect of serotonin on ciliary beating and intracellular calcium concentration in identified populations of embryonic ciliary cells. *The Journal of Experimental Biology* 207(Pt 8), pp. 1415–1429.

Dorresteijn, A.W.C. 1990. Quantitative analysis of cellular differentiation during early embryogenesis ofPlatynereis dumerilii. *Roux's Archives of Developmental Biology* 199(1), pp. 14–30.

Drewes, C.D. 1984. Escape reflexes in earthworms and other annelids. In: Eaton, R. C. ed. *Neural mechanisms of startle behavior*. Boston, MA: Springer US, pp. 43–91.

Dumont, J.P. and Robertson, R.M. 1986. Neuronal circuits: an evolutionary perspective. *Science* 233(4766), pp. 849–853.

Dur, G., Souissi, S., Schmitt, F.G., Beyrend-Dur, D. and Hwang, J.-S. 2011. Mating and mate choice in Pseudodiaptomus annandalei (Copepoda: Calanoida). *Journal of Experimental Marine Biology and Ecology* 402(1–2), pp. 1–11.

Duren, L.A. v., Stamhuis, E.J. and Videler, J.J. 1998. Reading the copepod personal ads: increasing encounter probability with hydromechanical signals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 353(1369), pp. 691–700.

Eakin, R.M. and Kuda, A. 1971. Ultrastructure of sensory receptors in Ascidian tadpoles. *Zeitschrift fur Zellforschung und mikroskopische Anatomie (Vienna, Austria : 1948)* 112(3), pp. 287–312.

Eaton, R.C., Bombardieri, R.A. and Meyer, D.L. 1977. The Mauthner-initiated startle response in teleost fish. *The Journal of Experimental Biology* 66(1), pp. 65–81.

Eaton, R.C., Lavender, W.A. and Wieland, C.M. 1982. Alternative neural pathways initiate fast-start responses following lesions of the mauthner neuron in goldfish. *Journal of Comparative Physiology ? A* 145(4), pp. 485–496.

Edwards, D. 2017. *Crayfish Escape*. Oxford University Press.

Edwards, D.H., Heitler, W.J. and Krasne, F.B. 1999. Fifty years of a command neuron: the neurobiology of escape behavior in the crayfish. *Trends in Neurosciences* 22(4), pp. 153–161.

Edwards, D.H., Yeh, S.R. and Krasne, F.B. 1998. Neuronal coincidence detection by voltage-sensitive electrical synapses. *Proceedings of the National Academy of Sciences of the United States of America* 95(12), pp. 7145–7150.

Effertz, T., Wiek, R. and Göpfert, M.C. 2011. NompC TRP channel is essential for Drosophila sound receptor function. *Current Biology* 21(7), pp. 592–597.

Ehlers, U. and Ehlers, B. 1977. Monociliary receptors in interstitial Proseriata and Neorhabdocoela (Turbellaria Neoophora). *Zoomorphologie* 86(3), pp. 197–222.

England, S.J., Campbell, P.C., Banerjee, S., Swanson, A.J. and Lewis, K.E. 2017. Identification and expression analysis of the complete family of zebrafish pkd genes. *Frontiers in cell and developmental biology* 5, p. 5.

Enright, J.T. 1965. Entrainment of a tidal rhythm. *Science* 147(3660), pp. 864–867.

Enright, J.T. 1961. Pressure sensitivity of an amphipod. *Science* 133(3455), pp. 758–760.

Ernstrom, G.G. and Chalfie, M. 2002. Genetics of sensory mechanotransduction. *Annual Review of Genetics* 36, pp. 411–

453.

Evans, G.T. 1989. The encounter speed of moving predator and prey. *Journal of plankton research* 11(2), pp. 415–417.

Fain, G.L., Hardie, R. and Laughlin, S.B. 2010. Phototransduction and the evolution of photoreceptors. *Current Biology* 20(3), pp. R114-24.

Faucherre, A., Nargeot, J., Mangoni, M.E. and Jopling, C. 2013. piezo2b regulates vertebrate light touch response. *The Journal of Neuroscience* 33(43), pp. 17089–17094.

Fetcho, J.R. 1991. Spinal network of the Mauthner cell. *Brain, Behavior and Evolution* 37(5), pp. 298–316.

Fetcho, J.R. and Faber, D.S. 1988. Identification of motoneurons and interneurons in the spinal network for escapes initiated by the mauthner cell in goldfish. *The Journal of Neuroscience* 8(11), pp. 4192–4213.

Field, S., Riley, K.-L., Grimes, D.T., Hilton, H., Simon, M., Powles-Glover, N., Siggers, P., Bogani, D., Greenfield, A. and Norris, D.P. 2011. Pkd1l1 establishes left-right asymmetry and physically interacts with Pkd2. *Development* 138(6), pp. 1131–1142.

Fields, D.M. and Yen, J. 1992. Escape responses of Acartia hudsonica (Copepoda) nauplii from the flow field of Temora longicornis (Copepoda). *Arch Hydrobiol Beih Ergebn Limnol* 36, pp. 123–134.

Fischer, A.H., Henrich, T. and Arendt, D. 2010. The normal development of Platynereis dumerilii (Nereididae, Annelida). *Frontiers in zoology* 7, p. 31.

Foggensteiner, L., Bevan, A.P., Thomas, R., Coleman, N., Boulter, C., Bradley, J., Ibraghimov-Beskrovnaya, O., Klinger, K. and Sandford, R. 2000. Cellular and subcellular distribution of polycystin-2, the protein product of the PKD2 gene. *Journal of the American Society of Nephrology* 11(5), pp. 814–827.

Friesen, W.O. 1981. Physiology of water motion detection in the medicinal leech. *The Journal of Experimental Biology* 92, pp. 255–275.

Fritzsch, B., Beisel, K.W., Pauley, S. and Soukup, G. 2007. Molecular evolution of the vertebrate mechanosensory cell and ear. *The International Journal of Developmental Biology* 51(6–7), pp. 663–678.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K. and Sander, J.D. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology* 31(9), pp. 822–826.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. and Joung, J.K. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology* 32(3), pp. 279–284.

Fuchs, H.L., Christman, A.J., Gerbi, G.P., Hunter, E.J. and Diez, F.J. 2015. Directional flow sensing by passively stable larvae. *The Journal of Experimental Biology* 218(Pt 17), pp. 2782–2792.

Fujiu, K., Nakayama, Y., Iida, H., Sokabe, M. and Yoshimura, K. 2011. Mechanoreception in motile flagella of Chlamydomonas. *Nature Cell Biology* 13(5), pp. 630–632.

Furshpan, E.J. and Furukawa, T. 1962. Intracellular and extracellular responses of the several regions of the Mauthner cell of the goldfish. *Journal of Neurophysiology* 25, pp. 732–771.

Gahtan, E., Sankrithi, N., Campos, J.B. and O'Malley, D.M. 2002. Evidence for a widespread brain stem escape network in larval zebrafish. *Journal of Neurophysiology* 87(1), pp. 608–614.

Gallager, S.M. 1993. Hydrodynamic disturbances produced by small zooplankton: case study for the veliger larva of a bivalve mollusc. *Journal of plankton research* 15(11), pp. 1277–1296.

Gallagher, A.R., Hoffmann, S., Brown, N., Cedzich, A., Meruvu, S., Podlich, D., Feng, Y., Könecke, V., de Vries, U., Hammes, H.-P., Gretz, N. and Witzgall, R. 2006. A truncated polycystin-2 protein causes polycystic kidney disease and retinal degeneration in transgenic rats. *Journal of the American Society of Nephrology* 17(10), pp. 2719–2730.

Gallio, M., Ofstad, T.A., Macpherson, L.J., Wang, J.W. and Zuker, C.S. 2011. The coding of temperature in the Drosophila brain. *Cell* 144(4), pp. 614–624.

Gaspar, I., Wippich, F. and Ephrussi, A. 2017. Enzymatic production of single-molecule FISH and RNA capture probes. *RNA (New York)* 23(10), pp. 1582–1591.

Gaylord, B., Hodin, J. and Ferner, M.C. 2013. Turbulent shear spurs settlement in larval sea urchins. *Proceedings of the National Academy of Sciences of the United States of America* 110(17), pp. 6901–6906.

Geffeney, S.L., Cueva, J.G., Glauser, D.A., Doll, J.C., Lee, T.H.-C., Montoya, M., Karania, S., Garakani, A.M., Pruitt, B.L. and Goodman, M.B. 2011. DEG/ENaC but not TRP channels are the major mechanoelectrical transduction channels in a C. elegans nociceptor. *Neuron* 71(5), pp. 845–857.

Gemmell, B.J., Jiang, H. and Buskey, E.J. 2015. A tale of the ciliate tail: investigation into the adaptive significance of this sub-cellular structure. *Proceedings. Biological Sciences / the Royal Society* 282(1812), p. 20150770.

Genin, A., Jaffe, J.S., Reef, R., Richter, C. and Franks, P.J.S. 2005. Swimming against the flow: a mechanism of zooplankton aggregation. *Science* 308(5723), pp. 860–862.

Giamarchi, A., Feng, S., Rodat-Despoix, L., Xu, Y., Bubenshchikova, E., Newby, L.J., Hao, J., Gaudioso, C., Crest, M., Lupas, A.N., Honoré, E., Williamson, M.P., Obara, T., Ong, A.C.M. and Delmas, P. 2010. A polycystin-2 (TRPP2) dimerization domain essential for the function of heteromeric polycystin complexes. *The EMBO Journal* 29(7), pp. 1176–1191.

Gilbert, J.J. 1985. Escape response of the rotifer Polyarthra: a high-speed cinematographic analysis. *Oecologia* 66(3), pp. 322–331.

Gilbert, J.J. 1987. The Polyarthra escape from response: Defense against interference from Daphnia. *Hydrobiologia* 147(1), pp. 235–238.

Gilbert, J.J. and Williamson, C.E. 1978. Predator-prey behavior and its effect on rotifer survival in associations of Mesocyclops edax, Asplanchna girodi, Polyarthra vulgaris, and Keratella cochlearis. *Oecologia* 37(1), pp. 13–22.

Gill, C.W. 1985. The response of a restrained copepod to tactile stimulation. *Marine Ecology Progress Series* 21, pp. 121–125.

Glantz, R.M. and Viancour, T. 1983. Integrative properties of crayfish medial giant neuron: steady-state model. *Journal of Neurophysiology* 50(5), pp. 1122–1142.

Goodman, M.B. 2006. Mechanosensation. *Wormbook: the Online Review of C. Elegans Biology*, pp. 1–14.

Goodman, M.B., Ernstrom, G.G., Chelur, D.S., O'Hagan, R., Yao, C.A. and Chalfie, M. 2002. MEC-2 regulates C. elegans DEG/ENaC channels needed for mechanosensation. *Nature* 415(6875), pp. 1039–1042.

Göpfert, M.C., Albert, J.T., Nadrowski, B. and Kamikouchi, A. 2006. Specification of auditory sensitivity by Drosophila TRP channels. *Nature Neuroscience* 9(8), pp. 999–1000.

Göpfert, M.C. and Robert, D. 2001. Active auditory mechanics in mosquitoes. *Proceedings. Biological Sciences / the Royal Society* 268(1465), pp. 333–339.

Göpfert, M.C. and Robert, D. 2003. Motion generation by Drosophila mechanosensory neurons. *Proceedings of the National Academy of Sciences of the United States of America* 100(9), pp. 5514–5519.

Grieben, M., Pike, A.C.W., Shintre, C.A., Venturi, E., El-Ajouz, S., Tessitore, A., Shrestha, L., Mukhopadhyay, S., Mahajan, P., Chalk, R., Burgess-Brown, N.A., Sitsapesan, R., Huiskonen, J.T. and Carpenter, E.P. 2017. Structure of the polycystic kidney disease TRP channel Polycystin-2 (PC2). *Nature Structural & Molecular Biology* 24(2), pp. 114–122.

Grienberger, C. and Konnerth, A. 2012. Imaging calcium in neurons. *Neuron* 73(5), pp. 862–885.

Grimes, D.T., Keynton, J.L., Buenavista, M.T., Jin, X., Patel, S.H., Kyosuke, S., Vibert, J., Williams, D.J., Hamada, H., Hussain, R., Nauli, S.M. and Norris, D.P. 2016. Genetic Analysis Reveals a Hierarchy of Interactions between Polycystin-Encoding Genes and Genes Controlling Cilia Function during Left-Right Determination. *PLoS Genetics* 12(6), p. e1006070.

Grosse, P. and Brown, P. 2003. Acoustic startle evokes bilaterally synchronous oscillatory EMG activity in the healthy human. *Journal of Neurophysiology* 90(3), pp. 1654–1661.

Grossman, Y., Alkon, D.L. and Heldman, E. 1979. A common origin of voltage noise and generator potentials in statocyst hair cells. *The Journal of General Physiology* 73(1), pp. 23–48.

Gühmann, M., Jia, H., Randel, N., Verasztó, C., Bezares-Calderón, L.A., Michiels, N.K., Yokoyama, S. and Jékely, G. 2015. Spectral Tuning of Phototaxis by a Go-Opsin in the Rhabdomeric Eyes of Platynereis. *Current Biology* 25(17), pp. 2265–2271.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59(3), pp. 307–321.

Guo, Y., Wang, Y., Zhang, W., Meltzer, S., Zanini, D., Yu, Y., Li, J., Cheng, T., Guo, Z., Wang, Q., Jacobs, J.S., Sharma, Y., Eberl, D.F., Göpfert, M.C., Jan, L.Y., Jan, Y.N. and Wang, Z. 2016. Transmembrane channel-like (tmc) gene regulates Drosophila larval locomotion. *Proceedings of the National Academy of Sciences of the United States of America* 113(26), pp. 7243–7248.

Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly, J.-S. and Concordet, J.-P. 2016. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology* 17(1), p. 148.

Häfker, N.S., Meyer, B., Last, K.S., Pond, D.W., Hüppe, L. and Teschke, M. 2017. Circadian clock involvement in zooplankton diel vertical migration. *Current Biology* 27(14), p. 2194–2201.e3.

Hale, M.E. 2014. Mapping circuits beyond the models: integrating connectomics and comparative neuroscience. *Neuron* 83(6), pp. 1256–1258.

Hale, M.E. 2002. S- and C-start escape responses of the muskellunge (Esox masquinongy) require alternative neuromotor mechanisms. *The Journal of Experimental Biology* 205(Pt 14), pp. 2005–2016.

Hale, M.E., Katz, H.R., Peek, M.Y. and Fremont, R.T. 2016. Neural circuits that drive startle behavior, with a focus on the Mauthner cells and spiral fiber neurons of fishes. *Journal of Neurogenetics* 30(2), pp. 89–100.

Hale, M.E., Long, J.H., McHenry, M.J. and Westneat, M.W. 2002. Evolution of behavior and neural control of the fast-start escape response. *Evolution* 56(5), pp. 993–1007.

Hansen, B.W., Jakobsen, H.H., Andersen, A., Almeda, R., Pedersen, T.M., Christensen, A.M. and Nilsson, B. 2010. Swimming behavior and prey retention of the polychaete larvae Polydora ciliata (Johnston). *The Journal of Experimental Biology* 213(Pt 18), pp. 3237–3246.

Hardy, A.C. and Bainbridge, R. 1951. Effect of pressure on the behavior of decapod larvae (Crustacea). *Nature* 167(4244), pp. 354–355.

Harris, P.C. and Torres, V.E. 2009. Polycystic kidney disease. *Annual Review of Medicine* 60, pp. 321–337.

Hart, A.C., Sims, S. and Kaplan, J.M. 1995. Synaptic code for sensory modalities revealed by C. elegans GLR-1 glutamate receptor. *Nature* 378(6552), pp. 82–85.

Hauenschild, C. and Fischer, A. 1969. *Platynereis dumerilii : mikroskopische Anatomie, Fortpflanzung, Entwicklung*. Stuttgart : G. Fischer, 1969.

Hausen, H. 2005. Chaetae and chaetogenesis in polychaetes (Annelida). *Hydrobiologia* 535–536(1), pp. 37–52.

Hawkins, R.D. and Byrne, J.H. 2015. Associative learning in invertebrates. *Cold Spring Harbor Perspectives in Biology* 7(5).

He, L., Gulyanon, S., Skanata, M.M., Karagyozov, D., Heckscher, E., Krieg, M., Tschekpenakis, G., Gershow, M. and Tracey, D. 2018. Proprioceptive neurons of larval Drosophila melanogaster show direction selective activity requiring the mechanosensory channel TMC. *bioRxiv*.

Heckscher, E.S., Zarin, A.A., Faumont, S., Clark, M.Q., Manning, L., Fushiki, A., Schneider-Mizell, C.M., Fetter, R.D., Truman, J.W., Zwart, M.F., Landgraf, M., Cardona, A., Lockery, S.R. and Doe, C.Q. 2015. Even-Skipped(+) Interneurons Are Core Components of a Sensorimotor Circuit that Maintains Left-Right Symmetric Muscle Contraction Amplitude. *Neuron* 88(2), pp. 314–329.

Hein, A.M., Gil, M.A., Twomey, C.R., Couzin, I.D. and Levin, S.A. 2018. Conserved behavioral circuits govern high-speed decision-making in wild fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*.

Heitler, W.J. and Fraser, K. 1993. Thoracic connections between crayfish giant fibres and motor giant neurones reverse abdominal pattern. *The Journal of Experimental Biology* 181, pp. 329–333.

Hemmrich, G. and Bosch, T.C.G. 2008. Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. *Bioessays: News and Reviews in Molecular, Cellular and Developmental Biology* 30(10), pp. 1010–1018.

Hempelmann, F. 1911. Zur Naturgeschichte von Nereis dumerilii Aud. et Edw. *Zoologica*.

Herberholz, J., Sen, M.M. and Edwards, D.H. 2004. Escape behavior and escape circuit activation in juvenile crayfish during prey-predator interactions. *The Journal of Experimental Biology* 207(Pt 11), pp. 1855–1863.

Horridge, G.A. 1959. Analysis of the rapid responses of Nereis and Harmothoë (Annelida). *Proceedings of the Royal Society of London. Series B, Containing Papers of A Biological Character. Royal Society (Great Britain)* 150(939), pp. 245–262.

Hsu, P.D., Lander, E.S. and Zhang, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6), pp. 1262–1278.

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z. and Liu, D.R. 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556(7699), pp. 57–63.

Huang, K., Diener, D.R., Mitchell, A., Pazour, G.J., Witman, G.B. and Rosenbaum, J.L. 2007. Function and dynamics of PKD2 in Chlamydomonas reinhardtii flagella. *The Journal of Cell Biology* 179(3), pp. 501–514.

Huang, M., Gu, G., Ferguson, E.L. and Chalfie, M. 1995. A stomatin-like protein necessary for mechanosensation in C. elegans. *Nature* 378(6554), pp. 292–295.

Hudspeth, A. 1997. Mechanical amplification of stimuli by hair cells. *Current Opinion in Neurobiology* 7(4), pp. 480–486.

Hudspeth, A.J. 1989. How the ear's works work. *Nature* 341(6241), pp. 397–404.

Hudspeth, A.J. 2014. Integrating the active process of hair cells with cochlear function. *Nature Reviews. Neuroscience* 15(9), pp. 600–614.

Hudspeth, A.J., Choe, Y., Mehta, A.D. and Martin, P. 2000. Putting ion channels to work: mechanoelectrical transduction, adaptation, and amplification by hair cells. *Proceedings of the National Academy of Sciences of the United States of America* 97(22), pp. 11765–11772.

Hudspeth, A.J. and Jacobs, R. 1979. Stereocilia mediate transduction in vertebrate hair cells (auditory system/cilium/vestibular system). *Proceedings of the National Academy of Sciences of the United States of America* 76(3), pp. 1506–1509.

Hulse, R.E., Li, Z., Huang, R.K., Zhang, J. and Clapham, D.E. 2018. Cryo-EM structure of the polycystin 2-l1 ion channel. *eLife* 7.

Hwang, R.Y., Zhong, L., Xu, Y., Johnson, T., Zhang, F., Deisseroth, K. and Tracey, W.D. 2007. Nociceptive neurons protect Drosophila larvae from parasitoid wasps. *Current Biology* 17(24), pp. 2105–2116.

Ikmi, A., McKinney, S.A., Delventhal, K.M. and Gibson, M.C. 2014. TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. *Nature Communications* 5, p. 5486.

Ishikawa, H. and Marshall, W.F. 2011. Ciliogenesis: building the cell's antenna. *Nature Reviews. Molecular Cell Biology*

12(4), pp. 222–234.

Jakobsen, H.H. 2002. Escape of protists in predator-generated feeding currents. *Aquatic Microbial Ecology* 26, pp. 271–281.

Jakobsen, H.H. 2001. Escape response of planktonic protists to fluid mechanical signals. *Marine Ecology Progress Series* 214, pp. 67–78.

Jarrell, T.A., Wang, Y., Bloniarz, A.E., Brittin, C.A., Xu, M., Thomson, J.N., Albertson, D.G., Hall, D.H. and Emmons, S.W. 2012. The connectome of a decision-making neural network. *Science* 337(6093), pp. 437–444.

Jékely, G. 2013. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proceedings of the National Academy of Sciences of the United States of America* 110(21), pp. 8702–8707.

Jékely, G., Colombelli, J., Hausen, H., Guy, K., Stelzer, E., Nédélec, F. and Arendt, D. 2008. Mechanism of phototaxis in marine zooplankton. *Nature* 456(7220), pp. 395–399.

Jennings, H.S. 1899. Studies on reactions to stimuli in unicellular organisms. ii.—the mechanism of the motor reactions of paramecium. *American Journal of Physiology-Legacy Content* 2(4), pp. 311–341.

Jennings, H.S. 1900. Studies on reactions to stimuli in unicellular organisms. V. On the movements and motor reflexes of the Flagellata and Ciliata. *The American Journal of Physiology* 3, pp. 229–260.

Jennings, H.S. 1904. The behavior of paramecium. Additional features and general relations. *Journal of Comparative Neurology and Psychology* 14(6), pp. 441–510.

Jiang, H. and Kiørboe, T. 2011. The fluid dynamics of swimming by jumping in copepods. *Journal of the Royal Society, Interface* 8(61), pp. 1090–1103.

Jin, Y., Yaguchi, S., Shiba, K., Yamada, L., Yaguchi, J., Shibata, D., Sawada, H. and Inaba, K. 2013. Glutathione transferase theta in apical ciliary tuft regulates mechanical reception and swimming behavior of Sea Urchin Embryos. *Cytoskeleton* 70(8), pp. 453–470.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096), pp. 816–821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. 2013. RNA-programmed genome editing in human cells. *eLife* 2, p. e00471.

Jouin, C., Tchernigovtzeff, C., Baucher, M.F. and Toulmond, A. 1985. Fine structure of probable mechano- and chemoreceptors in the caudal epidermis of the lugworm Arenicola marina (Annelida, Polychaeta). *Zoomorphology* 105(2), pp. 76–82.

Jovanic, T., Schneider-Mizell, C.M., Shao, M., Masson, J.-B., Denisov, G., Fetter, R.D., Mensh, B.D., Truman, J.W., Cardona, A. and Zlatic, M. 2016. Competitive disinhibition mediates behavioral choice and sequences in drosophila. *Cell* 167(3), p. 858–870.e19.

Kamura, K., Kobayashi, D., Uehara, Y., Koshida, S., Iijima, N., Kudo, A., Yokoyama, T. and Takeda, H. 2011. Pkd1l1 complexes with Pkd2 on motile cilia and functions to establish the left-right axis. *Development* 138(6), pp. 1121–1129.

Kang, L., Gao, J., Schafer, W.R., Xie, Z. and Xu, X.Z.S. 2010. C. elegans TRP family protein TRP-4 is a pore-forming subunit of a native mechanotransduction channel. *Neuron* 67(3), pp. 381–391.

Karak, S., Jacobs, J.S., Kittelmann, M., Spalthoff, C., Katana, R., Sivan-Loukianova, E., Schon, M.A., Kernan, M.J., Eberl, D.F. and Göpfert, M.C. 2015. Diverse roles of axonemal dyneins in drosophila auditory neuron function and mechanical amplification in hearing. *Scientific reports* 5, p. 17085.

Katta, S., Krieg, M. and Goodman, M.B. 2015. Feeling force: physical and physiological principles enabling sensory mechanotransduction. *Annual Review of Cell and Developmental Biology* 31, pp. 347–371.

Kawashima, Y., Géléoc, G.S.G., Kurima, K., Labay, V., Lelli, A., Asai, Y., Makishima, T., Wu, D.K., Della Santina,

C.C., Holt, J.R. and Griffith, A.J. 2011. Mechanotransduction in mouse inner ear hair cells requires transmembrane channel-like genes. *The Journal of Clinical Investigation* 121(12), pp. 4796–4809.

Keresztes, G., Mutai, H. and Heller, S. 2003. TMC and EVER genes belong to a larger novel family, the TMC gene family encoding transmembrane proteins. *BMC Genomics* 4(1), p. 24.

Kerfoot, W.C. 1978. Combat between predatory copepods and their prey: *Cyclops, Epischura*, and *Bosmina*. *Limnology and Oceanography Letters* 23(6), pp. 1089–1102.

Kernan, M., Cowan, D. and Zuker, C. 1994. Genetic dissection of mechanosensory transduction: mechanoreception-defective mutations of Drosophila. *Neuron* 12(6), pp. 1195–1206.

Kim, J.H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K.Y., Kim, M.-K., Shin, B.A. and Choi, S.-Y. 2011. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *Plos One* 6(4), p. e18556.

Kim, S.E., Coste, B., Chadha, A., Cook, B. and Patapoutian, A. 2012. The role of Drosophila Piezo in mechanical nociception. *Nature* 483(7388), pp. 209–212.

Kimura, Y., Okamura, Y. and Higashijima, S. 2006. alx, a zebrafish homolog of Chx10, marks ipsilateral descending excitatory interneurons that participate in the regulation of spinal locomotor circuits. *The Journal of Neuroscience* 26(21), pp. 5684–5697.

Kindt, K.S., Finch, G. and Nicolson, T. 2012. Kinocilia mediate mechanosensitivity in developing zebrafish hair cells. *Developmental Cell* 23(2), pp. 329–341.

Kindt, K.S., Quast, K.B., Giles, A.C., De, S., Hendrey, D., Nicastro, I., Rankin, C.H. and Schafer, W.R. 2007. Dopamine mediates context-dependent modulation of sensory plasticity in C. elegans. *Neuron* 55(4), pp. 662–676.

Kindt, K.S., Viswanath, V., Macpherson, L., Quast, K., Hu, H., Patapoutian, A. and Schafer, W.R. 2007. Caenorhabditis elegans TRPA-1 functions in mechanosensation. *Nature Neuroscience* 10(5), pp. 568–577.

Kiørboe, T. 2011. How zooplankton feed: mechanisms, traits and trade-offs. *Biological Reviews of the Cambridge Philosophical Society* 86(2), pp. 311–339.

Kiørboe, T., Andersen, A., Langlois, V.J., Jakobsen, H.H. and Bohr, T. 2009. Mechanisms and feasibility of prey capture in ambush-feeding zooplankton. *Proceedings of the National Academy of Sciences of the United States of America* 106(30), pp. 12394–12399.

Kiørboe, T., Jiang, H., Gonçalves, R.J., Nielsen, L.T. and Wadhwa, N. 2014. Flow disturbances generated by feeding and swimming zooplankton. *Proceedings of the National Academy of Sciences of the United States of America* 111(32), pp. 11738–11743.

Kiørboe, T. and Saiz, E. 1995. Planktivorous feeding in calm and turbulent environments, with emphasis on copepods. *Marine Ecology Progress Series* 122, pp. 135–145.

Kiørboe, T., Saiz, E. and Visser, A. 1999. Hydrodynamic signal perception in the copepod Acartia tonsa. *Marine Ecology Progress Series* 179, pp. 97–111.

Kiørboe, T. and Visser, A.W. 1999. Predator and prey perception in copepods due to hydromechanical signals. *Marine Ecology Progress Series* 179, pp. 81–95.

Kirk, K. and Gilbert, J. 1988. Escape Behavior of Polyarthra in Response to Artificial Flow Stimuli. *Bulletin of marine science* 43(3), pp. 551–560.

Kleene, N.K. and Kleene, S.J. 2012. A method for measuring electrical signals in a primary cilium. *Cilia* 1.

Kleene, S.J. and Kleene, N.K. 2017. The native TRPP2-dependent channel of murine renal primary cilia. *American Journal of Physiology. Renal Physiology* 312(1), pp. F96–F108.

Knapp, M.F. and Mill, P.J. 1971. The fine structure of ciliated sensory cells in the epidermis of the earthworm

Lumbricus terrestris. *Tissue & cell* 3(4), pp. 623–636.

Knight-Jones, E.W. and Qasim, S.Z. 1955. Responses of some marine plankton animals to changes in hydrostatic pressure. *Nature* 175(4465), pp. 941–942.

Kohashi, T. and Oda, Y. 2008. Initiation of Mauthner- or non-Mauthner-mediated fast escape evoked by different modes of sensory input. *The Journal of Neuroscience* 28(42), pp. 10641–10653.

König, P., Engel, A.K. and Singer, W. 1996. Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in Neurosciences* 19(4), pp. 130–137.

Koressaar, T. and Remm, M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23(10), pp. 1289–1291.

Korn, H. and Faber, D.S. 2005. The Mauthner cell half a century later: a neurobiological model for decision-making? *Neuron* 47(1), pp. 13–28.

Köttgen, M. and Walz, G. 2005. Subcellular localization and trafficking of polycystins. *Pflugers Archiv: European Journal of Physiology* 451(1), pp. 286–293.

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A. and Poeppel, D. 2017. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93(3), pp. 480–490.

Kramer, A.P. and Krasne, F.B. 1984. Crayfish escape behavior: production of tailflips without giant fiber activity. *Journal of Neurophysiology* 52(2), pp. 189–211.

Kramer, A.P., Krasne, F.B. and Wine, J.J. 1981. Interneurons between giant axons and motoneurons in crayfish escape circuitry. *Journal of Neurophysiology* 45(3), pp. 550–573.

Kristan, W. 2018. *Control of locomotion in annelids*. Byrne, J. H. ed. Oxford University Press.

Kristan, W.B., McGIRR, S.J. and Simpson, G.V. 1982. Behavioral and Mechanosensory Neurone Responses to Skin Stimulation in Leeches. *Journal of Experimental Biology*.

Kumagai, H., Nakanishi, T., Matsuura, T., Kato, Y. and Watanabe, H. 2017. CRISPR/Cas-mediated knock-in via non-homologous end-joining in the crustacean Daphnia magna. *Plos One* 12(10), p. e0186112.

Kung, C. 2005. A possible unifying principle for mechanosensation. *Nature* 436(7051), pp. 647–654.

Kupfermann, I. and Weiss, K.R. 1978. The command neuron concept. *Behavioral and Brain Sciences* 1(01), p. 3.

Kurima, K., Peters, L.M., Yang, Y., Riazuddin, Saima, Ahmed, Z.M., Naz, S., Arnaud, D., Drury, S., Mo, J., Makishima, T., Ghosh, M., Menon, P.S.N., Deshmukh, D., Oddoux, C., Ostrer, H., Khan, S., Riazuddin, Sheikh, Deininger, P.L., Hampton, L.L., Sullivan, S.L., Battey, J.F., Keats, B.J.B., Wilcox, E.R., Friedman, T.B. and Griffith, A.J. 2002. Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. *Nature Genetics* 30(3), pp. 277–284.

Kwan, K.Y., Glazer, J.M., Corey, D.P., Rice, F.L. and Stucky, C.L. 2009. TRPA1 modulates mechanotransduction in cutaneous sensory neurons. *The Journal of Neuroscience* 29(15), pp. 4808–4819.

Lacalli, T.C. 1986. Prototroch structure and innervation in the trochophore larva of*Phyllodoce* (Polychaeta). *Canadian journal of zoology* 64(1), pp. 176–184.

Lacalli, T.C. 1984. Structure and organization of the nervous system in the trochophore larva of spirobranchus. *Philosophical Transactions of the Royal Society B: Biological Sciences* 306(1126), pp. 79–135.

Lacalli, T.C. 1982. The nervous system and ciliary band of Müller's larva. *Proceedings of the Royal Society of London. Series B, Containing Papers of A Biological Character. Royal Society (Great Britain)* 217(1206), pp. 37–58.

Lacalli, T.C. and Gilmour, T.H.J. 1990. Ciliary Reversal and Locomotory Control in the Pluteus Larva of Lytechinus pictus. *Philosophical Transactions of the Royal Society B: Biological Sciences* 330(1258), pp. 391–396.

Lacalli, T.C., Gilmour, T.H.J. and West, J.E. 1990. Ciliary Band Innervation in the Bipinnaria Larva of Piaster

ochraceus. *Philosophical Transactions of the Royal Society B: Biological Sciences* 330(1258), pp. 371–390.

Lacalli, T.C. and Hou, S. 1999. A reexamination of the epithelial sensory cells of amphioxus (Branchiostoma). *Acta Zoologica* 80(2), pp. 125–134.

Lauga, E. and Powers, T.R. 2009. The hydrodynamics of swimming microorganisms. *Reports on Progress in Physics* 72(9), p. 096601.

Laumer, C.E., Gruber-Vodicka, H., Hadfield, M.G., Pearse, V.B., Riesgo, A., Marioni, J.C. and Giribet, G. 2018. Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. *eLife* 7.

Lauri, A., Brunet, T., Handberg-Thorsager, M., Fischer, A.H.L., Simakov, O., Steinmetz, P.R.H., Tomer, R., Keller, P.J. and Arendt, D. 2014. Development of the annelid axochord: insights into notochord evolution. *Science* 345(6202), pp. 1365–1368.

Laverack, M.S. 1962. Responses of cuticular sense organs of the lobster, Homarus vulgaris (crustacea)—II. Hair-fan organs as pressure receptors. *Comparative biochemistry and physiology* 6(2), pp. 137–145.

Lefort, V., Longueville, J.E. and Gascuel, O. 2017. SMS: smart model selection in phyml. *Molecular Biology and Evolution* 34(9), pp. 2422–2424.

Légier-Visser, M.F., Mitchell, J.G., Okubo, A. and Fuhrman, J.A. 1986. Mechanoreception in calanoid copepods. *Marine biology* 90(4), pp. 529–535.

Lehnert, B.P., Baker, A.E., Gaudry, Q., Chiang, A.-S. and Wilson, R.I. 2013. Distinct roles of TRP channels in auditory transduction and amplification in Drosophila. *Neuron* 77(1), pp. 115–128.

Lenz, P.H. and Hartline, D.K. 1999. Reaction times and force production during escape behavior of a calanoid copepod, Undinula vulgaris. *Marine biology* 133(2), pp. 249–258.

Lenz, P.H., Hower, A.E. and Hartline, D.K. 2004. Force production during pereiopod power strokes in Calanus finmarchicus. *Journal of Marine Systems* 49(1–4), pp. 133–144.

Lenz, P.H. and Yen, J. 1993. Distal setal mechanoreceptors of the first antennae of marine copepods. *Bulletin of marine science* 1(53), pp. 170–179.

Levina, N., Tötemeyer, S., Stokes, N.R., Louis, P., Jones, M.A. and Booth, I.R. 1999. Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *The EMBO Journal* 18(7), pp. 1730–1737.

Li, W., Kang, L., Piggott, B.J., Feng, Z. and Xu, X.Z.S. 2011. The neural circuits and sensory channels mediating harsh touch sensation in Caenorhabditis elegans. *Nature Communications* 2, p. 315.

Liang, X., Madrid, J., Saleh, H.S. and Howard, J. 2011. NOMPC, a member of the TRP channel family, localizes to the tubular body and distal cilium of Drosophila campaniform and chordotonal receptor cells. *Cytoskeleton* 68(1), pp. 1–7.

Liao, J.C. and Fetcho, J.R. 2008. Shared versus specialized glycinergic spinal interneurons in axial motor circuits of larval zebrafish. *The Journal of Neuroscience* 28(48), pp. 12982–12992.

Liden, W.H. and Herberholz, J. 2008. Behavioral and neural responses of juvenile crayfish to moving shadows. *The Journal of Experimental Biology* 211(Pt 9), pp. 1355–1361.

Lima, W.C., Vinet, A., Pieters, J. and Cosson, P. 2014. Role of PKD2 in rheotaxis in Dictyostelium. *Plos One* 9(2), p. e88682.

Lin, C.-Y. and Su, Y.-H. 2016. Genome editing in sea urchin embryos by using a CRISPR/Cas9 system. *Developmental Biology* 409(2), pp. 420–428.

Lin, M.Z. and Schnitzer, M.J. 2016. Genetically encoded indicators of neuronal activity. *Nature Neuroscience* 19(9), pp. 1142–1153.

Lin, S.-H., Cheng, Y.-R., Banks, R.W., Min, M.-Y., Bewick, G.S. and Chen, C.-C. 2016. Evidence for the involvement

of ASIC3 in sensory mechanotransduction in proprioceptors. *Nature Communications* 7, p. 11460.

Liu, K.S. and Fetcho, J.R. 1999. Laser ablations reveal functional relationships of segmental hindbrain neurons in zebrafish. *Neuron* 23(2), pp. 325–335.

Liu, X., Vien, T., Duan, J., Sheu, S.-H., DeCaen, P.G. and Clapham, D.E. 2018. Polycystin-2 is an essential ion channel subunit in the primary cilium of the renal collecting duct epithelium. *eLife* 7.

Liu, Y.-C., Bailey, I. and Hale, M.E. 2012. Alternative startle motor patterns and behaviors in the larval zebrafish (Danio rerio). *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 198(1), pp. 11–24.

Liu, Y.-C. and Hale, M.E. 2017. Local spinal cord circuits and bilateral mauthner cell activity function together to drive alternative startle behaviors. *Current Biology* 27(5), pp. 697–704.

Lu, W., Peissel, B., Babakhanlou, H., Pavlova, A., Geng, L., Fan, X., Larson, C., Brent, G. and Zhou, J. 1997. Perinatal lethality with kidney and pancreas defects in mice with a targetted Pkd1 mutation. *Nature Genetics* 17(2), pp. 179–181.

Lu, Y., Ma, X., Sabharwal, R., Snitsarev, V., Morgan, D., Rahmouni, K., Drummond, H.A., Whiteis, C.A., Costa, V., Price, M., Benson, C., Welsh, M.J., Chapleau, M.W. and Abboud, F.M. 2009. The ion channel ASIC2 is required for baroreceptor and autonomic control of the circulation. *Neuron* 64(6), pp. 885–897.

Ludeman, D.A., Farrar, N., Riesgo, A., Paps, J. and Leys, S.P. 2014. Evolutionary origins of sensation in metazoans: functional evidence for a new sensory organ in sponges. *BMC Evolutionary Biology* 14, p. 3.

Luo, L., Callaway, E.M. and Svoboda, K. 2008. Genetic dissection of neural circuits. *Neuron* 57(5), pp. 634–660.

Luo, L., Callaway, E.M. and Svoboda, K. 2018. Genetic dissection of neural circuits: A decade of progress. *Neuron* 98(2), pp. 256–281.

Luo, Y.-J., Kanda, M., Koyanagi, R., Hisata, K., Akiyama, T., Sakamoto, H., Sakamoto, T. and Satoh, N. 2018. Nemertean and phoronid genomes reveal lophotrochozoan evolution and the origin of bilaterian heads. *Nature ecology & evolution* 2(1), pp. 141–151.

Lynagh, T., Mikhaleva, Y., Colding, J.M., Glover, J.C. and Pless, S.A. 2018. Acid-sensing ion channels emerged over 600 Mya and are conserved throughout the deuterostomes. *Proceedings of the National Academy of Sciences of the United States of America* 115(33), pp. 8430–8435.

Lyons, K.M. 1973. Collar cells in planula and adult tentacle ectoderm of the solitary coral Balanophyllia regia (anthozoa eupsammiidae). *Zeitschrift f*[illegible]*r Zellforschung und Mikroskopische Anatomie* 45(1), pp. 57–74.

Machemer, H. and Ogura, A. 1979. Ionic conductances of membranes in ciliated and deciliated Paramecium. *The Journal of Physiology* 296, pp. 49–60.

Mackie, G. and Meech, R. 1995. Central circuitry in the jellyfish Aglantha. II: The ring giant and carrier systems. *The Journal of Experimental Biology* 198(Pt 11), pp. 2271–2278.

Mackie, G.O. 1970. Neuroid conduction and the evolution of conducting tissues. *The Quarterly Review of Biology* 45(4), pp. 319–332.

Mackie, G.O., Paul, D.H., Singla, C.M., Sleigh, M.A. and Williams, D.E. 1974. Branchial innervation and ciliary control in the ascidian Corella. *Proceedings of the Royal Society of London. Series B, Containing Papers of A Biological Character. Royal Society (Great Britain)* 187(1086), pp. 1–35.

Mackie, G.O., Singla, C.L. and Thiriot-Quievreux, C. 1976. Nervous control of ciliary activity in gastropod larvae. *The Biological Bulletin* 151(1), pp. 182–199.

Maguire, S.M., Clark, C.M., Nunnari, J., Pirri, J.K. and Alkema, M.J. 2011. The C. elegans touch response facilitates escape from predacious fungi. *Current Biology* 21(15), pp. 1326–1330.

Mahmoodian Sani, M.R., Hashemzadeh-Chaleshtori, M., Saidijam, M., Jami, M.-S. and Ghasemi-Dehkordi, P. 2016. MicroRNA-183 Family in Inner Ear: Hair Cell Development and Deafness. *Journal of audiology & otology* 20(3), pp. 131–

138.

Maingret, F., Fosset, M., Lesage, F., Lazdunski, M. and Honoré, E. 1999. TRAAK is a mammalian neuronal mechano-gated K+ channel. *The Journal of Biological Chemistry* 274(3), pp. 1381–1387.

Maksimovic, S., Nakatani, M., Baba, Y., Nelson, A.M., Marshall, K.L., Wellnitz, S.A., Firozi, P., Woo, S.-H., Ranade, S., Patapoutian, A. and Lumpkin, E.A. 2014. Epidermal Merkel cells are mechanosensory cells that tune mammalian touch receptors. *Nature* 509(7502), pp. 617–621.

Marder, E. 2007. Searching for Insight: Using Invertebrate Nervous Systems to Illuminate Fundamental Principles in Neuroscience. In: North, G. and Greenspan, R. J. eds. *Invertebrate Neurobiology*. 1st ed. pp. 1–18.

Markl, H. 1978. Adaptive radiation of mechanoreception. In: Ali, M. A. ed. *Sensory Ecology*. Boston, MA: Springer US, pp. 319–344.

Markl, H. 1973. [Sense of vibration in invertebrates]. *Fortschritte der Zoologie* 21(2), pp. 100–120.

Marlétaz, F., Peijnenburg, K.T.C.A., Goto, T., Satoh, N. and Rokhsar, D.S. 2019. A New Spiralian Phylogeny Places the Enigmatic Arrow Worms among Gnathiferans. *Current Biology*.

Marlow, H., Tosches, M.A., Tomer, R., Steinmetz, P.R., Lauri, A., Larsson, T. and Arendt, D. 2014. Larval body patterning and apical organs are conserved in animal evolution. *BMC Biology* 12, p. 7.

Marshall, K.L. and Lumpkin, E.A. 2012. The molecular basis of mechanosensory transduction. *Advances in Experimental Medicine and Biology* 739, pp. 142–155.

Martens, E.A., Wadhwa, N., Jacobsen, N.S., Lindemann, C., Andersen, K.H. and Visser, A. 2015. Size structures sensory hierarchy in ocean life. *Proceedings. Biological Sciences / the Royal Society* 282(1815).

Martin, A., Serano, J.M., Jarvis, E., Bruce, H.S., Wang, J., Ray, S., Barker, C.A., O'Connell, L.C. and Patel, N.H. 2016. Crispr/cas9 mutagenesis reveals versatile roles of hox genes in crustacean limb specification and evolution. *Current Biology* 26(1), pp. 14–26.

Martinac, B., Buechner, M., Delcour, A.H., Adler, J. and Kung, C. 1987. Pressure-sensitive ion channel in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* 84(8), pp. 2297–2301.

Masters, W.M., Aicher, B., Tautz, J. and Markl, H. 1982. A new type of water vibration receptor on the crayfish antenna. *Journal of Comparative Physiology ? A* 149(3), pp. 409–422.

Mauthner, S.E., Hwang, R.Y., Lewis, A.H., Xiao, Q., Tsubouchi, A., Wang, Y., Honjo, K., Skene, J.H.P., Grandl, J. and Tracey, W.D. 2014. Balboa binds to pickpocket in vivo and is required for mechanical nociception in Drosophila larvae. *Current Biology* 24(24), pp. 2920–2925.

McDougall, C., Chen, W.-C., Shimeld, S.M. and Ferrier, D.E.K. 2006. The development of the larval nervous system, musculature and ciliary bands of Pomatoceros lamarckii (Annelida): heterochrony in polychaetes. *Frontiers in zoology* 3, p. 16.

McLaughlin, S. 2017. Evidence that polycystins are involved in Hydra cnidocyte discharge. *Invertebrate Neuroscience* 17(1), p. 1.

Mettam, C. 1967. Segmental musculature and parapodial movement of Nereis diversicolor and Nephthys hombergi (Annelida: Polychaeta). *Journal of zoology* 153(2), pp. 245–275.

Michalec, F.-G., Fouxon, I., Souissi, S. and Holzner, M. 2017. Zooplankton can actively adjust their motility to turbulent flow. *Proceedings of the National Academy of Sciences of the United States of America* 114(52), pp. E11199–E11207.

Mire, P. and Watson, G.M. 1997. Mechanotransduction of hair bundles arising from multicellular complexes in anemones. *Hearing Research* 113(1–2), pp. 224–234.

Mire-Thibodeaux, P. and Watson, G.M. 1994. Morphodynamic hair bundles arising from sensory cell/supporting cell complexes frequency-tune nematocyst discharge in sea anemones. *The Journal of Experimental Zoology* 268(4), pp. 282–

292.

Momose, T. and Concordet, J.-P. 2016. Diving into marine genomics with CRISPR/Cas9 systems. *Marine genomics* 30, pp. 55–65.

Momose, T., De Cian, A., Shiba, K., Inaba, K., Giovannangeli, C. and Concordet, J.-P. 2018. High doses of CRISPR/Cas9 ribonucleoprotein efficiently induce gene knockout with low mosaicism in the hydrozoan Clytia hemisphaerica through microhomology-mediated deletion. *Scientific reports* 8(1), p. 11734.

Montalvo, S., Junoy, J., Roldán, C. and García-Corrales, P. 1996. Ultrastructural study of sensory cells of the proboscidial glandular epithelium ofRiseriellus occultus (Nemertea, Heteronemertea). *Journal of Morphology* 229(1), pp. 83–96.

Morgan, C.P., Zhao, H., LeMasurier, M., Xiong, W., Pan, B., Kazmierczak, P., Avenarius, M.R., Bateschell, M., Larisch, R., Ricci, A.J., Müller, U. and Barr-Gillespie, P.G. 2018. TRPV6, TRPM6 and TRPM7 Do Not Contribute to Hair-Cell Mechanotransduction. *Frontiers in Cellular Neuroscience* 12, p. 41.

Moritz, K. and Storch, V. 1971. Elektronenmikroskopische Untersuchung eines Mechanorezeptors von Evertebraten (Priapuliden, Oligochaeten). *Zeitschrift f*[illegible]*r Zellforschung und Mikroskopische Anatomie* 17(2), pp. 226–234.

Mu, L., Bacon, J.P., Ito, K. and Strausfeld, N.J. 2014. Responses of Drosophila giant descending neurons to visual and mechanical stimuli. *The Journal of Experimental Biology* 217(Pt 12), pp. 2121–2129.

Murchison, D., Chrachri, A. and Mulloney, B. 1993. A separate local pattern-generating circuit controls the movements of each swimmeret in crayfish. *Journal of Neurophysiology* 70(6), pp. 2620–2631.

Murthy, S.E., Dubin, A.E., Whitwam, T., Jojoa Cruz, S., Cahalan, S.M., Mosavi, S.A., Ward, A.B. and Patapoutian, A. 2018. OSCA/TMEM63 are an evolutionarily conserved family of mechanically activated ion channels. *eLife* 7.

Naitoh, Y. and Eckert, R. 1969. Ionic mechanisms controlling behavioral responses of paramecium to mechanical stimulation. *Science* 164(3882), pp. 963–965.

Naitoh, Y. and Eckert, R. 1973. Sensory Mechanisms in Paramecium : II. Ionic Basis of the Hyperpolarizing Mechanoreceptor Potential. *Journal of Experimental Biology*.

Nakai, J., Ohkura, M. and Imoto, K. 2001. A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. *Nature Biotechnology* 19(2), pp. 137–141.

Nakanishi, N. and Martindale, M.Q. 2018. CRISPR knockouts reveal an endogenous role for ancient neuropeptides in regulating developmental timing in a sea anemone. *eLife* 7.

Nakanishi, T., Kato, Y., Matsuura, T. and Watanabe, H. 2014. CRISPR/Cas-mediated targeted mutagenesis in Daphnia magna. *Plos One* 9(5), p. e98363.

Nauli, S.M., Alenghat, F.J., Luo, Y., Williams, E., Vassilev, P., Li, X., Elia, A.E.H., Lu, W., Brown, E.M., Quinn, S.J., Ingber, D.E. and Zhou, J. 2003. Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nature Genetics* 33(2), pp. 129–137.

Newbury, T.K. 1972. Vibration Perception by Chaetognaths. *Nature* 236(5348), pp. 459–460.

Newby, L.J., Streets, A.J., Zhao, Y., Harris, P.C., Ward, C.J. and Ong, A.C.M. 2002. Identification, characterization, and localization of a novel kidney polycystin-1-polycystin-2 complex. *The Journal of Biological Chemistry* 277(23), pp. 20763–20773.

Newland, P.L., Aonuma, H. and Nagayama, T. 1997. Monosynaptic excitation of lateral giant fibres by proprioceptive afferents in the crayfish. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* 181(2), pp. 103–109.

Nezlin, L.P. and Voronezhskaya, E.E. 2017. Early peripheral sensory neurons in the development of trochozoan animals. *Russian journal of developmental biology* 48(2), pp. 130–143.

Nezlin, L.P. and Yushin, V.V. 2004. Structure of the nervous system in the tornaria larva of Balanoglossus

proterogonius (Hemichordata: Enteropneusta) and its phylogenetic implications. *Zoomorphology* 123(1), pp. 1–13.

Nicol, J.A.C. 1948. The giant axons of annelids. *The Quarterly Review of Biology* 23(4), pp. 291–323.

Nicolson, T., Rüsch, A., Friedrich, R.W., Granato, M., Ruppersberg, J.P. and Nüsslein-Volhard, C. 1998. Genetic analysis of vertebrate sensory hair cell mechanosensation: the zebrafish circler mutants. *Neuron* 20(2), pp. 271–283.

Nissanov, J., Eaton, R.C. and DiDomenico, R. 1990. The motor output of the Mauthner cell, a reticulospinal command neuron. *Brain Research* 517(1–2), pp. 88–98.

Noël, J., Zimmermann, K., Busserolles, J., Deval, E., Alloui, A., Diochot, S., Guy, N., Borsotto, M., Reeh, P., Eschalier, A. and Lazdunski, M. 2009. The mechano-activated K+ channels TRAAK and TREK-1 control both warm and cold perception. *The EMBO Journal* 28(9), pp. 1308–1318.

Nørrevang, A. 1964. Choanocytes in the Skin of Harrimania kupfferi (Enteropneusta). *Nature* 204(4956), pp. 398–399.

Ogura, A. and Machemer, H. 1980. Distribution of mechanoreceptor channels in theParamecium surface membrane. *Journal of Comparative Physiology ? A* 135(3), pp. 233–242.

Ogura, A. and Takahashi, K. 1976. Artificial deciliation causes loss of calcium-dependent responses in Paramecium. *Nature* 264(5582), pp. 170–172.

O'Hagan, R., Chalfie, M. and Goodman, M.B. 2005. The MEC-4 DEG/ENaC channel of Caenorhabditis elegans touch receptor neurons transduces mechanical signals. *Nature Neuroscience* 8(1), pp. 43–50.

Ohman, M. 1988. Behavioral responses of zooplankton to predation. *Bulletin of Marine Science* 43, pp. 530–550.

Ohyama, T., Schneider-Mizell, C.M., Fetter, R.D., Aleman, J.V., Franconville, R., Rivera-Alba, M., Mensh, B.D., Branson, K.M., Simpson, J.H., Truman, J.W., Cardona, A. and Zlatic, M. 2015. A multilevel multimodal circuit enhances action selection in Drosophila. *Nature* 520(7549), pp. 633–639.

Okuda, S. 1946. Studies on the Development of Annelida Polychaeta I . *Journal of the Faculty of Science Hokkaido Zoology.* 6(9), pp. 115–219.

Olson, G.C. and Krasne, F.B. 1981. The crayfish lateral giants as command neurons for escape behavior. *Brain Research* 214(1), pp. 89–100.

O'Malley, D.M., Kao, Y.H. and Fetcho, J.R. 1996. Imaging the functional organization of zebrafish hindbrain segments during escape behaviors. *Neuron* 17(6), pp. 1145–1155.

Owen, G. and McCrae, J.M. 1979. Sensory Cell/Gland Cell Complexes Associated with the Pallial Tentacles of the Bivalve Lima hians (Gmelin), with a Note on Specialized Cilia on the Pallial Curtains. *Philosophical Transactions of the Royal Society B: Biological Sciences* 287(1017), pp. 45–62.

Özpolat, B.D., Handberg-Thorsager, M., Vervoort, M. and Balavoine, G. 2017. Cell lineage and cell cycling analyses of the 4d micromere using live imaging in the marine annelid Platynereis dumerilii. *eLife* 6.

Palmer, C.P., Zhou, X.L., Lin, J., Loukin, S.H., Kung, C. and Saimi, Y. 2001. A TRP homolog in Saccharomyces cerevisiae forms an intracellular Ca(2+)-permeable channel in the yeast vacuolar membrane. *Proceedings of the National Academy of Sciences of the United States of America* 98(14), pp. 7801–7805.

Pan, B., Akyuz, N., Liu, X.-P., Asai, Y., Nist-Lund, C., Kurima, K., Derfler, B.H., György, B., Limapichat, W., Walujkar, S., Wimalasena, L.N., Sotomayor, M., Corey, D.P. and Holt, J.R. 2018. TMC1 forms the pore of mechanosensory transduction channels in vertebrate inner ear hair cells. *Neuron* 99(4), p. 736–753.e6.

Pan, B., Géléoc, G.S., Asai, Y., Horwitz, G.C., Kurima, K., Ishikawa, K., Kawashima, Y., Griffith, A.J. and Holt, J.R. 2013. TMC1 and TMC2 are components of the mechanotransduction channel in hair cells of the mammalian inner ear. *Neuron* 79(3), pp. 504–515.

Patel, A.J., Honoré, E., Maingret, F., Lesage, F., Fink, M., Duprat, F. and Lazdunski, M. 1998. A mammalian two pore domain mechano-gated S-like K+ channel. *The EMBO Journal* 17(15), pp. 4283–4290.

Patella, P. and Wilson, R.I. 2018. Functional maps of mechanosensory features in the drosophila brain. *Current Biology* 28(8), p. 1189–1203.e5.

Pazour, G.J., San Agustin, J.T., Follit, J.A., Rosenbaum, J.L. and Witman, G.B. 2002. Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Current Biology* 12(11), pp. R378-80.

Peng, G., Shi, X. and Kadowaki, T. 2015. Evolution of TRP channels inferred by their classification in diverse animal species. *Molecular Phylogenetics and Evolution* 84, pp. 145–157.

Pennington, J.T. and Chia, F.-S. 1984. Morphological and behavioral defenses of trochophore larvae of *Sabellaria cementarium* (Polychaeta) against four planktonic predators. *The Biological Bulletin* 167(1), pp. 168–175.

Perry, K.J. and Henry, J.Q. 2015. CRISPR/Cas9-mediated genome modification in the mollusc, Crepidula fornicata. *Genesis* 53(2), pp. 237–244.

Peterka, D.S., Takahashi, H. and Yuste, R. 2011. Imaging voltage in neurons. *Neuron* 69(1), pp. 9–21.

Peyronnet, R., Sharif-Naeini, R., Folgering, J.H.A., Arhatte, M., Jodar, M., El Boustany, C., Gallian, C., Tauc, M., Duranton, C., Rubera, I., Lesage, F., Pei, Y., Peters, D.J.M., Somlo, S., Sachs, F., Patel, A., Honoré, E. and Duprat, F. 2012. Mechanoprotection by polycystins against apoptosis is mediated through the opening of stretch-activated K(2P) channels. *Cell reports* 1(3), pp. 241–250.

Phelan, P., Nakagawa, M., Wilkin, M.B., Moffat, K.G., O'Kane, C.J., Davies, J.A. and Bacon, J.P. 1996. Mutations in shaking-B prevent electrical synapse formation in the Drosophila giant fiber system. *The Journal of Neuroscience* 16(3), pp. 1101–1113.

Phillips, C.E. and Friesen, W.O. 1982. Ultrastructure of the water-movement-sensitive sensilla in the medicinal leech. *Journal of Neurobiology* 13(6), pp. 473–486.

Pichon, X., Lagha, M., Mueller, F. and Bertrand, E. 2018. A growing toolbox to image gene expression in single cells: sensitive approaches for demanding challenges. *Molecular Cell* 71(3), pp. 468–480.

Pickles, J.O., Comis, S.D. and Osborne, M.P. 1984. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hearing Research* 15(2), pp. 103–112.

Pierce, M.L., Weston, M.D., Fritzsch, B., Gabel, H.W., Ruvkun, G. and Soukup, G.A. 2008. MicroRNA-183 family conservation and ciliated neurosensory organ expression. *Evolution & Development* 10(1), pp. 106–113.

Pirschel, F. and Kretzberg, J. 2016. Multiplexed population coding of stimulus properties by leech mechanosensory cells. *The Journal of Neuroscience* 36(13), pp. 3636–3647.

Poulet, S.A. and Marsot, P. 1978. Chemosensory grazing by marine calanoid copepods (arthropoda: crustacea). *Science* 200(4348), pp. 1403–1405.

Praetorius, H.A. and Spring, K.R. 2001. Bending the MDCK cell primary cilium increases intracellular calcium. *The Journal of Membrane Biology* 184(1), pp. 71–79.

Pratt, J.W. 1959. Remarks on zeros and ties in the wilcoxon signed rank procedures. *Journal of the American Statistical Association* 54(287), p. 655.

Preibisch, S., Saalfeld, S., Schindelin, J. and Tomancak, P. 2010. Software for bead-based registration of selective plane illumination microscopy data. *Nature Methods* 7(6), pp. 418–419.

Price, M.P., McIlwrath, S.L., Xie, J., Cheng, C., Qiao, J., Tarr, D.E., Sluka, K.A., Brennan, T.J., Lewin, G.R. and Welsh, M.J. 2001. The DRASIC cation channel contributes to the detection of cutaneous touch and acid stimuli in mice. *Neuron* 32(6), pp. 1071–1083.

Purcell, E.M. 1976. Life at low Reynolds number. In: *AIP Conference Proceedings*. AIP, pp. 49–64.

Purschke, G. 2005. Sense organs in polychaetes (Annelida). *Hydrobiologia* 535–536(1), pp. 53–78.

Purschke, G., Hugenschütt, M., Ohlmeyer, L., Meyer, H. and Weihrauch, D. 2017. Structural analysis of the branchiae

and dorsal cirri in Eurythoe complanata (Annelida, Amphinomida). *Zoomorphology* 136(1), pp. 1–18.

Qian, F., Germino, F.J., Cai, Y., Zhang, X., Somlo, S. and Germino, G.G. 1997. PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nature Genetics* 16(2), pp. 179–183.

Quiroga Artigas, G., Lapébie, P., Leclère, L., Takeda, N., Deguchi, R., Jékely, G., Momose, T. and Houliston, E. 2018. A gonad-expressed opsin mediates light-induced spawning in the jellyfish Clytia. *eLife* 7.

Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., Ferrier, D., Benes, V., de Jong, P., Weissenbach, J., Bork, P. and Arendt, D. 2005. Vertebrate-type intron-rich genes in the marine annelid Platynereis dumerilii. *Science* 310(5752), pp. 1325–1326.

Ranade, S.S., Syeda, R. and Patapoutian, A. 2015. Mechanically activated ion channels. *Neuron* 87(6), pp. 1162–1179.

Ranade, S.S., Woo, S.-H., Dubin, A.E., Moshourab, R.A., Wetzel, C., Petrus, M., Mathur, J., Bégay, V., Coste, B., Mainquist, J., Wilson, A.J., Francisco, A.G., Reddy, K., Qiu, Z., Wood, J.N., Lewin, G.R. and Patapoutian, A. 2014. Piezo2 is the major transducer of mechanical forces for touch sensation in mice. *Nature* 516(7529), pp. 121–125.

Randel, N., Asadulina, A., Bezares-Calderón, L.A., Verasztó, C., Williams, E.A., Conzelmann, M., Shahidi, R. and Jékely, G. 2014. Neuronal connectome of a sensory-motor circuit for visual navigation. *eLife* 3.

Randel, N., Shahidi, R., Verasztó, C., Bezares-Calderón, L.A., Schmidt, S. and Jékely, G. 2015. Inter-individual stereotypy of the Platynereis larval visual connectome. *eLife* 4, p. e08069.

von Reyn, C.R., Breads, P., Peek, M.Y., Zheng, G.Z., Williamson, W.R., Yee, A.L., Leonardo, A. and Card, G.M. 2014. A spike-timing mechanism for action selection. *Nature Neuroscience* 17(7), pp. 962–970.

von Reyn, C.R., Nern, A., Williamson, W.R., Breads, P., Wu, M., Namiki, S. and Card, G.M. 2017. Feature integration drives probabilistic behavior in the drosophila escape response. *Neuron* 94(6), p. 1190–1204.e6.

Rice, A.L. 1964. Observations on the effects of changes of hydrostatic pressure on the behavior of some marine animals. *Journal of the Marine Biological Association of the UK* 44(01), p. 163.

Ringelberg, J. 2010. *Diel vertical migration of zooplankton in lakes and oceans*. Dordrecht: Springer Netherlands.

Roberts, A. and Mackie, G.O. 1980. The giant axon escape system of a hydrozoan medusa, Aglantha digitale. *The Journal of Experimental Biology* 84, pp. 303–318.

Robertson, J.L., Tsubouchi, A. and Tracey, W.D. 2013. Larval defense against attack from parasitoid wasps requires nociceptive neurons. *Plos One* 8(10), p. e78704.

Rodat-Despoix, L., Hao, J., Dandonneau, M. and Delmas, P. 2013. Shear stress-induced Ca^{2+} mobilization in MDCK cells is ATP dependent, no matter the primary cilium. *Cell Calcium* 53(5–6), pp. 327–337.

Rothschild, B.J. and Osborn, T.R. 1988. Small-scale turbulence and plankton contact rates. *Journal of plankton research* 10(3), pp. 465–474.

Russell, I.J. and Roberts, B.L. 1974. Active reduction of lateral-line sensitivity in swimming dogfish. *Journal of Comparative Physiology ? A* 94(1), pp. 7–15.

Ryan, K., Lu, Z. and Meinertzhagen, I.A. 2017. Circuit Homology between Decussating Pathways in the Ciona Larval CNS and the Vertebrate Startle-Response Pathway. *Current Biology* 27(5), pp. 721–728.

Ryan, K., Lu, Z. and Meinertzhagen, I.A. 2016. The CNS connectome of a tadpole larva of Ciona intestinalis (L.) highlights sidedness in the brain of a chordate sibling. *eLife* 5.

Saalfeld, S., Cardona, A., Hartenstein, V. and Tomancak, P. 2009. CATMAID: collaborative annotation toolkit for massive amounts of image data. *Bioinformatics* 25(15), pp. 1984–1986.

Saiz, E., Calbet, A. and Broglio, E. 2003. Effects of small-scale turbulence on copepods: The case of*Oithona davisae*. *Limnology and Oceanography Letters* 48(3), pp. 1304–1311.

Sánchez-Alcañiz, J.A., Zappia, G., Marion-Poll, F. and Benton, R. 2017. A mechanosensory receptor required for food

texture detection in Drosophila. *Nature Communications* 8, p. 14192.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K. and Dobbs, D. 2010. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Research* 38(Web Server issue), pp. W462-8.

Sanders, S.M., Ma, Z., Hughes, J.M., Riscoe, B.M., Gibson, G.A., Watson, A.M., Flici, H., Frank, U., Schnitzler, C.E., Baxevanis, A.D. and Nicotra, M.L. 2018. CRISPR/Cas9-mediated gene knockin in the hydroid Hydractinia symbiolongicarpus. *BMC Genomics* 19(1), p. 649.

Sarin, S., Bertrand, V., Bigelow, H., Boyanov, A., Doitsidou, M., Poole, R.J., Narula, S. and Hobert, O. 2010. Analysis of multiple ethyl methanesulfonate-mutagenized Caenorhabditis elegans strains by whole-genome sequencing. *Genetics* 185(2), pp. 417–430.

Sarkar, S. 1999. Focal adhesions. *Current Biology* 9(12), p. R428.

Sasaki, H., Yoshida, K., Hozumi, A. and Sasakura, Y. 2014. CRISPR/Cas9-mediated gene knockout in the ascidian Ciona intestinalis. *Development, Growth & Differentiation* 56(7), pp. 499–510.

Satir, P. 1963. Studies on cilia. the fixation of the metachronal wave. *The Journal of Cell Biology* 18, pp. 345–365.

Satou, C., Kimura, Y., Kohashi, T., Horikawa, K., Takeda, H., Oda, Y. and Higashijima, S. 2009. Functional role of a specialized class of spinal commissural inhibitory neurons during fast escapes in zebrafish. *The Journal of Neuroscience* 29(21), pp. 6780–6793.

Sawin, E.R., Ranganathan, R. and Horvitz, H.R. 2000. C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. *Neuron* 26(3), pp. 619–631.

Schafer, W.R. 2015. Mechanosensory molecules and circuits in C. elegans. *Pflugers Archiv: European Journal of Physiology* 467(1), pp. 39–48.

Schikorski, T. and Stevens, C.F. 1997. Quantitative ultrastructural analysis of hippocampal excitatory synapses. *The Journal of Neuroscience* 17(15), pp. 5858–5867.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods* 9(7), pp. 676–682.

Schlawny, A., Grünig, C. and Pfannenstiel, H.D. 1991. Sensory and secretory cells of *Ophryotrocha puerilis* (Polychaeta). *Zoomorphology* 110(4), pp. 209–215.

Schmidt, A., Bauknecht, P., Williams, E.A., Augustinowski, K., Gründer, S. and Jékely, G. 2018. Dual signaling of Wamide myoinhibitory peptides through a peptide-gated channel and a GPCR in Platynereis. *The FASEB Journal* 32(10), pp. 5338–5349.

Schmidtberg, H. and Dorresteijn, A.W.C. 2010. Ultrastructure of the nuchal organs in the polychaete Platynereis dumerilii (Annelida, Nereididae). *Invertebrate Biology* 129(3), pp. 252–265.

Schneider-Mizell, C.M., Gerhard, S., Longair, M., Kazimiers, T., Li, F., Zwart, M.F., Champion, A., Midgley, F.M., Fetter, R.D., Saalfeld, S. and Cardona, A. 2016. Quantitative neuroanatomy for connectomics in Drosophila. *eLife* 5.

Schüler, A., Schmitz, G., Reft, A., Özbek, S., Thurm, U. and Bornberg-Bauer, E. 2015. The Rise and Fall of TRP-N, an Ancient Family of Mechanogated Ion Channels, in Metazoa. *Genome Biology and Evolution* 7(6), pp. 1713–1727.

Sebe-Pedros, A., Chomsky, E., Saudemont, B., Mailhe, M.-P., Pleisser, F., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., Nouvault, S., Spitz, F., Tanay, A. and Marlow, H. 2017. Cnidarian cell type diversity revealed by whole-organism single-cell RNA-seq analysis. *BioRxiv*.

Segrove, F. 1941. Memoirs: The Development of the Serpiilid Pomatoeeros Triqueter L. *Journal of Cell Science*.

Seuront, L. 2013. Chemical and hydromechanical components of mate-seeking behavior in the calanoid copepod Eurytemora affinis. *Journal of plankton research* 35(4), pp. 724–743.

Shahidi, R., Williams, E.A., Conzelmann, M., Asadulina, A., Verasztó, C., Jasek, S., Bezares-Calderón, L.A. and Jékely, G. 2015. A serial multiplex immunogold labeling method for identifying peptidergic neurons in connectomes. *eLife* 4.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E. and Tsien, R.Y. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* 22(12), pp. 1567–1572.

Sharif-Naeini, R., Folgering, J.H.A., Bichet, D., Duprat, F., Lauritzen, I., Arhatte, M., Jodar, M., Dedman, A., Chatelain, F.C., Schulte, U., Retailleau, K., Loufrani, L., Patel, A., Sachs, F., Delmas, P., Peters, D.J.M. and Honoré, E. 2009. Polycystin-1 and -2 dosage regulates pressure sensing. *Cell* 139(3), pp. 587–596.

Shaw, B.K. and Kristan, W.B. 1995. The whole-body shortening reflex of the medicinal leech: motor pattern, sensory basis, and interneuronal pathways. *Journal of comparative physiology. A, Sensory, neural, and behavioral physiology* 177(6), pp. 667–681.

Shen, P.S., Yang, X., DeCaen, P.G., Liu, X., Bulkley, D., Clapham, D.E. and Cao, E. 2016. The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. *Cell* 167(3), p. 763–773.e11.

Short, G. and Tamm, S.L. 1991. On the nature of paddle cilia and discocilia. *The Biological Bulletin* 180(3), pp. 466–474.

Sidi, S., Friedrich, R.W. and Nicolson, T. 2003. NompC TRP channel required for vertebrate sensory hair cell mechanotransduction. *Science* 301(5629), pp. 96–99.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7, p. 539.

Sigger, J.N. and Dorsett, D.A. 1986a. Physiological and anatomical relationships of the 4B motoneurons of Nereis. *Comparative Biochemistry and Physiology Part A: Physiology* 84(4), pp. 783–788.

Sigger, J.N. and Dorsett, D.A. 1986b. The morphology and physiological properties of nereid giant fibres. *Comparative Biochemistry and Physiology Part A: Physiology* 83(3), pp. 589–595.

Silen, L. 1954. On the nervous system of Phoronis. *Arkiv för Zoologi* (2), p. 6.

Sillar, K.T. (Keith T., Picton, L. and Heitler, W.J. 2016. *The neuroethology of predation and escape*. Chichester, UK ; WILEY/Blackwell.

Simakov, O., Marletaz, F., Cho, S.-J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D.-H., Larsson, T., Lv, J., Arendt, D., Savage, R., Osoegawa, K., de Jong, P., Grimwood, J., Chapman, J.A., Shapiro, H., Aerts, A., Otillar, R.P., Terry, A.Y., Boore, J.L., Grigoriev, I.V., Lindberg, D.R., Seaver, E.C., Weisblat, D.A., Putnam, N.H. and Rokhsar, D.S. 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* 493(7433), pp. 526–531.

Simionato, E., Kerner, P., Dray, N., Le Gouar, M., Ledent, V., Arendt, D. and Vervoort, M. 2008. atonal- and achaete-scute-related genes in the annelid Platynereis dumerilii: insights into the evolution of neural basic-Helix-Loop-Helix genes. *BMC Evolutionary Biology* 8, p. 170.

Singarajah, K.V. 1969. Escape reactions of zooplankton: The avoidance of a pursuing siphon tube. *Journal of Experimental Marine Biology and Ecology* 3(2), pp. 171–178.

Singla, C.L. 1983. Fine structure of the sensory receptors of Aglantha digitale (Hydromedusae: Trachylina). *Cell and Tissue Research* 231(2), pp. 415–425.

Singla, V. and Reiter, J.F. 2006. The primary cilium as the cell's antenna: signaling at a sensory organelle. *Science* 313(5787), pp. 629–633.

Sinigaglia, C., Busengdal, H., Leclère, L., Technau, U. and Rentzsch, F. 2013. The bilaterian head patterning gene six3/6 controls aboral domain development in a cnidarian. *PLoS Biology* 11(2), p. e1001488.

Sinigaglia, C., Busengdal, H., Lerner, A., Oliveri, P. and Rentzsch, F. 2015. Molecular characterization of the apical

organ of the anthozoan Nematostella vectensis. *Developmental Biology* 398(1), pp. 120–133.

Smith, J.E. 1957. The Nervous Anatomy of the Body Segments of Nereid Polychaetes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 240(671), pp. 135–196.

Smith, J.M. and Szathmáry, E. 1998. *The Major Transitions in Evolution*. Oxford University Press.

Song, R., Walentek, P., Sponer, N., Klimke, A., Lee, J.S., Dixon, G., Harland, R., Wan, Y., Lishko, P., Lize, M., Kessel, M. and He, L. 2014. miR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. *Nature* 510(7503), pp. 115–120.

Spasic, M. and Jacobs, C.R. 2017. Primary cilia: Cell and molecular mechanosensors directing whole tissue function. *Seminars in Cell & Developmental Biology* 71, pp. 42–52.

Spierts, I.L. and Leeuwen, J.L. 1999. Kinematics and muscle dynamics of C- and S-starts of carp (Cyprinus carpio L.). *The Journal of Experimental Biology* 202(Pt 4), pp. 393–406.

Starunov, V.V., Voronezhskaya, E.E. and Nezlin, L.P. 2017. Development of the nervous system in Platynereis dumerilii (Nereididae, Annelida). *Frontiers in zoology* 14, p. 27.

Steigelman, K.A., Lelli, A., Wu, X., Gao, J., Lin, S., Piontek, K., Wodarczyk, C., Boletta, A., Kim, H., Qian, F., Germino, G., Géléoc, G.S.G., Holt, J.R. and Zuo, J. 2011. Polycystin-1 is required for stereocilia structure but not for mechanotransduction in inner ear hair cells. *The Journal of Neuroscience* 31(34), pp. 12241–12250.

Steinmetz, P.R.H., Kostyuchenko, R.P., Fischer, A. and Arendt, D. 2011. The segmental pattern of otx, gbx, and Hox genes in the annelid Platynereis dumerilii. *Evolution & Development* 13(1), pp. 72–79.

Sternberg, J.R., Prendergast, A.E., Brosse, L., Cantaut-Belarif, Y., Thouvenin, O., Orts-Del'Immagine, A., Castillo, L., Djenoune, L., Kurisu, S., McDearmid, J.R., Bardet, P.-L., Boccara, C., Okamoto, H., Delmas, P. and Wyart, C. 2018. Pkd2l1 is required for mechanoception in cerebrospinal fluid-contacting neurons and maintenance of spine curvature. *Nature Communications* 9(1), p. 3804.

Stewart, W.J., Nair, A., Jiang, H. and McHenry, M.J. 2014. Prey fish escape by sensing the bow wave of a predator. *The Journal of Experimental Biology* 217(Pt 24), pp. 4328–4336.

Stommel, E.W. 1986. Mechanical stimulation activates beating in calcium-arrested lateral cilia ofMytilus edulis gill. *Journal of muscle research and cell motility* 7(3), pp. 237–244.

Stommel, E.W., Stephens, R.E. and Alkon, D.L. 1980. Motile statocyst cilia transmit rather than directly transduce mechanical stimuli. *The Journal of Cell Biology* 87(3 Pt 1), pp. 652–662.

Strathmann, R.R. 1971. The feeding behavior of planktotrophic echinoderm larvae: Mechanisms, regulation, and rates of suspensionfeeding. *Journal of Experimental Marine Biology and Ecology* 6(2), pp. 109–160.

Strickler, J.R. 1975. Swimming of Planktonic Cyclops Species (Copepoda, Crustacea): Pattern, Movements and Their Control. In: Wu, T. Y.-T., Brokaw, C. J., and Brennen, C. eds. *Swimming and Flying in Nature*. Boston, MA: Springer US, pp. 599–613.

Strickler, J.R. and Bal, A.K. 1973. Setae of the First Antennae of the Copepod Cyclops scutifer (Sars): Their Structure and Importance. *Proceedings of the National Academy of Sciences of the United States of America* 70(9), pp. 2656–2659.

Strickler, J.R. and Balázsi, G. 2007. Planktonic copepods reacting selectively to hydrodynamic disturbances. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 362(1487), pp. 1947–1958.

Su, Q., Hu, F., Ge, X., Lei, J., Yu, S., Wang, T., Zhou, Q., Mei, C. and Shi, Y. 2018. Structure of the human PKD1-PKD2 complex. *Science* 361(6406).

Su, S.-H., Gibbs, N.M., Jancewicz, A.L. and Masson, P.H. 2017. Molecular mechanisms of root gravitropism. *Current Biology* 27(17), pp. R964–R972.

Sukharev, S. and Sachs, F. 2012. Molecular force transduction by ion channels: diversity and unifying principles. *Journal*

of Cell Science 125(Pt 13), pp. 3075–3083.

Sundermann, G. 1983. The fine structure of epidermal lines on arms and head of postembryonic Sepia officinalis and Loligo vulgaris (Mollusca, Cephalopoda). *Cell and Tissue Research* 232(3), pp. 669–677.

Svensen, C. and Kiørboe, T. 2000. Remote prey detection in Oithona similis: hydromechanical versus chemical cues. *Journal of plankton research* 22(6), pp. 1155–1166.

Takagi, D. and Hartline, D.K. 2018. Directional Hydrodynamic Sensing by Free-Swimming Organisms. *Bulletin of Mathematical Biology* 80(1), pp. 215–227.

Takagi, S., Cocanougher, B.T., Niki, S., Miyamoto, D., Kohsaka, H., Kazama, H., Fetter, R.D., Truman, J.W., Zlatic, M., Cardona, A. and Nose, A. 2017. Divergent Connectivity of Homologous Command-like Neurons Mediates Segment-Specific Touch Responses in Drosophila. *Neuron*.

Talavera, G. and Castresana, J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56(4), pp. 564–577.

Tamar, H. 1974. Further Studies on Halteria. *Acta Protozoologica* 13, pp. 177–191.

Tamar, H. 1965. The culture, structure, and locomotion of Halteria grandinella. *Acta Protozoologica*.

Tamar, H. 1979. The movements of jumping ciliates. *Archiv für Protistenkunde* 122(3–4), pp. 290–327.

Tamm, S.L. 1982. Ctenophora. In: Shelton, G. A. B. ed. *Electrical conduction and behavior in "simple" invertebrates*. Oxford: Clarendon Press, pp. 266–358.

Tamm, S.L. 2014. Formation of the statolith in the ctenophore Mnemiopsis leidyi. *The Biological Bulletin* 227(1), pp. 7–18.

Tanouye, M.A. and Wyman, R.J. 1980. Motor outputs of giant nerve fiber in Drosophila. *Journal of Neurophysiology* 44(2), pp. 405–421.

Tautz, J., Masters, W.M., Aicher, B. and Markl, H. 1981. A new type of water vibration receptor on the crayfish antenna. *Journal of Comparative Physiology ? A* 144(4), pp. 533–541.

Tautz, J. and Sandeman, D.C. 1980. The Detection of Waterborne Vibration by Sensory Hairs on the Chelae of the Crayfish. *Journal of Experimental Biology*.

Temereva, E.N. and Tsitrin, E.B. 2014. Organization and metamorphic remodeling of the nervous system in juveniles of Phoronopsis harmeri (Phoronida): insights into evolution of the bilaterian nervous system. *Frontiers in zoology* 11, p. 35.

Tessmar-Raible, K., Raible, F., Christodoulou, F., Guy, K., Rembold, M., Hausen, H. and Arendt, D. 2007. Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. *Cell* 129(7), pp. 1389–1400.

Thiel, D., Bauknecht, P., Jékely, G. and Hejnol, A. 2017. An ancient FMRFamide-related peptide-receptor pair induces defence behavior in a brachiopod larva. *Open biology* 7(8).

Thomas, J.B. and Wyman, R.J. 1984. Mutations altering synaptic connectivity between identified neurons in Drosophila. *The Journal of Neuroscience* 4(2), pp. 530–538.

Tinbergen, N. 1963. On aims and methods of Ethology . *Ethology : formerly Zeitschrift fur Tierpsychologie* (20), pp. 410–433.

Titelman, J. and Kiørboe, T. 2003. Predator avoidance by nauplii. *Marine Ecology Progress Series* 247, pp. 137–149.

Tomer, R., Denes, A.S., Tessmar-Raible, K. and Arendt, D. 2010. Profiling by image registration reveals common origin of annelid mushroom bodies and vertebrate pallium. *Cell* 142(5), pp. 800–809.

Tosches, M.A., Bucher, D., Vopalensky, P. and Arendt, D. 2014. Melatonin signaling controls circadian swimming behavior in marine zooplankton. *Cell* 159(1), pp. 46–57.

Tracey, W.D., Wilson, R.I., Laurent, G. and Benzer, S. 2003. painless, a Drosophila gene essential for nociception. *Cell*

113(2), pp. 261–273.

Trager, G., Achituv, Y. and Genin, A. 1994. Effects of prey escape ability, flow speed, and predator feeding mode on zooplankton capture by barnacles. *Marine biology* 120(2), pp. 251–259.

Trang, N.V., Choisy, M., Nakagomi, T., Chinh, N.T.M., Doan, Y.H., Yamashiro, T., Bryant, J.E., Nakagomi, O. and Anh, D.D. 2015. Determination of cut-off cycle threshold values in routine RT-PCR assays to assist differential diagnosis of norovirus in children hospitalized for acute gastroenteritis. *Epidemiology and Infection* 143(15), pp. 3292–3299.

Troconis, E.L., Ordoobadi, A.J., Sommers, T.F., Aziz-Bose, R., Carter, A.R. and Trapani, J.G. 2017. Intensity-dependent timing and precision of startle response latency in larval zebrafish. *The Journal of Physiology* 595(1), pp. 265–282.

Tsiokas, L., Kim, E., Arnould, T., Sukhatme, V.P. and Walz, G. 1997. Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. *Proceedings of the National Academy of Sciences of the United States of America* 94(13), pp. 6965–6970.

Turner, H.N., Armengol, K., Patel, A.A., Himmel, N.J., Sullivan, L., Iyer, S.C., Bhattacharya, S., Iyer, E.P.R., Landry, C., Galko, M.J. and Cox, D.N. 2016. The TRP Channels Pkd2, NompC, and Trpm Act in Cold-Sensing Neurons to Mediate Unique Aversive Behaviors to Noxious Cold in Drosophila. *Current Biology* 26(23), pp. 3116–3128.

Tuthill, J.C. and Wilson, R.I. 2016. Mechanosensation and adaptive motor control in insects. *Current Biology* 26(20), pp. R1022–R1038.

Tytell, E.D. and Lauder, G.V. 2002. The C-start escape response of Polypterus senegalus: bilateral muscle activity and variation during stage 1 and 2. *The Journal of Experimental Biology* 205(Pt 17), pp. 2591–2603.

Tzetlin, A.B. and Filippova, A.V. 2005. Muscular system in polychaetes (Annelida). *Hydrobiologia* 535–536(1), pp. 113–126.

Uexküll, J.B. and Kriszat, G. 1934. *Streifzüge durch die Umwelten von Tieren und Menschen Ein Bilderbuch unsichtbarer Welten*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Veedin-Rajan, V.B., Fischer, R.M., Raible, F. and Tessmar-Raible, K. 2013. Conditional and specific cell ablation in the marine annelid Platynereis dumerilii. *Plos One* 8(9), p. e75811.

Venkatachalam, K. and Montell, C. 2007. TRP channels. *Annual Review of Biochemistry* 76, pp. 387–417.

Verasztó, C., Gühmann, M., Jia, H., Rajan, V.B.V., Bezares-Calderón, L.A., Piñeiro-Lopez, C., Randle, N., Shahidi, R., Michiels, N.K., Yokoyama, S., Tessmar, K. and Jékely, G. 2018. Ciliary and rhabdomeric photoreceptor-cell circuits form a spectral depth gauge in marine zooplankton. *eLife* 7.

Verasztó, C., Ueda, N., Bezares-Calderón, L.A., Panzera, A., Williams, E.A., Shahidi, R. and Jékely, G. 2017. Ciliomotor circuitry underlying whole-body coordination of ciliary activity in the Platynereis larva. *eLife* 6.

Vergara, H.M., Bertucci, P.Y., Hantz, P., Tosches, M.A., Achim, K., Vopalensky, P. and Arendt, D. 2017. Whole-organism cellular gene-expression atlas reveals conserved cell types in the ventral nerve cord of Platynereis dumerilii. *Proceedings of the National Academy of Sciences of the United States of America* 114(23), pp. 5878–5885.

Viitasalo, M., Kiørboe, T., Flinkman, J., Pedersen, L.W. and Visser, A.W. 1998. Predation vulnerability of planktonic copepods:consequences of predator foraging strategies and prey sensory abilities. *Marine Ecology Progress Series* 175, pp. 129–142.

Vij, S., Rink, J.C., Ho, H.K., Babu, D., Eitel, M., Narasimhan, V., Tiku, V., Westbrook, J., Schierwater, B. and Roy, S. 2012. Evolutionarily ancient association of the FoxJ1 transcription factor with the motile ciliogenic program. *PLoS Genetics* 8(11), p. e1003019.

Visser, A.W. 2001. Hydromechanical signals in the plankton. *Marine Ecology Progress Series* 222, pp. 1–24.

Vopalensky, P. and Kozmik, Z. 2009. Eye evolution: common use and independent recruitment of genetic components. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364(1531), pp. 2819–2832.

Voronezhskaya, E.E., Tsitrin, E.B. and Nezlin, L.P. 2003. Neuronal development in larval polychaete Phyllodoce maculata (Phyllodocidae). *The Journal of Comparative Neurology* 455(3), pp. 299–309.

Walker, R.G., Willingham, A.T. and Zuker, C.S. 2000. A Drosophila mechanosensory transduction channel. *Science* 287(5461), pp. 2229–2234.

Wang, L.-X., Niu, C.-D., Zhang, Yan, Jia, Y.-L., Zhang, Y.-J., Zhang, Yue, Zhang, Y.-Q., Gao, C.-F. and Wu, S.-F. 2018. The NompC channel regulates Nilaparvata lugens proprioception and gentle-touch response. *Insect Biochemistry and Molecular Biology*.

Wang, Q., Shui, B., Kotlikoff, M.I. and Sondermann, H. 2008. Structural basis for calcium sensing by GCaMP2. *Structure* 16(12), pp. 1817–1827.

Watson, G.M., Mire, P. and Hudson, R.R. 1997. Hair bundles of sea anemones as a model system for vertebrate hair bundles. *Hearing Research* 107(1–2), pp. 53–66.

Webb, P.W. 1976. The effect of size on the fast-start performance of rainbow trout Salmo cairdneri, and a consideration of piscivorous predator-prey interactions. *The Journal of Experimental Biology* 65(1), pp. 157–177.

Weihs, D. 1973. The mechanism of rapid starting of slender fish1. *Biorheology* 10(3), pp. 343–350.

Westfall, J.A., Sayyar, K.L. and Elliott, C.F. 1998. Cellular origins of kinocilia, stereocilia, and microvilli on tentacles of sea anemones of the genus calliactis (cnidaria: anthozoa). *Invertebrate Biology* 117(3), p. 186.

Westneat, M.W., Hale, M.E., Mchenry, M.J. and Long, J.H. 1998. Mechanics of the fast-start: muscle function and the role of intramuscular pressure in the escape behavior of amia calva and polypterus palmas. *The Journal of Experimental Biology* 201 (Pt 22), pp. 3041–3055.

Wiederhold, M.L. 1976. Mechanosensory transduction in "sensory" and "motile" cilia. *Annual review of biophysics and bioengineering* 5, pp. 39–62.

Wiersma, C.A.G. 1938. Function of the giant fibers of the central nervous system of the crayfish. *Experimental biology and medicine* 38(5), pp. 661–662.

Wiersma, C.A.G. 1947. Giant nerve fiber system of the crayfish; a contribution to comparative physiology of synapse. *Journal of Neurophysiology* 10(1), pp. 23–38.

Wiersma, C.A.G. and Ikeda, K. 1964. Interneurons commanding swimmeret movements in the crayfish, Procambarus clarki (girard). *Comparative biochemistry and physiology* 12(4), pp. 509–525.

Wiese, K. 1976. Mechanoreceptors for near-field water displacements in crayfish. *Journal of Neurophysiology* 39(4), pp. 816–833.

Wilkes, M., Madej, M.G., Kreuter, L., Rhinow, D., Heinz, V., De Sanctis, S., Ruppel, S., Richter, R.M., Joos, F., Grieben, M., Pike, A.C.W., Huiskonen, J.T., Carpenter, E.P., Kühlbrandt, W., Witzgall, R. and Ziegler, C. 2017. Molecular insights into lipid-assisted Ca2+ regulation of the TRP channel Polycystin-2. *Nature Structural & Molecular Biology* 24(2), pp. 123–130.

Williams, E.A., Conzelmann, M. and Jékely, G. 2015. Myoinhibitory peptide regulates feeding in the marine annelid Platynereis. *Frontiers in zoology* 12(1), p. 1.

Williams, E.A., Verasztó, C., Jasek, S., Conzelmann, M., Shahidi, R., Bauknecht, P., Mirabeau, O. and Jékely, G. 2017. Synaptic and peptidergic connectome of a neurosecretory center in the annelid brain. *eLife* 6.

Williamson, C.E. 1987. Predator-prey interactions between omnivorous diaptomid copepods and rotifers: The role of prey morphology and behavior1. *Limnology and Oceanography Letters* 32(1), pp. 167–177.

Williamson, C.E. 1980. The predatory behavior of*Mesocyclops edax*: Predator preferences, prey defenses, and starvation-

induced changes1. *Limnology and Oceanography Letters* 25(5), pp. 903–909.

Wilson, D.M. 1960. Nervous Control of Movement in Annelids. *Journal of Experimental Biology*.

Wilson, D.M. and Bullock, T.H. 1959. Electrical Recordings from giant fibers and muscle in phoronids. In: pp. 518–519.

Wilson, D.P. 1929. The larvæ of the british sabellarians. *Journal of the Marine Biological Association of the UK* 16(01), p. 221.

Windoffer, R. and Westheide, W. 1988. The Nervous System of the Male *Dinophilus gyrociliatus* (Annelida: Polychaeta). I. Number, Types and Distribution Pattern of Sensory Cells. *Acta Zoologica* 69(1), pp. 55–64.

Wine, J.J. and Krasne, F.B. 1972. The organization of escape behavior in the crayfish. *The Journal of Experimental Biology* 56(1), pp. 1–18.

Woo, S.-H., Lukacs, V., de Nooij, J.C., Zaytseva, D., Criddle, C.R., Francisco, A., Jessell, T.M., Wilkinson, K.A. and Patapoutian, A. 2015. Piezo2 is the principal mechanotransduction channel for proprioception. *Nature Neuroscience* 18(12), pp. 1756–1762.

Woo, S.-H., Ranade, S., Weyer, A.D., Dubin, A.E., Baba, Y., Qiu, Z., Petrus, M., Miyamoto, T., Reddy, K., Lumpkin, E.A., Stucky, C.L. and Patapoutian, A. 2014. Piezo2 is required for Merkel-cell mechanotransduction. *Nature* 509(7502), pp. 622–626.

Wood, J.M. 2015. Bacterial responses to osmotic challenges. *The Journal of General Physiology* 145(5), pp. 381–388.

Xu, J., Mathur, J., Vessières, E., Hammack, S., Nonomura, K., Favre, J., Grimaud, L., Petrus, M., Francisco, A., Li, J., Lee, V., Xiang, F.-L., Mainquist, J.K., Cahalan, S.M., Orth, A.P., Walker, J.R., Ma, S., Lukacs, V., Bordone, L., Bandell, M., Laffitte, B., Xu, Y., Chien, S., Henrion, D. and Patapoutian, A. 2018. GPR68 senses flow and is essential for vascular physiology. *Cell* 173(3), p. 762–775.e16.

Yan, Z., Zhang, W., He, Y., Gorczyca, D., Xiang, Y., Cheng, L.E., Meltzer, S., Jan, L.Y. and Jan, Y.N. 2013. Drosophila NOMPC is a mechanotransduction channel subunit for gentle-touch sensation. *Nature* 493(7431), pp. 221–225.

Yang, W. and Yuste, R. 2017. In vivo imaging of neural activity. *Nature Methods* 14(4), pp. 349–359.

Yeomans, J.S., Li, L., Scott, B.W. and Frankland, P.W. 2002. Tactile, acoustic and vestibular systems sum to elicit the startle reflex. *Neuroscience and Biobehavioral Reviews* 26(1), pp. 1–11.

Yoder, B.K., Hou, X. and Guay-Woodford, L.M. 2002. The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *Journal of the American Society of Nephrology* 13(10), pp. 2508–2516.

Yoshiba, S., Shiratori, H., Kuo, I.Y., Kawasumi, A., Shinohara, K., Nonaka, S., Asai, Y., Sasaki, G., Belo, J.A., Sasaki, H., Nakai, J., Dworniczak, B., Ehrlich, B.E., Pennekamp, P. and Hamada, H. 2012. Cilia at the node of mouse embryos sense fluid flow for left-right determination via Pkd2. *Science* 338(6104), pp. 226–231.

Yoshimura, K. 1996. A novel type of mechanoreception by the flagella of Chlamydomonas. *The Journal of Experimental Biology* 199(Pt 2), pp. 295–302.

Young, S.R., Dedwylder, R.D. and Friesen, W.O. 1981. Responses of the medicinal leech to water waves. *Journal of Comparative Physiology ? A* 144(1), pp. 111–116.

Yu, X., Ng, C.P., Habacher, H. and Roy, S. 2008. Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nature Genetics* 40(12), pp. 1445–1453.

Zaret, T.M. and Suffern, J.S. 1976. Vertical migration in zooplankton as a predator avoidance mechanism1. *Limnology and Oceanography Letters* 21(6), pp. 804–813.

Zeng, W.-Z., Marshall, K.L., Min, S., Daou, I., Chapleau, M.W., Abboud, F.M., Liberles, S.D. and Patapoutian, A. 2018. PIEZOs mediate neuronal sensing of blood pressure and the baroreceptor reflex. *Science* 362(6413), pp. 464–467.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Flicek, P., et al. 2018. Ensembl 2018. *Nucleic Acids Research*

46(D1), pp. D754–D761.

Zhadan, P.M. and Semen'kov, P.G. 1984. An electrophysiological study of the mechanoreceptory function of abdominal sense organ of the scallop Patinopecten yessoensis (JAY). *Comparative Biochemistry and Physiology Part A: Physiology* 78(4), pp. 865–870.

Zhadan, P.M., Sizov, A.V. and Dautov, S.S. 2004. Ultrastructure of the abdominal sense organ of the scallop Mizuchopecten yessoensis (Jay). *Cell and Tissue Research* 318(3), pp. 617–629.

Zhang, H., Yue, X., Cheng, H., Zhang, X., Cai, Y., Zou, W., Huang, G., Cheng, L., Ye, F. and Kang, L. 2018. OSM-9 and an amiloride-sensitive channel, but not PKD-2, are involved in mechanosensation in C. elegans male ray neurons. *Scientific reports* 8(1), p. 7192.

Zhang, L., Hastings, M.H., Green, E.W., Tauber, E., Sladek, M., Webster, S.G., Kyriacou, C.P. and Wilcockson, D.C. 2013. Dissociation of circadian and circatidal timekeeping in the marine crustacean Eurydice pulchra. *Current Biology* 23(19), pp. 1863–1873.

Zhang, M., Li, X., Zheng, H., Wen, X., Chen, S., Ye, J., Tang, S., Yao, F., Li, Y. and Yan, Z. 2018. Brv1 is required for drosophila larvae to sense gentle touch. *Cell reports* 23(1), pp. 23–31.

Zhang, S., Arnadottir, J., Keller, C., Caldwell, G.A., Yao, C.A. and Chalfie, M. 2004. MEC-2 is recruited to the putative mechanosensory complex in C. elegans touch receptor neurons through its stomatin-like domain. *Current Biology* 14(21), pp. 1888–1896.

Zhang, W., Cheng, L.E., Kittelmann, M., Li, J., Petkovic, M., Cheng, T., Jin, P., Guo, Z., Göpfert, M.C., Jan, L.Y. and Jan, Y.N. 2015. Ankyrin repeats convey force to gate the NOMPC mechanotransduction channel. *Cell* 162(6), pp. 1391–1403.

Zhang, W., Yan, Z., Jan, L.Y. and Jan, Y.N. 2013. Sound response mediated by the TRP channels NOMPC, NANCHUNG, and INACTIVE in chordotonal organs of Drosophila larvae. *Proceedings of the National Academy of Sciences of the United States of America* 110(33), pp. 13612–13617.

Zhang, X.-H., Tee, L.Y., Wang, X.-G., Huang, Q.-S. and Yang, S.-H. 2015. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular therapy. Nucleic acids* 4, p. e264.

Zhang, Y.V., Aikin, T.J., Li, Z. and Montell, C. 2016. The basis of food texture sensation in drosophila. *Neuron* 91(4), pp. 863–877.

Zhong, L., Hwang, R.Y. and Tracey, W.D. 2010. Pickpocket is a DEG/ENaC protein required for mechanical nociception in Drosophila larvae. *Current Biology* 20(5), pp. 429–434.

Žídek, R., Machoň, O. and Kozmik, Z. 2018. Wnt/β-catenin signalling is necessary for gut differentiation in a marine annelid, Platynereis dumerilii. *EvoDevo* 9, p. 14.

Zischewski, J., Fischer, R. and Bortesi, L. 2017. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnology advances* 35(1), pp. 95–104.

Zottoli, S.J. 1977. Correlation of the startle reflex and Mauthner cell auditory responses in unrestrained goldfish. *The Journal of Experimental Biology* 66(1), pp. 243–254.