pezo-1 function in *Caenorhabditis elegans*

Thesis by

Katherine I. Brugman

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended May 11, 2020)

Katherine I. Brugman
ORCID: 0000-0003-2625-2903

ACKNOWLEDGEMENTS

I would like to thank my family for bearing with me through this very long process and working with my, often insane, scheduling concerns with an understanding that has made my time here so much more bearable. I would also like to thank my friends, both in person and online, for being there to listen to science they likely didn't understand or particularly care about, but nevertheless responded with enthusiasm and support. Many of the figures in this thesis have been emailed or posted or scribbled on napkins at least once before to a friend because I knew they'd at least appreciate that it looked like science.

I would also like to thank the lab, for their endless good advice and critiques that have helped in fixing the things that have needed to be fixed and for their support when I didn't believe my own results. There are so many of you that I'd like to thank, and everyone has helped in some way, but in particular, I would like to thank Daniel Leighton, James Lee, Hillel Schwartz, Mengyi Cao, Christopher Cronen, Barbara Perry, and of course Paul Sternberg, without whose constant belief in my data I might still be running the same assay, attempting to get one hundred percent accuracy.

I would also like to thank my collaborators, Xiaofei Bai and Vladislav Susoy, who have helped me tremendously in both confirming my results and solidifying the stories I have come up with. Vladislav Susoy has kindly provided data for Figure 6 of Chapter 2 to help tie up that particular story.

And lastly, I would like to thank my brother, John Brugman, who at the last minute wrote a training algorithm to measure the acceleration of worms to test a hypothesis I had. He has provided amazing tech support in these trying times.

ABSTRACT

The piezo class of mechanosensitive ion channels is a recently discovered class of cation channels with orthologs found in every phylogenic clade aside from yeast and bacteria. They are large channels, both in gene length and in overall diameter, with the diameter of the human PIEZO2 protein measuring in at 280 Å in its full homotrimeric form. In addition, like many similar mechanosensitive channels, such as the DEG/ENaC channels, TRP channels, and TREK/TRAAK channels, they have been linked to a number of different functions within *drosophila*, zebrafish, and mice, including light touch, nociception, blood cell volume regulation, vascular development, and neuropathic pain. Structurally, this channel is intriguing as it possesses no known structural motifs and is organized into a central pore with a cap, surrounded by three "propeller blade" regions that are theorized to anchor the channel to the membrane and control gating through hydrophobic mismatch based on membrane curvature.

The *C. elegans* piezo, *pezo-1*, has not yet been fully characterized, even though a crystal structure of part of this particular piezo was used to assist in the resolution of the first set of cryo-EM images. Here, I generated a number of GFP transcriptional fusions of non-coding potential promoter regions to track the expression of the *pezo-1* gene in the animal, using these expression patterns to design further experiments. From these expression patterns, I identified expression in a number of neurons of the *C. elegans* male tail, the primary mating apparatus of the male, and identified these neurons as key neurons as relating to mating. In particular, I identified neurons HOB, PCB, PCC, and various ray neurons as potential candidates, which are ciliated neurons theorized to have

mechanosensitive properties. In addition, I also identified expression in the vulva muscle and spermatheca of the hermaphrodite, both theorized to be involved in ovulation and egg-laying processes.

From there, I designed CRISPR/Cas9 mutants with defects in *pezo-1* in order to investigate the potential link between the *pezo-1* expression in those neurons and mating behavior via a mating assay. Similarly, I devised a fecundity assay to investigate the link between *pezo-1* expression in ovulation organs and progeny survival. I have discovered that *pezo-1* has function in both of these areas, with *pezo-1* mutant males demonstrating discrete mating defects that correlate with the expression pattern seen from the GFP transcriptional fusion mutants and with *pezo-1* mutant hermaphrodites having much smaller brood sizes than wildtype hermaphrodites. However, I have also discovered that the processes this mechanotransducer is involved in are also more complex than I originally believed, as I also discovered that *pezo-1* appears to interact with another mechanotransducer, *trp-4*, illuminating some potentially novel pathway considerations for how these channels overlap in function.

PUBLISHED CONTENT AND CONTRIBUTIONS

Wan-Rong Wong, Katherine I Brugman, Shayda Maher, Jun Young Oh, Kevin Howe, Mihoko Kato, Paul W Sternberg, Autism-associated missense genetic variants impact locomotion and neurodevelopment in *Caenorhabditis elegans*, *Human Molecular Genetics*, Volume 28, Issue 13, 1 July 2019, Pages 2271-2281, https://doi.org/10.1093/hmg/ddz051

Katherine designed the fecundity assay, made a large number of the mutant strains used in the study, and coordinated between different team members. She also assisted in writing and editing the final paper.

Xiaofei Bai, Jeff Bouffard, Avery Lord, Katherine Brugman, Paul W. Sternberg, Erin J. Cram, Andy Golden, Caenorhabditis elegans PIEZO Channel Coordinates Multiple Reproductive Tissues to Govern Ovulation (Unpublished)

Katherine designed the STOP-IN *pezo-1* mutant and the small isoform deletion *pezo-1* mutant and performed the fecundity assays with both of these mutants.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND/OR TABLES

Chapter 1

INTRODUCTION

1.1 Background

Why Study Mechanosensation?

Thinking back to the beginning of my interest in science, I remember a lesson in high school in which my biology teacher asked the class what we considered to be the most important sensory modality. Most of the class, myself included, chose eyesight, thinking of the primary way we perceive the world. But after spending the rest of the class discussing the myriad ways mechanosensation affects how we exist in the world and respond to everchanging stimuli, everyone had changed their answer to touch.

The first question to answer is what, exactly, defines a mechanosensory response. At its base level, mechanotransduction has been defined by Voglis and Tavernarakis as a "conversion of mechanical stimulus into biological response," which covers both the obvious cases such as touch sensation, nociception, proprioception, and hearing, but also less obvious uses of mechanotransduction, such as in development and homeostasis (1–7). Just from this definition, we can see the sheer versatility of mechanosensitive channels and their functions, not only in rapid response to sudden changes in environment but also in stabilizing prolonged processes. Therefore, it should come as little surprise there exists a large variety of mechanotranducers, such that they can cover a variety of different roles, even within the same organism.

The second question to answer is what are the general features of mechanotranducers themselves. First of all, mechanotransducers tend to exist in the

plasma membrane of cells, where the mechanical stimuli that trigger them occur and where they are best positioned to activate in response to those stimuli (8). Second, mechanotransducers tend to be ion channels, which amplify the speed of the response, which is often necessary in the case of a changing environment (9). Third, mechanotransducers tend to involve some manner of anchoring, either to extracellular structures as is the case with MET channels in the ear, the cytoskeleton and extracellular matrix as in degenerin channels in *C. elegans*, or hydrophobic mismatch as in TREK1 and TRAAK channels (9–12). This anchoring is essential to the channel's function as it is the method by which the channel can open or close under an external mechanical change by way of giving it some form of leverage.

So, now that we have some idea of how we define a mechanotransducer and what it requires to function, our third question is how this basic information translates into these channels offering such a wide array of mechanical responses to stimuli. To answer this, I use two key examples.

The most basic examples for physiological processes tend to revert to bacteria, as their single-celled nature renders them vulnerable to system failure, forcing them to have robust cellular processes, and this is no different for processes performed by mechanotransducers. Bacteria use the mechanosensitive ion channels MscS (Mechanosensitive channel Small conductance) and MscL (Mechanosensitive channel Large conductance) to sense and respond to changes in osmolarity, which is a fairly important homeostatic process as the bacteria traverses various environments (8, 13–15). Both channels are made up of oligomeric subunits—MscS forms a homoheptamer, MscL

a homopentamer—that come together and utilize hydrophobic mismatch of their T1 and T2 transmembrane helices crisscrossed in the plasma membrane (16–18) (Figure 1.1a). Should the membrane expand due to osmotic pressure caused by low external osmolarity, the firm bonds between those helices and the rest of the channel provide the tension necessary to force the channel to spiral open, releasing excess cytoplasm into the environment and maintaining homeostasis (19–24). The primary difference between these channels, and the one that informs their specific function, is size. MscL, as a much larger channel with a pore over twice the diameter of that of MscS, provokes a far stronger response when activated and requires a far greater shift in membrane curvature in order to activate (17, 23, 25, 26). As a result, MscS acts as a regulatory valve, responding to smaller shifts in osmotic stress, and MscL acts as an emergency valve, preventing lysis from hypo-osmotic shock (2, 27).

My second example is more complex and features prominently in *Caenorhabditis elegans* light touch sensation, the DEG/ENaC sodium channel (1). Unlike MscS and MscL, this mechanotransducer is not constructed from a single gene, but instead requires a complex of different proteins in order to function properly (28–30). In *C. elegans*, the main channel consists of heterotrimer composed of two different proteins, MEC-4 and MEC-10, while accessory proteins, MEC-2 and UNC-24, associate with the channel to alter its gating mechanism with a stomatin-like domain (28–33). MEC-6, meanwhile, appears to act as a chaperone and stabilizing agent to the channel, preventing it from degrading and recruiting MEC-2 to bind to it (29, 30, 32, 34, 35). Part of the reason this complex requires so much assistance is that, also unlike MscS and MscL, the MEC-

4/MEC-10 complex requires external tethering to provide the leverage and tension to open the channel. MEC-2 and UNC-24 complex together in the intracellular part of the channel, which may bind the channel to MEC-7/MEC-12 microtubule structures inside the cell, while EGF containing proteins, MEC-1, likely binds to the extracellular portion of the complex to link it to a MEC-5 collagen-type protein and EGF containing MEC-9 in the extracellular matrix (32, 36–38) (Figure 1.1b). While there is still some debate as to the degree that the inner tether is involved in the mechanosensory portion of the channel opening, pressure on the extracellular matrix could be sufficient to push the channel into or out of the cell, distorting the membrane and forcing a hydrophobic mismatch tension that would force the channel open (29, 39, 40). As this channel responds to light touch sensation, this particular type of gating ties into its function.

These are just two of the many examples of mechanosensitive ion channels and the differences abound, from basic structure, to complexity, to gating. And among the others that have so far been discovered, more intricacies and complexities that enable them to carry out the varied responses for which these channels are responsible, which leads to the final question of this section: why study mechanosensation? I would like to say that because these channels are important in how we perceive the world and our own bodies, they have an inherent beauty in both their complexity and simplicity, and with the development of CRISPR/Cas9 and GCamP, we have better genetic tools than ever before to study them *in vivo.*

Now, there's one more mechanosensitive ion channel class that I would like to discuss, piezo.

The Piezo Class of Mechanosensative Ion Channels

Coming to graduate school, was already aware of a number of mechanosensitive ion channels. My first neurobiology class covered the TRP superfamily of mechanoreceptors and the DEG/ENaC family as the mechanoreceptors responsible for many sensory modalities such as pain sensation, heat sensation, touch, and osmotic and volume regulation (39, 41). Therefore, I was intrigued when I learned about a newly discovered class of mechanosensative ion channels, dubbed piezo channels. From the very start of my relationship with this protein family, this channel perplexed me: it was expressed in a wide variety of structures in humans, namely bladder, colon, kidney, lung, and skin and based on the gene sequence, it lacked any easily identifiable motifs more complex than alpha helixes and beta sheets, which left early structural prediction a mystery (42). And yet, from early *in vitro* experiments, the channel was non-selective for all cations (with a preference for calcium) and responded to mechanosensory impulses (42–44).

This feature made piezo channels an intriguing subject for research as they implied an entirely novel structure for a mechanosensative ion channel, and with that, a potentially novel function. At the time, early studies of piezo genomics showed that piezo protein orthologs were extant in nearly every clade of organism other than yeast and bacteria (42). Vertebrates were shown to possess two piezo proteins with 50% sequence

identity with each other, invertebrates had only one, and piezo orthologs between different species typically had five residues conserved between all species (45, 46). This conservation was mostly located in the C-terminus, leading to the early conclusion that this region of the gene encoded the pore or some similarly essential part of the protein (45, 46) (Figure 1.2).

The observed functions in other model organisms at the time also made piezo a compelling research subject due to the range of observed functions in which piezo was implicated. The *Drosophila* piezo was the first piezo to have a mechanosensory modality categorized *in vivo*, with early experiments showing piezo function in mechanosensory nociception (47). This finding was closely followed by a study in zebrafish showing a function in light touch response in its *piezo2b* and another study showing *piezo1* involved in the regulation of erythrocyte volume (48, 49). Similarly to zebrafish, mouse Piezo1 functioned in vascular development while the Piezo2 had a prominent role in light touch response as well as neuropathic pain independent of nociception (50–52).

Still, it was not until 2014 that there was any real movement toward deciphering the actual structure of the protein itself: the paper by Kamajaya et al demonstrated that a highly conserved C-terminal loop of the *C. elegans* piezo was crystallized, showing a previously uncharacterized β sandwich fold (53). I entered the fray after Aron Kamajaya contacted me to ask what piezo did in *C. elegans*, and I was similarly intrigued by what functions might be linked to a *C. elegans* piezo, as I will later discuss in this thesis. Around the same period of time, Piezo1 was shown via a mixture of cryo-electron microscopy and some crystallography in two different studies to form a homotrimeric

channel, with the conserved C-terminus forming the pore and cap regions, and the N-terminus forming flexible “propeller blades” made of transmembrane helices that were hypothesized to be the force-sensing components of the channel (54, 55). It was also confirmed in these studies that the previous *C. elegans* domain crystallized to form the cap portion of the protein, and shortly thereafter in another study, that Piezo2 followed the same structural motifs (54, 56) (Figure 1.3).

Since then, there have been several mechanisms proposed for how piezo channels might be activated, mostly focusing on Piezo1, as the existence of a structure earlier presented a base for models. Possible mechanisms included direct sensation of membrane tension via hydrophobic mismatch of propeller blades, a cytoskeleton gating pulling the channel open, pushing the cap region of the protein out of the way of the pore, or changing membrane composition entirely. The “propeller blade” transmembrane helices were a compelling aspect of the structure to be involved in the gating of the channel from the beginning, using the model of these embedded helices warping the membrane surrounding the channel such that further membrane distortions would cause the central pore to pull apart due to hydrophobic mismatch forces and allow the transmission of cations (57, 58). Electrophysiology experiments confirmed that Piezo1 could be activated by both convex and concave membrane curvature and respond to changes in membrane tension in both configurations, providing evidence to support this model (58). In addition, Piezo1 was capable of being activated in membrane “blebs”, regions of the membrane lacking cytoskeleton connections, making it unlikely that activation would come from a cytoskeleton connection (59).

As for the other mechanisms, magnetic pulling via nanoparticles revealed that both the cap and the propeller blade transmembrane regions are mechanosensitive, allowing for the possibility of sheer forces to push the cap aside for some activation alongside hydrophobic mismatch, possibly in conjunction with each other (60). In addition, activating TRPV1 channels in a system also expressing piezo channels led to inhibition of the piezo channels via depletion of phosphoinositides, an abundant lipid in the plasma membrane (61). The depletion of these membrane lipids and subsequent change in composition of the plasma membrane could alter piezo's ability to create a pucker as in the first proposed mechanism, or perhaps the lipids act on the channel in a more direct manner.

For my project, I decided to focus specifically on the function of the *C. elegans* piezo.

Caenorhabditis elegans **Physiology and Mating Behavior**

For an *in vivo* study of a complex membrane protein, *Caenorhabditis elegans*, a free-living bacterivorous roundworm of about a millimeter in length, is the ideal model organism (62). It is a self-fertilizing androdioecious species with six chromosomes—five autosomes and one sex chromosome as discovered by Victor Nigon—and males are produced through a nondisjunction event that leaves the oocyte with a single X chromosome (63, 64). It has a short generation time, can be easily grown on agar plates, has neurons that can be visualized easily with GFP markers in a living animal, and most

recently can be easily genetically manipulated via Crispr-Cas9 (62, 65–69). These features, combined with the fact that the species has only 302 neurons in the hermaphrodite and 385 neurons in the male, as well as being the only species to have a fully characterized connectome, makes it ideal for neurobiology studies (70–73).

While *C. elegans* males only make up 0.1% of the population, they have a strikingly different physiology compared to the hermaphrodite animal and represent an effective way of extending the reproductive lifecycle of the hermaphrodite adult, which typically lasts three days (64, 74–79). In addition, male mating is the primary way that *C. elegans* possesses for swapping genetic material between otherwise clonal strains (77). It is a complex behavior—the most complex that the worm undertakes—and utilizes 170 neurons, 81 of which are male specific, and 64 muscles (80).

The hermaphrodite and male differ in body morphology as well (79, 81). The hermaphrodite possesses a two-armed gonad in which the oocytes develop before they are passed through either the anterior or posterior spermatheca, held in place by the distal and spermathecal-uterine valve to be fused with a sperm, and then released into the uterus (79, 82, 83). There, the fertilized egg matures and is eventually expelled by the central vulva (81). The male, in contrast, has a single-armed gonad that only creates sperm and is connected to a copulatory fan-like tail structure, containing two spicules that aid in sperm transfer (79, 80) (Figure 1.4).

So much of this machinery is necessary, however, as mating behavior in *C. elegans* is an involved process comprised of a number of discrete steps and checkpoints. After the male detects the hermaphrodite via chemosensory processes, he moves towards

her until he contacts her, and then orients his body so that the ventral portion of the fan contacts the hermaphrodite's cuticle (84). He then begins to back around the hermaphrodite—a process that is referred to as "scanning"—and executes turns at her nose or tail when he reaches them to scan her other side. When he reaches the vulva, the behavior changes. He stops at this point and beings probing the area with his spicules for fine scale location of the vulva. When he locates it, he inserts the spicules and releases sperm (81, 85) (Figure 1.5).

Some neurological aspects of this behavior have been characterized already via previous studies. The fan structure and associated ray neurons, R(1-9)A/B, have been implicated in attachment to the hermaphrodite and appropriate turning (71, 80, 85–87). The spicule hook structure, which is located slightly anterior and ventral to the cloaca, and the two neurons attached to it, named HOA and HOB, have been implicated via ablation experiments to have a function in the male making the initial stop at the vulva (80, 85, 88, 89). The post cloacal sensilla, located on the other side of the cloaca, and its six associated neurons, PCAL/R, PCBL/R, and PCCL/R, have been implicated by other ablation experiments to be involved in the subsequent fine-scale vulva location (80, 85, 90–92). One of the neuron pairs attached to the spicules, SPCL/R, has been implicated to be involved in sensing the change in spicule position during prodding and inducing full insertion (80, 85, 90–92). The other spicule associated neurons, SPCVL/R and SPDL/R, sense vulva pheromones and coordinate sperm release (80, 85, 90–92) (Figure 1.6).

Of these neurons, the ray neurons, HOA, HOB, PCAL/R, PCBL/R, PCCL/R, and SPCL/R are all potentially mechanosensory neurons due to the nature of the response

they generate and the fact that they have all been confirmed to be ciliated neurons , which has been linked to possessing a mechanosensory property (93). However, at the point when I began my study, only HOB and a few ray neurons had been linked to specific mechanosensative ion channels (88, 94). This presented a starting point for my research, which I pursued to discover associations with the *C. elegans* piezo protein, *PEZO-1*, as well as a surprise addition of a second mechanosensor in the *trp-4* gene.

1.2 Figures

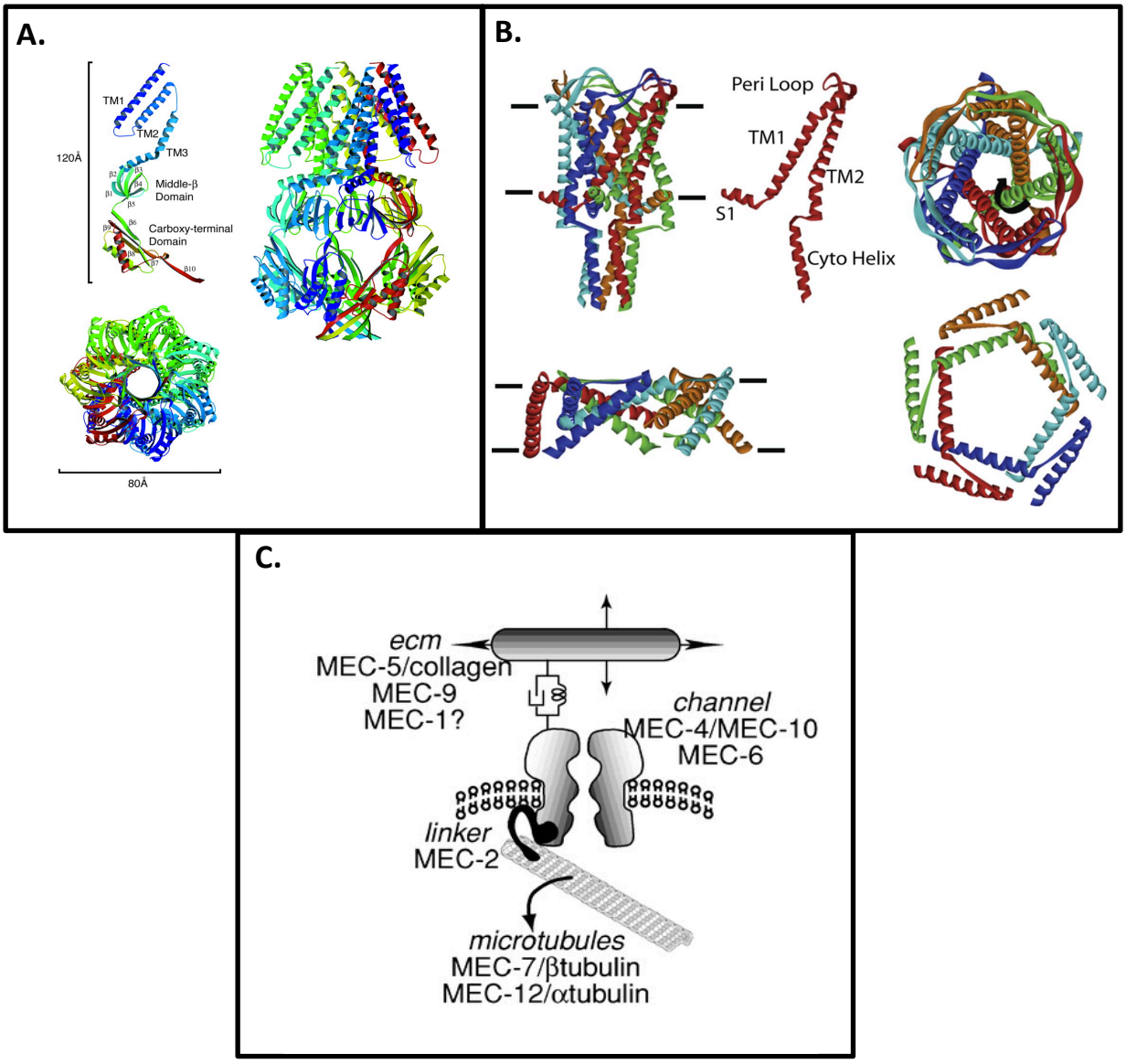


Figure 1.1 (A) MscS ribbon diagram showing homoheptameric structure and tension-bearing amphipathic helix (16). (B) MscL ribbon diagram model showing homopentameric structure and S1 amphipathic helix (2). (C) Cartoon diagram of a model

of MEC-4/MEC-10 DEG/ENaC channel gating mechanism with MEC-2 linkage (95). There is still some debate over the role of MEC-7/MEC-12 in the gating of this channel.

Figure 1.2 A sequence alignment of human, mouse, zebrafish, *drosophila*, and *C. elegans* piezo proteins. Shown here are sample alignments from the N-terminus (top) and the C-terminal alignments (bottom). The N-terminus of piezo proteins is generally less conserved across all species than the C-terminus. Alignment done with Clustal Omega, Uniprot (96, 97).

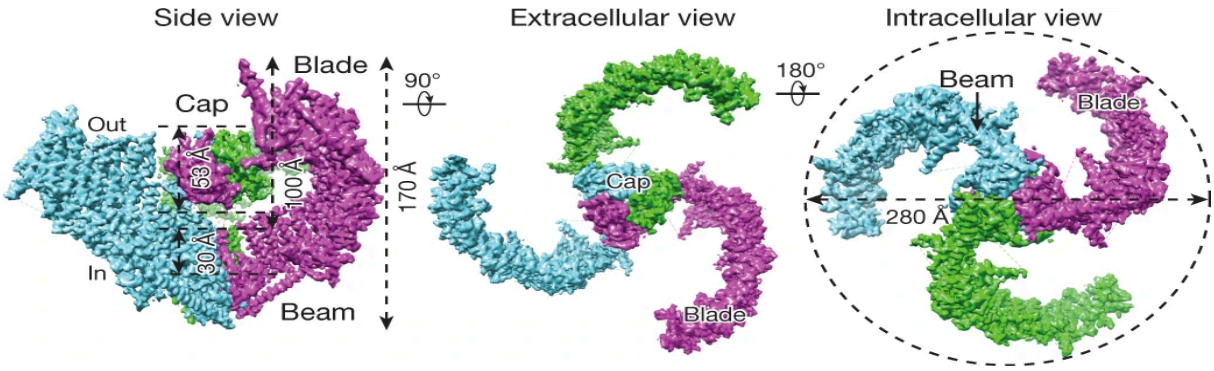


Figure 1.3 Higher resolution cryo-EM structure of Piezo2. Here we can see the central pore and cap, with the blades linked together in the homotrimeric structure and the beams (56). The current model for gating involves the blade regions acting as the primary mechanosensory apparatus by using hydrophobic mismatch to detect changes in membrane curvature.

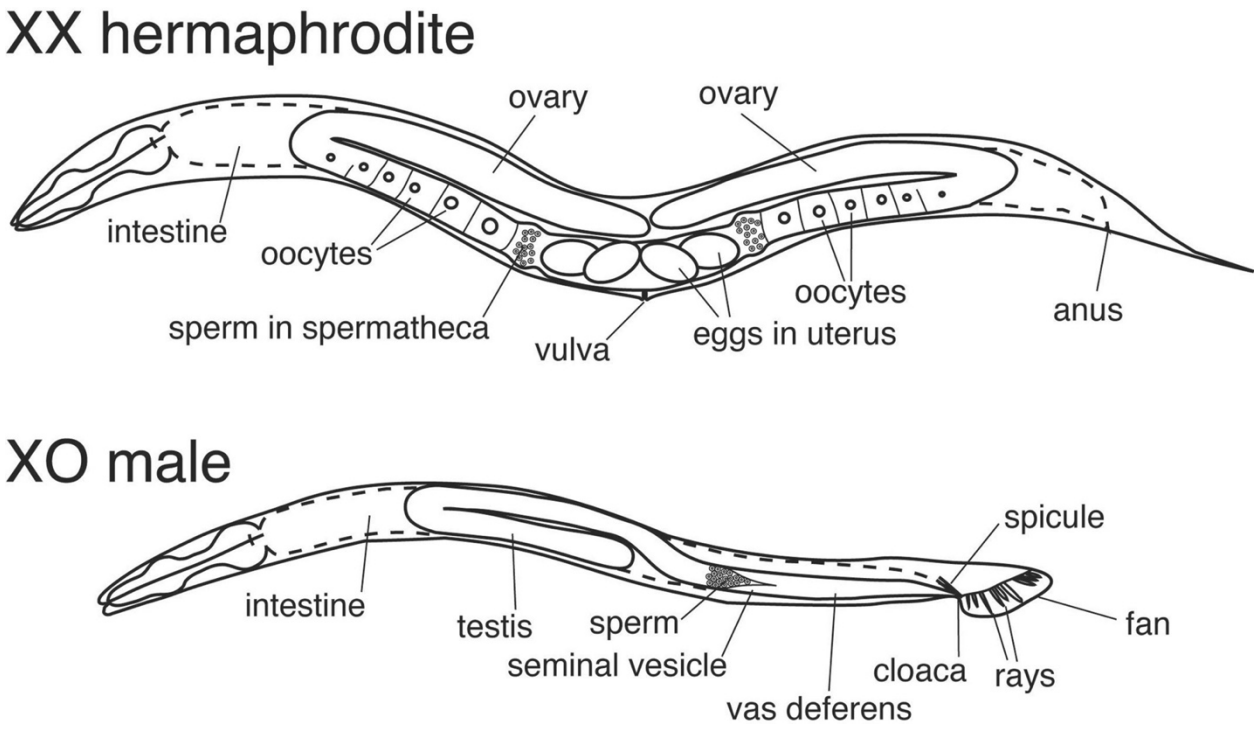


Figure 1.4 Comparison image of the hermaphrodite and male in *Caenorhabditis elegans* (98). The hermaphrodite animal has a two-armed gonad along which oocytes develop and are fertilized, while the males have a one-armed gonad that only produces sperm, along with a mating specific tail apparatus.

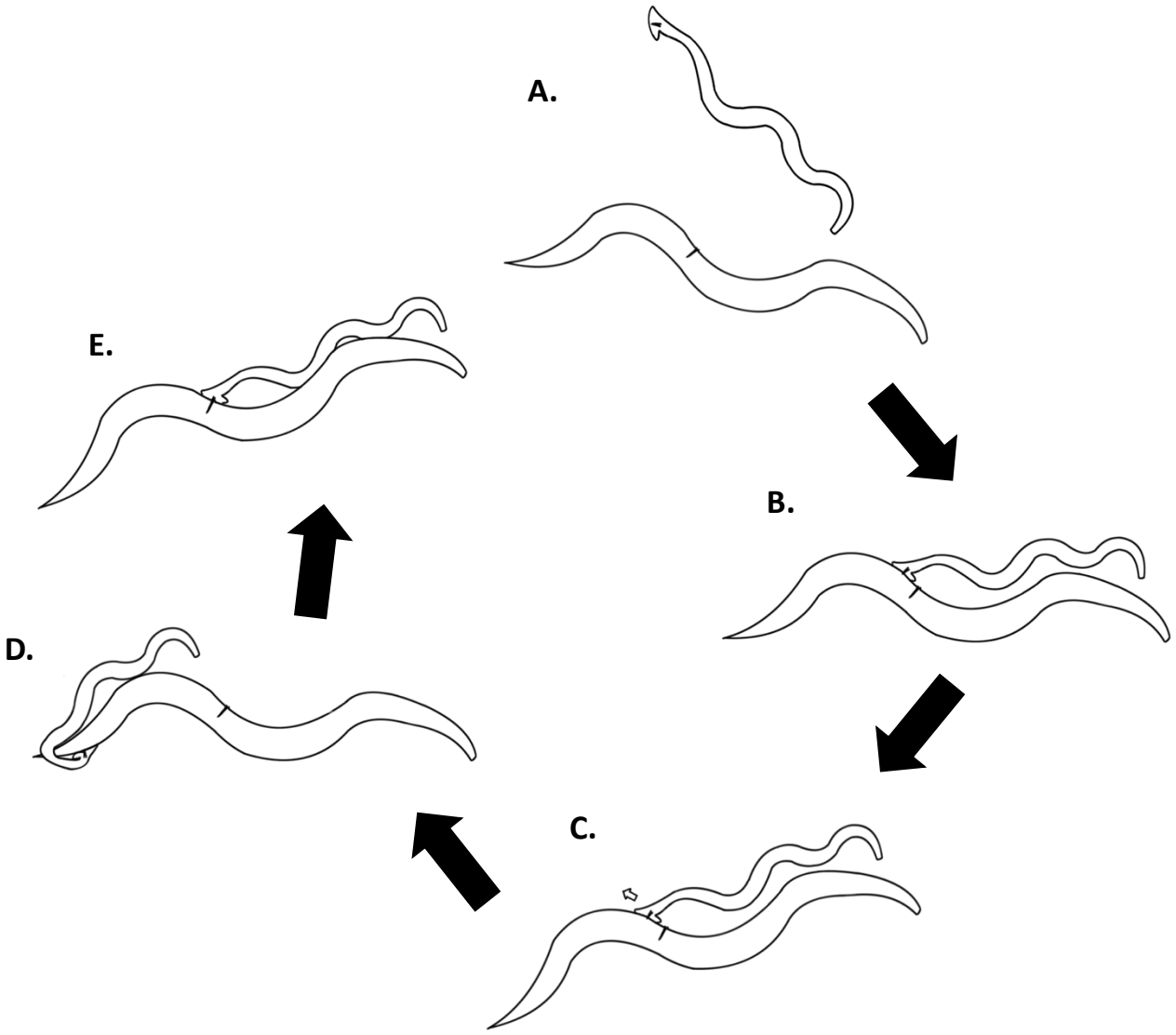
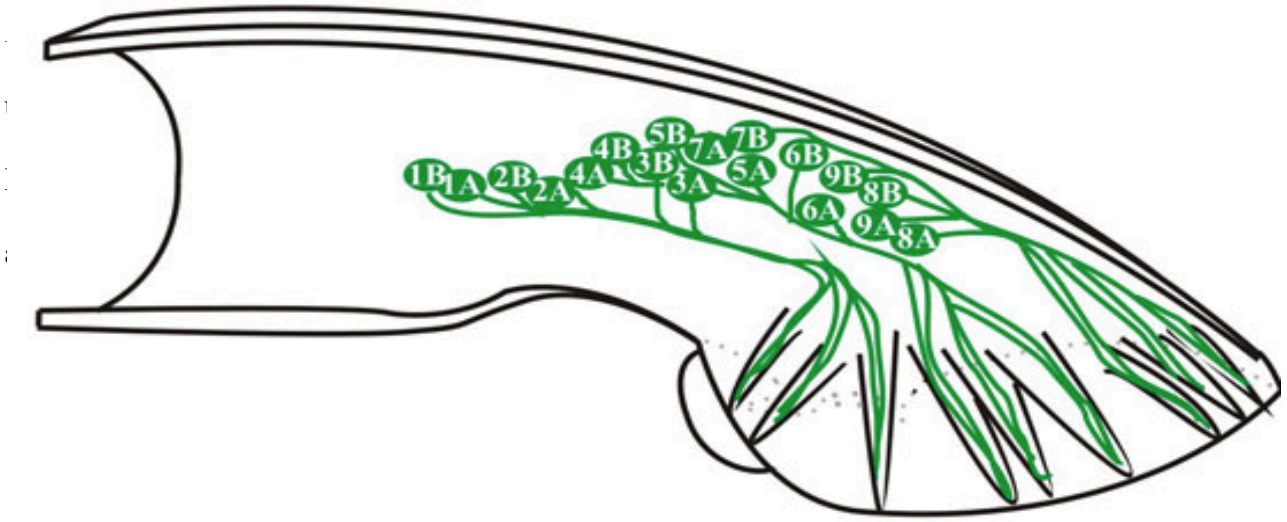


Figure 1.5 (A) The male approaches the hermaphrodite. (B) Initial contact by the male with the ventral portion of his tail. (C) The male commences backing along the hermaphrodite, searching for the vulva. (D) Having reached the tail of the hermaphrodite, the male executes a sharp turn and resumes tracing on the other side. When he reaches the nose of the hermaphrodite, he will execute another such turn. (E) The male reaches

the vulva and stops, begins probing for fine scale vulva location, then inserts his spicules and releases sperm.

A.



B.

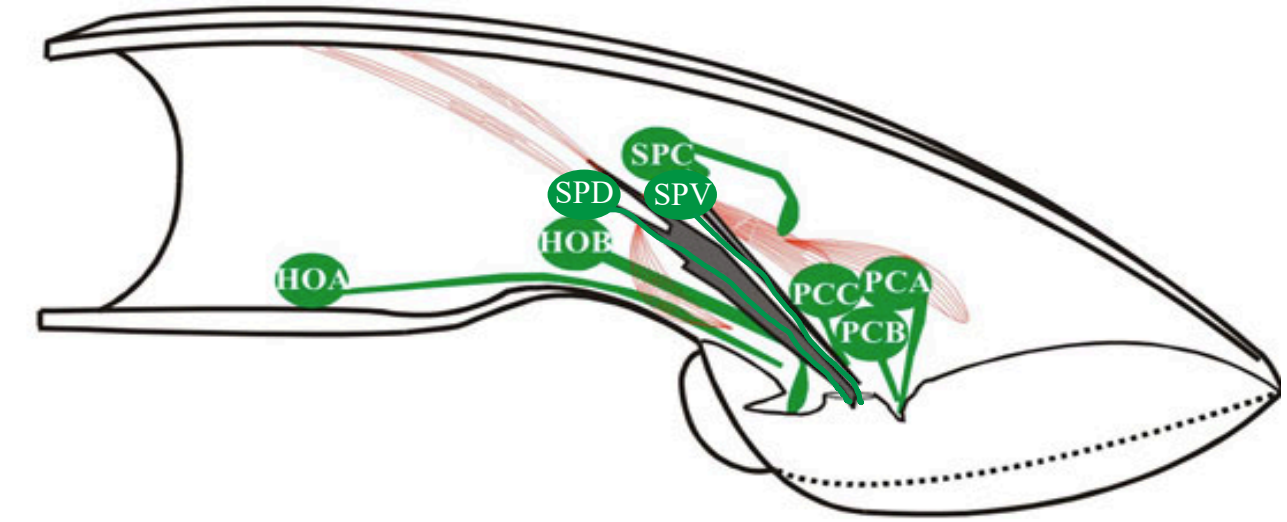


Figure 1.6 (A) The ray neurons of the male tail, used in turning and tracing. (B) Neurons used in stopping at the vulva (HOA, HOB), fine-scale vulva location (PCAL/R, PCBL/R,

PCCL/R), and spicule insertion and sperm release (SPDL/R, SPVL/R, SPCL/R). Images adapted from Wormbook. (99)

1.3 References

1. G. Voglis, N. Tavernarakis, *Mechanotransduction in the Nematode Caenorhabditiselegans* (Academia, 2005) (March 29, 2020).
2. I. R. Booth, P. Blount, The MscS and MscL families of mechanosensitive channels act as microbial emergency release valves. *J. Bacteriol.* **194**, 4802–4809 (2012).
3. M. Schwander, B. Kachar, U. Muller, The cell biology of hearing. *J. Cell Biol.* **190**, 9–20 (2010).
4. J. C. Tuthill, E. Azim, Proprioception. *Curr. Biol.* **28**, R194–R203 (2018).
5. A. I. Basbaum, D. M. Bautista, G. Scherrer, D. Julius, Cellular and Molecular Mechanisms of Pain. *Cell* **139**, 267–284 (2009).
6. E. A. Lumpkin, K. L. Marshall, A. M. Nelson, The cell biology of touch. *J. Cell Biol.* **191**, 237–248 (2010).
7. M. A. Wozniak, C. S. Chen, Mechanotransduction in development: A growing role for contractility. *Nat. Rev. Mol. Cell Biol.* **10**, 34–43 (2009).
8. S. I. Sukharev, B. Martinac, P. Blount, C. Kung, Functional Reconstitution as an Assay for Biochemical Isolation of Channel Proteins: Application to the Molecular Identification of a Bacterial Mechanosensitive Channel. *Methods A Companion to Methods Enzymol.* **6**, 51–59 (1994).
9. P. G. Gillespie, R. G. Walker, Molecular basis of mechanosensory transduction. *Nature* **413**, 194–202 (2001).
10. X. Qiu, U. Müller, Mechanically gated ion channels in mammalian hair cells.

Front. Cell. Neurosci. **12**, 1–10 (2018).

11. S. G. Brohawn, Z. Su, R. MacKinnon, Mechanosensitivity is mediated directly by the lipid membrane in TRAAK and TREK1 K+ channels. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3614–3619 (2014).

12. J. García-Añoveros, D. P. Corey, Mechanosensation: Touch at the molecular level. *Curr. Biol.* **6**, 541–543 (1996).

13. B. Martinac, M. Buechner, A. H. Delcour, J. Adler, C. Kung, Pressure-sensitive ion channel in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 2297–2301 (1987).

14. N. Levina, Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J.* **18**, 1730–1737 (1999).

15. S. I. Sukharev, P. Blount, B. Martinac, F. R. Blattner, C. Kung, A large-conductance mechanosensitive channel in E. coli encoded by mscL alone. *Nature* **368**, 265–268 (1994).

16. R. B. Bass, P. Strop, M. Barclay, D. C. Rees, Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel. *Science (80-.).* **298**, 1582–1587 (2002).

17. W. Wang, *et al.*, The structure of an open form of an E. coli mechanosensitive channel at 3.45 å resolution. *Science (80-.).* **321**, 1179–1183 (2008).

18. G. Chang, R. H. Spencer, A. T. Lee, M. T. Barclay, D. C. Rees, Structure of the MscL homolog from Mycobacterium tuberculosis: A gated mechanosensitive ion

channel. *Science (80-.).* **282**, 2220–2226 (1998).

19. E. Perozo, A. Kloda, D. M. Cortes, B. Martinac, Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nat. Struct. Biol.* **9**, 696–703 (2002).

20. E. Perozo, D. M. Cortes, P. Sompornpisut, A. Kloda, B. Martinac, Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature* **418**, 942–948 (2002).

21. I. Iscla, R. Wray, P. Blount, On the structure of the N-terminal domain of the MscL channel: Helical bundle or membrane interface. *Biophys. J.* **95**, 2283–2291 (2008).

22. B. Akitake, A. Anishkin, N. Liu, S. Sukharev, Straightening and sequential buckling of the pore-lining helices define the gating cycle of MscS. *Nat. Struct. Mol. Biol.* **14**, 1141–1149 (2007).

23. C. C. Häse, A. C. Le Dain, B. Martinac, Molecular dissection of the large mechanosensitive ion channel (MscL) of E. coli: Mutants with altered channel gating and pressure sensitivity. *J. Membr. Biol.* **157**, 17–25 (1997).

24. X. Ou, P. Blount, R. J. Hoffman, C. Kung, One face of a transmembrane helix is crucial in mechanosensitive channel gating. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 11471–11475 (1998).

25. C. C. Cruickshank, R. F. Minchin, A. C. Le Dain, B. Martinac, Estimation of the pore size of the large-conductance mechanosensitive ion channel of Escherichia coli. *Biophys. J.* **73**, 1925–1931 (1997).

26. K. Yoshimura, M. Sokabe, Mechanosensitivity of ion channels based on protein-lipid interactions. *J. R. Soc. Interface* **7**, S307 (2010).

27. C. Berrier, M. Besnard, B. Ajouz, A. Coulombe, A. Ghazi, Multiple Mechanosensitive Ion Channels from Escherichia coli, Activated at Different Thresholds of Applied Pressure. *J. Membr. Biol.* **151**, 175–187 (1996).

28. M. Huang, M. Chalfie, Gene interactions affecting mechanosensory transduction in Caenorhabditis elegans. *Nature* **367**, 467–470 (1994).

29. R. O'Hagan, M. Chalfie, M. B. Goodman, The MEC-4 DEG/ENaC channel of Caenorhabditis elegans touch receptor neurons transduces mechanical signals. *Nat. Neurosci.* **8**, 43–50 (2005).

30. Y. Chen, S. Bharill, E. Y. Isacoff, M. Chalfie, Subunit composition of a DEG/ENaC mechanosensory channel of Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11690–11695 (2015).

31. M. B. Goodman, *et al.*, MEC-2 regulates C. elegans DEG/ENaC channels needed for mechanosensation. *Nature* **415**, 1039–1042 (2002).

32. S. Zhang, *et al.*, MEC-2 is recruited to the putative mechanosensory complex in C. elegans touch receptor neurons through its stomatin-like domain. *Curr. Biol.* **14**, 1888–1896 (2004).

33. S. Shi, C. J. Luke, M. T. Miedel, G. A. Silverman, T. R. Kleyman, Activation of the caenorhabditis elegans degenerin channel by shear stress requires the MEC-10 subunit. *J. Biol. Chem.* **291**, 14012–14022 (2016).

34. D. S. Chelur, *et al.*, The mechanosensory protein MEC-6 is a subunit of the C.

elegans touch-cell degenerin channel. *Nature* **420**, 669–673 (2002).

35. A. L. Brown, Z. Liao, M. B. Goodman, MEC-2 and MEC-6 in the Caenorhabditis elegans sensory mechanotransduction complex: Auxiliary subunits that enable channel activity. *J. Gen. Physiol.* **131**, 605–616 (2008).

36. H. Du, G. Gu, C. M. William, M. Chalfie, Extracellular proteins needed for C. elegans mechanosensation. *Neuron* **16**, 183–194 (1996).

37. L. Emtage, G. Gu, E. Hartwieg, M. Chalfie, Extracellular proteins organize the mechanosensory channel complex in C. elegans touch receptor neurons. *Neuron* **44**, 795–807 (2004).

38. T. Fukushige, *et al.*, Mech-12, an α-tubulin required for touch sensitivity in C. elegans. *J. Cell Sci.* **112**, 395–403 (1999).

39. A. Bounoutas, M. Chalfie, Touch sensitivity in Caenorhabditis elegans. *Pflugers Arch. Eur. J. Physiol.* **454**, 691–702 (2007).

40. M. B. Goodman, P. Sengupta, How caenorhabditis elegans senses mechanical stress, temperature, and other physical stimuli. *Genetics* **212**, 25–51 (2019).

41. C. Montell, The TRP Superfamily of Cation Channels. *Sci. STKE* **2005**, re3 (2005).

42. B. Coste, *et al.*, Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science (80-.).* **330**, 55–60 (2010).

43. B. Coste, *et al.*, Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* **483**, 176–181 (2012).

44. P. A. Gottlieb, C. Bae, F. Sachs, Gating the mechanical channel Piezo1: A

comparison between whole-cell and patch recording. *Channels* **6**, 282–289 (2012).

45. S. N. Bagriantsev, E. O. Gracheva, P. G. Gallagher, Piezo proteins: Regulators of mechanosensation and other cellular processes. *J. Biol. Chem.* **289**, 31673–31681 (2014).

46. D. L. Prole, C. W. Taylor, Identification and Analysis of Putative Homologues of Mechanosensitive Channels in Pathogenic Protozoa. *PLoS One* **8**, 66068 (2013).

47. S. E. Kim, B. Coste, A. Chadha, B. Cook, A. Patapoutian, The role of Drosophila Piezo in mechanical nociception. *Nature* **483**, 209–212 (2012).

48. A. Faucherre, J. Nargeot, M. E. Mangoni, C. Jopling, piezo2b regulates vertebrate light touch response. *J. Neurosci.* **33**, 17089–17094 (2013).

49. A. Faucherre, K. Kissa, J. Nargeot, M. E. Mangoni, C. Jopling, Piezo1 plays a role in erythrocyte volume homeostasis. *Haematologica* **99**, 70–75 (2014).

50. S. S. Ranade, *et al.*, Piezo2 is the major transducer of mechanical forces for touch sensation in mice. *Nature* **516**, 121–125 (2014).

51. S. S. Ranade, *et al.*, Piezo1, a mechanically activated ion channel, is required for vascular development in mice. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10347–10352 (2014).

52. N. Eijkelkamp, *et al.*, A role for Piezo2 in EPAC1-dependent mechanical allodynia. *Nat. Commun.* **4** (2013).

53. A. Kamajaya, J. T. Kaiser, J. Lee, M. Reid, D. C. Rees, The structure of a conserved piezo channel domain reveals a topologically distinct β sandwich fold. *Structure* **22**, 1520–1527 (2014).

54. J. Ge, *et al.*, Architecture of the mammalian mechanosensitive Piezo1 channel. *Nature* **527**, 64–69 (2015).

55. K. Saotome, *et al.*, Structure of the mechanically activated ion channel Piezo1. *Nature* **554**, 481–486 (2018).

56. L. Wang, *et al.*, Structure and mechanogating of the mammalian tactile channel PIEZO2. *Nature* **573**, 225–229 (2019).

57. J. Wu, A. H. Lewis, J. Grandl, Touch, Tension, and Transduction – The Function and Regulation of Piezo Ion Channels. *Trends Biochem. Sci.* **42**, 57–71 (2017).

58. A. H. Lewis, J. Grandl, Mechanical sensitivity of Piezo1 ion channels can be tuned by cellular membrane tension. *Elife* **4** (2015).

59. C. D. Cox, *et al.*, Removal of the mechanoprotective influence of the cytoskeleton reveals PIEZO1 is gated by bilayer tension. *Nat. Commun.* **7** (2016).

60. J. Wu, R. Goyal, J. Grandl, Localized force application reveals mechanically sensitive domains of Piezo1. *Nat. Commun.* **7** (2016).

61. I. Borbiro, D. Badheka, T. Rohacs, Activation of TRPV1 channels inhibits mechanosensitive piezo channel activity by depleting membrane phosphoinositides. *Sci. Signal.* **8** (2015).

62. S. Brenner, The Genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).

63. J. Hodgkin, Karyotype, ploidy, and gene dosage. *WormBook*, 1–9 (2005).

64. J. Hodgkin, H. R. Horvitz, S. Brenner, Nondisjunction mutants of the nematode Caenorhabditis elegans. *Genetics* **91** (1979).

65. M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, D. C. Prasher, Green Fluorescent

Protein as a Marker for Gene Expression. *Science (80-.).* **263**, 802–805 (1994).

66. D. J. Dickinson, B. Goldstein, CRISPR-based methods for caenorhabditis elegans genome engineering. *Genetics* **202**, 885–901 (2016).

67. M. Kosicki, K. Tomberg, A. Bradley, Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36** (2018).

68. A. E. Friedland, Y. B. Tzur, M. P. Colaiacovo, G. M. Church, J. A. Calarco, Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat. Commun.* **10**, 741–743 (2013).

69. H. Wang, H. Park, J. Liu, P. W. Sternberg, An efficient genome editing strategy to generate putative null mutants in Caenorhabditis elegans using CRISPR/Cas9. *G3 Genes, Genomes, Genet.* **8**, 3607–3616 (2018).

70. S. Ward, N. Thomson, J. G. White, S. Brenner, Electron microscopical reconstruction of the anterior sensory anatomy of the nematodecaenorhabditis elegans. *J. Comp. Neurol.* **160**, 313–337 (1975).

71. J. E. Sulston, H. R. Horvitz, Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol.* **56**, 110–156 (1977).

72. T. A. Jarrell, *et al.*, The connectome of a decision-making neural network. *Science (80-.).* **337**, 437–444 (2012).

73. S. J. Cook, *et al.*, Whole-animal connectomes of both Caenorhabditis elegans sexes. *Nature* **571**, 63–71 (2019).

74. J. Hodgkin, Male Phenotypes and Mating Efficiency in Caenorhabditis Elegans.

Genetics **103**, 43–64 (1983).

75. A. M. Rose, D. L. Baillie, The Effect of Temperature and Parental Age on Recombination and Nondisjunction in CAENORHABDITIS ELEGANS. *Genetics* **92**, 409–18 (1979).

76. J. L. Anderson, L. T. Morran, P. C. Phillips, Outcrossing and the maintenance of males within C. elegans populations in *Journal of Heredity*, (2010), pp. 62–74.

77. J. R. Chasnov, K. L. Chow, Why are there males in the hermaphroditic species Caenorhabditis elegans? *Genetics* **160**, 983–994 (2002).

78. D. H. W. Leighton, A. Choe, S. Y. Wu, P. W. Sternberg, Communication between oocytes and somatic cells regulates volatile pheromone production in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17905–17910 (2014).

79. J. Kimble, D. Hirsh, The Postembryonic Cell Lineages of the Hermaphrodite and Male Gonads in Caenorhabditis elegans. *Dev. Biol.*, 396–417 (1979).

80. J. E. Sulston, D. G. Albertson, J. N. Thomson, The Caenorhabditis elegans male: Postembryonic development of nongonadal structures. *Dev. Biol.* **78**, 542–576 (1980).

81. M. M. Barr, L. R. García, D. S. Portman, Sexual dimorphism and sex differences in caenorhabditis elegans neuronal development and behavior. *Genetics* **208**, 909–935 (2018).

82. J. McCarter, B. Bartlett, T. Dang, T. Schedl, On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans. *Dev. Biol.* **205**, 111–128 (1999).

83. J. McCarter, B. Bartlett, T. Dang, T. Schedl, Soma-germ cell interactions in Caenorhabditis elegans: Multiple events of hermaphrodite germline development require the somatic sheath and spermathecal lineages. *Dev. Biol.* **181**, 121–143 (1997).

84. J. M. Simon, P. W. Sternberg, Evidence of a mate-finding cue in the hermaphrodite nematode Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 1598–1603 (2002).

85. K. S. Liu, P. W. Sternberg, Sensory regulation of male mating behavior in caenorhabditis elegans. *Neuron* **14**, 79–89 (1995).

86. R. Lints, L. Jia, K. Kim, C. Li, S. W. Emmons, Axial patterning of C. elegans male sensilla identities by selector genes. *Dev. Biol.* **269**, 137–151 (2004).

87. P. K. Koo, X. Bian, A. L. Sherlekar, M. R. Bunkers, R. Lints, The robustness of Caenorhabditis elegans male mating behavior depends on the distributed properties of ray sensory neurons and their output through core and male-specific targets. *J. Neurosci.* **31**, 7497–7510 (2011).

88. M. M. Barr, P. W. Sternberg, A polycystic kidney-disease gene homologue required for male mating behaviour in C. elegans. *Nature* **401**, 386–389 (1999).

89. H. Yu, R. F. Pretot, T. R. Burglin, P. W. Sternberg, Distinct roles of transcription factors EGL-46 and DAF-19 in specifying the functionality of a polycystin-expressing sensory neuron necessary for C. elegans male vulva location behavior. *Development* **130**, 5217–5227 (2003).

90. L. R. Garcia, P. Mehta, P. W. Sternberg, Regulation of distinct muscle behaviors

controls the C. elegans male's copulatory spicules during mating. *Cell* **107**, 777–788 (2001).

91. B. Leboeuf, P. Correa, C. Jee, L. R. García, Caenorhabditis elegans male sensory-motor neurons and dopaminergic support cells couple ejaculation and post-ejaculatory behaviors. *Elife* **2014** (2014).

92. Y. Liu, *et al.*, Caenorhabditis elegans Male Copulation. *PLoS Genet* **7**, 1001326 (2011).

93. Y. K. Bae, M. M. Barr, Sensory roles of neuronal cilia: Cilia development, morphogenesis, and function in C. elegans. *Front. Biosci.* **13**, 5959–5974 (2008).

94. M. M. Barr, *et al.*, The Caenorhabditis elegans autosomal dominant polycystic kidney disease gene homologs lov-1 and pkd-2 act in the same pathway. *Curr. Biol.* **11**, 1341–1346 (2001).

95. M. B. Goodman, E. M. Schwarz, Transducing Touch in Caenorhabditis elegans . *Annu. Rev. Physiol.* **65**, 429–452 (2003).

96. F. Sievers, *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

97. T. UniProt Consortium, UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **46**, 2699–2699 (2018).

98. D. Zarkower, Somatic sex determination. *WormBook*, 1–12 (2006).

99. M. M. Barr, L. R. Garcia, Male mating behavior, 10.1895/wormbook.1.78.1 (2005) https:/doi.org/10.1895/wormbook.1.78.1 (August 26, 2019).

C h a p t e r 2

PEZO-1 AND *TRP-4* AFFECT MATING IN *CAENORHABDITIS ELEGANS*

Katherine I Brugman, Vladislav Susoy, Aravinthan DT Samuel, Paul W Sternberg

2.1 Abstract

Piezo channels are a class of homotrimeric mechanosensitive ion channels with a highly conserved pore structure found across all phylogenetic clades except for yeast and bacteria. Piezo channels have been implicated in a number of functions in development and touch sensation in other organisms, though previously no function has been determined in *C. elegans*. Using expression patterns from a GFP transcriptional fusion of several non-coding regions with a 5' to *pezo-1*, we were able to orient our investigation to specific neurons in the male tail, notably the sensory rays, HOB, and the PCS neurons, which are known to be important in male mating. We used CRISPR/Cas9 genomic editing to generate a deletion of a large segment of the highly conserved pore region as well as to create a putative stop mutation in the same area. We then performed a series of mating assays to test the responsiveness and mating ability of the *pezo-1* mutant males. Males with a mutated *pezo-1* gene have increased mating time, exhibit an off-target probing phenotype reminiscent of a PCS ablation, lose contact with the hermaphrodite more often, and pass the vulva in a similar manner to a HOB ablation. In addition to providing some insight into the role of

mechanosensation in *C. elegans* mating, these phenotypes may provide a path to further studying critical aspects of the PEZO-1 protein structure-function *in vivo*.

2.2 Introduction

Caenorhabditis elegans is a bacterivorous self-fertilizing hermaphroditic nematode that possesses an alternative reproductive strategy based on a low incidence of males, approximately one male per thousand hermaphroditic animals (1–4). This mating strategy arguably represents the most complex behavior observed in the worm and can be subdivided into several discrete stages as first described by Jonathan Hodgkin. These stages may be roughly summarized as initial attraction, tracing, turning, stopping at the vulva, probing, and finally insertion of the spicule and release of sperm (5, 6). In more detail, each step comes with its own particular behavior. Tracing is the base state of mating behavior, and the one the male returns to whenever a subsequent stage fails, and is where the male contacts the hermaphrodite with the ventral part of his tail and backs along her cuticle. Turning occurs if he should reach her head or tail, this is the behavior he must execute to perform a sharp turn around these points and resume tracing on the other side. Stopping at the vulva is self-explanatory, when the male reaches the vulva, he stops. Then he commences probing, a behavior of rapidly pressing his spicules in the vulval region for fine-scale vulva location. When he has located it, he inserts the spicules and releases sperm. This is a complex behavior

and apart from the initial attraction, largely driven by mechanosensory inputs, both due to the fairly mechanical nature of mating and that key neurons involved by the behavior, such as the ray neurons and PCS neurons, have no external access that would be necessary for chemosensation (5, 7).

Many of these sensory neurons appear to be both chemical and mechanical sensors. For example, the hook neurons HOA and HOB have ciliated sensory endings but are nonetheless embedded in what looks morphologically like a mechanosensory structure (the hook). The ray neurons (R[1-9]A and R[1-9]B), which are also ciliated, are open to the exterior of the worm, save for the ray 6 neurons, and are embedded in elongated sensory rays. The post-cloacal sensilla are a pair of elongated sensory structures just posterior to the cloaca, and each have three ciliated sensory neurons (PCA, PCB and PCC) (7, 8). Ablation by laser microbeam irradiation of each of these structures results in a defect in particular aspects of mating (5). Specifically, ablation of dorsal opening rays results in males that fail to respond to dorsal contact, and ablation of ventral opening rays along with other ventral organs results in males that fail to observe appropriate backing behavior (5). Ablation of either the hook sensillum or either associated neuron (HOA and HOB) results in a male that is unable to stop at the vulva normally, instead backing slowly with spicule partially extruded until contacting

the vulva (5). Ablation of post cloacal sensilla neurons results in males that appear to have wild type mating ability except that they tend to "lose" the vulva at a greater rate than their wild type counterparts (5). Ablation of both the hook sensillum and post cloacal sensillum results in males who are unable to stop at the vulva or locate it through slow searching (5).

Some known or suspected mechanoreceptor-associated proteins are involved in mating. Specifically, *mec-4* gain-of-function, *mec-4* reduction-of-function, and *mec-10* loss-of-function mutants, which showed mutant animals with repetitive turning defects, suggesting the involvement of the MEC-4/MEC-10 complex in regulating the turning behavior in male mating (9). Since the phenotypes of these mutants doesn't account for all the ablations, we sought out examine other potential mechanoreceptors, namely the piezo mechanoreceptor in *C. elegans,* PEZO-1, and a transient receptor potential channel protein, TRP-4.

Piezo mechanoreceptors are nonselective cation channels that have been categorized in every phylogenetic clade other than yeast and bacteria (10–15). Vertebrates have two piezo proteins, while invertebrates have only one (10, 11). Structurally, they form a homotrimeric structure that features a central pore through which cations pass, three N-terminal propeller blades spiraling out from this pore, and a cap region that covers the pore (16–18). They have

been implicated in a number of functions among various model organisms, from nociception in *drosophila*, to blood cell volume regulation in mice, and erythrocyte volume homeostasis in zebrafish (11, 14, 19–22). In humans, deficiencies in piezo have been linked to dehydrated hereditary xerocytosis (23–26). *C. elegans*, like all other invertebrates, has a single piezo protein, PEZO-1, though no function has currently been linked to it in the animal (12).

Channels in the TRP superfamily share some core common features, namely being large tetramers, possessing six transmembrane domains, and having the ability to pass cations into the neuron (27, 28). Past that, they're divided into six subfamilies based on function and sequence identity (27). Like piezo channels, TRP channels are implicated in a number of sensory modalities, such as thermosensation, mechanosensation, taste, smell, and hearing (27, 29–34). In *C. elegans*, the TRP-4 channel, from the mechanosensative TRPN subfamily, has been implicated in locomotion defects, where the animals exhibit exaggerated body bending, likely due to a defect in proprioception, and the protein is expressed in the DVA neuron (29, 35). There has currently been no published evidence of mating defects linked to *trp-4.*

Here we construct mutants defective in *pezo-1* and *trp-4* and analyze them.

2.3 Methods

General methods

Worms were cultured at 20˚C as described by Brenner (6). Assays were also conducted at this temperature.

pezo-1::GFP Transcriptional Fusions

To visualize the cellular expression patterns of *pezo-1* in *c. elegans*, a transcriptional fusion plasmid of several theorized promoter regions and GFP were generated and injected into N2 hermaphrodites along with a *myo-2* dsred co-injection marker (Figure 2.1). Several suspected promoter regions were selected for this procedure, including the region 2 kb preceding the ATG site of the primary isoform as well as those of proposed alternate isoforms. Ultimately, a combined transcriptional fusion plasmid was generated from a 9 kb combination of the regions that exhibited the best expression patterns (TCCTACCTTGTGCCTGTCTACGTA…TCTGTTGTTGCTGTGTTACGACCA and GCACATGCTGCAGGTTAGTTTATT...AGGGTATCGCCTGAAAAAACAATG) in both *him-5* background with a *myo-2* dsred co-injection marker and in a temperature sensitive *pha-*

1 background with a *pha-1* rescue plasmid as a marker. These two strains, along with previous expression data, were imaged at 100x to identify the target cells to design subsequent experiments.

CRISPR-Cas9 deletion mutant

To generate the 2kb deletion mutant, *pezo-1 (sy1113)*, we injected a 46168 Crispr Cas9 plasmid (peft-3::cas9-sv40_NLS::ttb-2 UTR) along with pRB1017 plasmids (pU6::gRNA construct) containing our guide RNA sequences (CAGAAGCTCGTAAGCCAGG, AGGTCGAGGTCGTGAGCGG, CCACCACTTTACGAGATGG, ATAGGCAGCTTCGAACTGG) into young adult hermaphrodites (36, 37). In addition, as part of the aforementioned injection, we included a *dpy-10* guide RNA plasmid, also inserted into pRB1017, and a repair oligo with the *dpy-10* mutation as a marker which was later crossed out prior to use. Successful injections were identified by the presence of dumpy or roller progeny on the plate. 120 dumpy and roller progeny were singled onto individual plates and screened for the deletion mutation. Worm lysate was split between a control PCR with one internal and external primer (GGCGACCGTTGGTTGGGCTGGTTT, AATACGAGAGCCTTCACATCATC) and a

detection PCR with two primers external to the proposed deletion (GGCGACCGTTGGTTGGGCTGGTTT, ATCCTGTGTCCGATCCTGACG). A homozygous mutation line was determined by the absence of a band in the control PCR and the presence of a 500 bp band in the detection PCR. The strain was then crossed and outcrossed into *him-5* to remove the dumpy and roller marker mutations and verified by PCR with the above primers. The deletion was also sequenced and verified to be in frame and at the location, IV:9347253...9345244.

CRISPR-Cas9 STOP-IN mutant

The CRISPR-Cas9 STOP-IN mutant, *pezo-1 (sy1199),* was generated using purified protein Cas9 at 10 ug/uL concentration, a purified guide RNA near the mutation location (CCAGAAGCTCGTAAGCCAGG), and a single stranded DNA repair oligo containing three stop codons, one in every reading frame (cttatcgctgtttctgaaccagaagctcgtaagccGGGAAGTTTGTCCAGAGCAGAGGTGACTAAGTG ATAAgctagcaggaggcactgaagaaacggatggtgatgaag) (38). These reagents were injected into N2 young adults along with a *dpy-10* guide and repair oligo. Successful injections were identified by the presence of dumpy and roller progeny. 30 roller progeny were singled out

from “jackpot” plates (plates with a high incidence of dumpy and roller progeny) and screened via PCR (GACAGGACTTTCCCGCCAACTTAA, ATCATTCGCCGATTGCACAAGTTG) and an internal primer (GCTTATCACTTAGTCACCTCTGCTC).

Other strains

trp-4 (sy695); him-5 (e1490)

OH10235 bxIs19 [*trp-4*p::GFP]

ZM9624 (Prgef-1 GCaMP6::3xNLS::mNeptune)

BB92 uuEx18 [*dcr-1*(wild-type) + *dpy-30*::mCherry]

ADS1014 otIs377 [*myo-3*p::mCherry]; *unc-64*(e246) III

Imaging strains

For panneuronal imaging, we used ZM9624 and PS8041 males. ZM9624 was created and generously shared by Mei Zhen’s laboratory (Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Canada). The strain was designed to co-express calcium indicator GCaMP6s and red fluorescent protein mNeptune in all neuronal nuclei under the rgef-

1 promoter. PS8041 was generated by crossing ZM9624 into *him-5* and *pezo-1* mutant backgrounds. To track mating events, we used hermaphrodites that expressed red fluorescent markers in their cuticle or muscles. To encourage the richness of male behavior we used partners with normal motility or showing mild unc phenotypes.

Recording of neuronal activity in freely moving males

To record activity of neurons in freely-moving males, we used a modified version of a custom-built spinning disc confocal microscopy setup described in Venkatachalam et al. (39). Virgin L4 males were picked onto a separate plate with OP50 and were kept on the plate for 15 to 20 hours before imaging. For the imaging experiments, a single virgin male was placed onto a 10 cm NGM agar plate containing a small amount of OP50 for food, 10 ul of NGM buffer, and 3 to 5 fluorescently labeled hermaphrodites. The plate was covered with a No1 coverslip which allowed the use of an immersion oil objective while also letting the animals to navigate freely. The volumetric imaging was performed with a 40x 0.95 NA oil objective (Nikon Plan Apo Lambda series) controlled by a piezoelectric stage. We imaged 10 brain volumes per second, each volume consisting of 20 optical slices approximately 1.75 micrometers apart. The emitted light from the objective was split into green and red channels

and imaged with two separate Andor Zyla 4.2 sCMOS cameras at 200 Hz. Each camera recorded a 256x512 pixel region with 0.45 um pixel size. The posterior nervous system of the male was tracked continuously throughout the imaging session. This was achieved by adjusting the microscope stage position in x, y, and z. Imaging experiments lasted from 1.5 to 10 minutes and specifically aimed at capturing scanning and vulva detection events.

Extraction of neuronal activity signals

Neuronal activity traces were extracted from the raw data following image preprocessing and registration. Registered data were downsampled to 5 volumes per second. Neuronal segmentation and tracking were performed using manual and semi-automated methods with the help of Fiji plugin MaMuT 0.27 (40, 41). Fluorescent signals from green and red channels were extracted for 2.25x2.25x3.5 um regions of interest centered on neuronal nuclei. Savitzky-Golay filtering with polynomial order of 1 and frame length 13 was used for noise reduction (42). To minimize motion artifacts, a ratiometric approach was used for calculating neuronal activity traces.

Activity of eight neurons implicated in vulva detection was extracted (5, 43)(Susoy et al, unpublished). These neurons included post-cloacal sensilla neurons PCB, PCC, PCA, the hook neurons HOA, HOB, and the spicule neuron SPC, the ray neuron R2B and the interneuron PVX. From the recordings, we also manually tabulated discrete behavioral motifs expressed by the male. Here we specifically focused on scanning, vulva detection, and searching (no ventral contact).

Mating assay

Fresh, hand-poured NG plates were spotted with 13 uL OP50, and L4 males, both controls and test males, were selected and placed on separate plates one day before the assay. At least a half an hour prior to the assay, three males were placed on each mating plate and left to acclimate. On the first control plate, a number of *unc-31* hermaphrodites were placed with the males, and the first three to be mated were separated and transferred to subsequent mating plates. After hermaphrodites were transferred to a mating plate, the plate was then video recorded for either fifteen minutes in total or up to five minutes following the initial mating attempt by any of the three males present. Video recording was stopped on successful transfer of sperm, at the end of the fifteen minute period in which no males attempted a

mating, or after a male had been unsuccessful in completing a mating for five minutes. The videos were then analyzed for mating defects.

2.4 Results

Transcriptional GFP reporters reveal the pezo-1 and trp-4 expression pattern

Analysis of the transcriptional fusion constructs from several proposed promoter regions for the *pezo-1* gene showed different expression patterns with some overlap (Figure 2.1, Figure 2.2). Promoter number 2 in particular was of interest because, despite sharing significant overlap with promoter 3, it lacked a significant portion of the expression. A difference of approximately 100 bp seems to have determined a significant amount of the expression in the male tail neurons. A fusion of the region 4kb upstream of the gene and the largest intron contained all the variant expression patterns seen in the other transcriptional fusions and was used as a basis for designing further experiments.

Expression was noted in several head neurons, pharynx-intestinal valve, vulval muscle, spermatheca, CAN neuron, and a number of male tail neurons. For the purpose of this paper, we will be focusing on the expression found in the PCS neurons, HOB, and the sensory ray neurons (Figure 2.2). Due to the key role these neurons play in mating behavior, *pezo-1* expression in these neurons implied a link between piezo and this behavior (5).

An existing *trp-4* transcriptional fusion, OH10235 bxIs19 [trp-4p::GFP], was also examined, and expression was noted in HOB and several ray neurons (Figure 2.3). In addition, my collaborator reported PCA expression, but I was unable to confirm this expression.

Pezo-1 and Trp-4 mutant males display multiple mating phenotypes

We tested two *pezo-1* mutants and one *trp-4* mutant along with a *pezo-1; trp-4* double mutant strain and observed a defect in mating in a subset of the males tested. This defect was first seen in approximately 30% of the *pezo-1* males taking a significantly longer time to complete a mating event than the wild type males. To properly analyze this phenotype, males were binned into those who were able to insert their spicules and release sperm within five minutes, and those who took longer than five minutes to do so or never succeeded in completing a copulation. This showed that while wild type mating time does seem to contain some environmental factor, the mutant males still succeeded far less in attaining this time—about 30% less often than the wild type controls—and the double mutant attained this time only half as often as the wild type controls (Figure 2.4).

To understand why the mutants took longer to mate or failed to mate on these occasions, the video recordings were analyzed to uncover more specific defects. Three main defects were highlighted as a result of this analysis, consisting of off-target probing behavior, passing the vulva, and loss of contact (Figure 2.5).

The off-target probing behavior was defined as when the male, instead of probing at the vulva, would probe at a location not touching any part of the vulva. The time that the male spent probing at inappropriate locations was measured as a percent of the total time the male was recorded mating so as to give a more standardized view between various males who may have spent different amounts of time on a mating attempt. *pezo-1* and *trp-4* males were found to linger far longer and more frequently at off-target locations than wild type males. However, the *pezo-1; trp-4* double mutant had a reduced incidence of this defect.

The behavior of passing the vulva was defined as when the male, while tracing the hermaphrodite, would simply pass over the vulva instead of stopping at that location. This behavior has been previously identified in *lov-1* animals and, while the phenotype in *pezo-1* and *trp-4* males was less severe, it was still in excess of wild type males, which would, at most, pass the vulva once or twice. The *pezo-1; trp-4* double mutant had no increased incidence of this phenotype compared with the single mutants.

The loss of contact behavior was defined as when the male would lose contact with the hermaphrodite, but not leave entirely. This contact loss was usually seen as sloppy turning attempts, though it was also observed during standard tracing. *Pezo-1* and *trp-4* mutants were seen to lose contact with their hermaphrodites far more often than wild type males. The *pezo-1; trp-4* double mutant had no increased incidence of this phenotype compared with the single mutants. An experiment with the *pezo-1 (sy1113)* males was partially repeated (interrupted due to pandemic) and seemed to have a much lower incidence of mutant phenotypes (data in Appendix – A).

Pan-Neuronal GCamP imaging of pezo-1 mutants shows aberrant activity in PCB and PCC compared to wild-type but not PCA

Pezo-1 mutant animals imaged using a pan-neuronal GCamP construct showed aberrant activity in the neurons PCB and PCC compared to wild-type animals. In wild-type animals, these two neurons activate when they contact the vulva region. In *pezo-1* animals, they activated at other regions along the worm as the male scanned the hermaphrodite (Figure 2.6).

2.5 Discussion

We see a striking correlation between *pezo-1* expression and mutant phenotype: We have shown that the *pezo-1* expression pattern of male-specific neurons in the male tail corresponds to a mating defect in *pezo-1* mutant males. This defect has been more specifically categorized into three primary sub-defects, off-target probing, passing the vulva during tracing, and loss of contact, which all contribute to the inability to complete a copulation within the wild type time.

There is an equally prominent *trp-4* function in mating that correlates to trp-4 expression patterns: We have shown that *trp-4* mutants exhibit a high rate of turning, passing vulva, and presenting off-target defects similar to the ones exhibited by *pezo-1* mutants. These defects correlate to *trp*-4 expression in the ray neurons, HOB, and possibly PCA (expression unclear), which as discussed earlier, have been linked to these defects in previous ablation studies.

The *trp-4; pezo-1* double mutant does not exhibit an increase in turning and passing the vulva mating defects compared to the single mutants: We have shown that the *trp-4; pezo-1* double mutant does not exhibit an exacerbated turning or passing the vulva phenotype compared to the single mutants of either gene. This suggests that these two genes operate within the same pathway and in such a manner that knocking out either disrupts the pathway to the same degree. There is almost certainly some third factor involved in this process, either a chemical signal or another mechanoreceptor (we have not yet found a mechanoreceptor linked to HOA, for example), that is capable of maintaining mating function despite these defects.

The *trp-4; pezo-1* double mutant has a reduced off-target probing phenotype: We have shown that the *trp-4; pezo-1* double mutant has a reduced off-target phenotype compared to either of the single mutant phenotypes. At first, this might seem to be counter-intuitive until we take into account the expression pattern data and the pan-neuronal GCamP data.

The function of *pezo-1* and *trp-4* in mating: Why are *pezo-1* and *trp-4* mutants incompletely penetrant? This lingering question of why both these mutants only seem to run

into this defect a portion of the time instead of their mating ability being constantly affected has some intriguing possibilities. While this could be an example of incomplete penetrance, as seen in *lov-1* mutants, it's also feasible that this could be the result of some regulatory mechanism being affected. The *pezo-1* mutants were generally able to accomplish their mating in a wild type manner when a mating attempt was flawless, that is to say, the hermaphrodite moved very little in response to the male's approach, the male either started tracing on the same side as the hermaphrodite's vulva or executed his first turn around her nose (generally an easier turn for the male), and the male stopped directly at the vulva. The moment a mistake was made, such as passing the vulva, or beginning to probe at the wrong location, or making the first turn at the tail, mutant males did not easily recover into a successful attempt and continued to make similar or worse errors. When wild type males made similar errors, they were nearly always able to recover quickly and mate successfully. The *trp-4* mutant followed in a similar pattern, though with a more regular incidence of sloppy or incomplete turns.

This can be seen in the data on the incidences of mating defects per worm (Figure 2.4). The *pezo-1* worms tended to have cascading defects that resulted in unsuccessful or prolonged mating while the wild type worms rarely would have more than one error. This

implies that one error increases the likelihood of further errors in mutant worms, while wild type worms will simply adjust to the mistake and continue on. This follows into the data of individual defects, where wild type defects are much less severe than mutant ones. Many wild type worms, for example, might pass the vulva once or twice in the course of a mating attempt, but once a *pezo-1* mutant male missed the vulva, he would often continue this pattern, seemingly unable to break the cycle. A *pezo-1* mutant male that probed at the wrong location, would often come back to the same incorrect location repeatedly, even if he passed over the actual vulva in the process. A male that failed to complete a turn might make the attempt multiple times, falling off and restarting where a wild type male would have adjusted and taken a different approach.

The other persisting questions lie in why the *pezo-1* and *trp-4* double mutant appears to exhibit no increase in severity of phenotype for passing the vulva and loss of contact defects, yet it shows a reduction in the severity for the off-target probing defect. This puzzle can somewhat be solved by the expression data observed in both *pezo-1* and *trp-4*. Both genes are expressed in ray neurons (likely connected to the loss of contact defects) and HOB (connected to passing the vulva) and, as part of the same pathway, additional defects do not exacerbate the function. The fact that neither mechanosensor is present in HOA and that there

are multiple missing ray neurons suggests that there are other elements at play in this process, such as additional mechanosensors or chemosensors.

The reduction in the severity of the off-target probing defect can be traced to the neurons that *pezo-1* and *trp-4* are expressed in, namely the PCB/C neurons for *pezo-1* and the PCA neurons for *trp-4*. From the pan-neuronal GCamP experiments, we can see more specifically that these two sets of neurons detect different aspects of fine-scale vulva location. The PCB/C neurons activate when the male is at the vulva, as shown in Figure 2.5, while the PCA neurons activate when the male is in contact with the hermaphrodite at any location other than the vulva. We propose that a defect in mechanosensation in either set causes an off-target probing effect as the male receives misleading cues as to the appropriate location of the vulva, while a defect in both sets of neurons has the net effect of the males probing less altogether, leading to a muted off-target phenotype. This can be supported by the fact that the *pezo-1; trp-4* double mutants also possess a far lower incidence of completed mating compared to either of their single mutant peers.

Implications for piezo biology: Currently, there have been no reports of piezo's involvement in mating behavior in any other species, so the discovery that it has a role in *C.*

elegans mating is particularly significant. While this function could be unique to the *C. elegans* piezo, it is possible that it may play some role in the mechanosensory aspects of mating behavior of other species and represent the missing mechanosensory channel in mapping these behaviors.

Acknowledgements.

We thank WormBase for genome information, CGC for providing strains, Male Brain grant and NIH Training Grant for funding, and Tsui-Fen Chou for providing Cas9 protein.

2.6 Figures

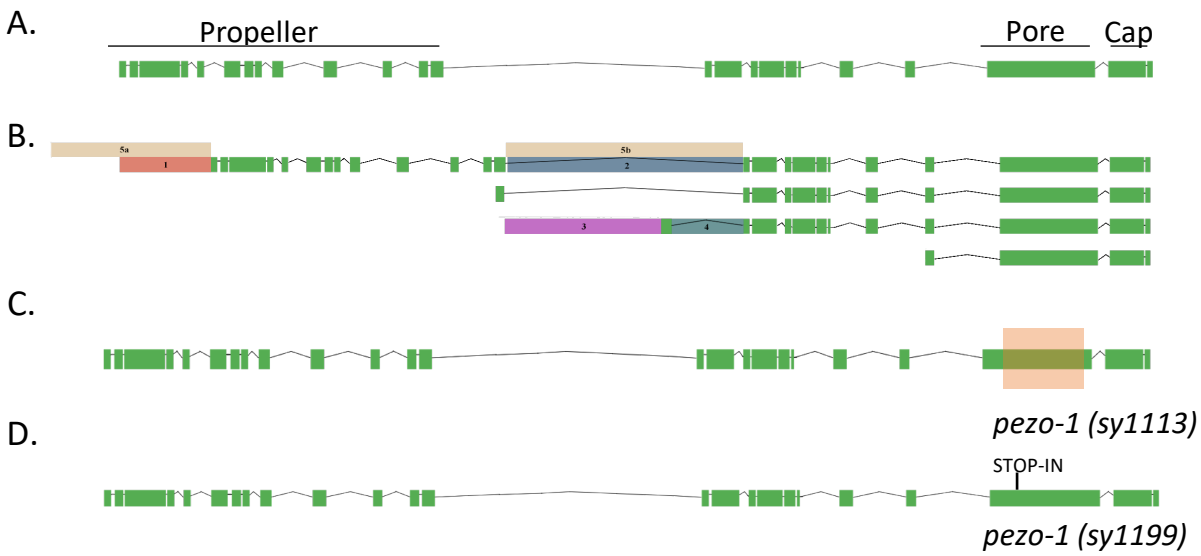


Figure 2.1 (A) Rough gene layout for *pezo-1* in *C. elegans*. (B) Non-coding regions used for the purposes of designing transcriptional fusion markers. Regions marked 5a and 5b were combined together to create strain *pha-1; him-5; syEx1771*[pBX-1 + *pezo-1*::GFP], region marked 3 was used to create strain *pha-1; him-5; syEx1772*[pBX-1 + *pezo-1*::GFP]. Strains created from regions marked 1, 2, and 4 are non-extant. (C) Crispr/Cas9 deletion mutation region. (D) Crispr/Cas9 STOP-IN location.

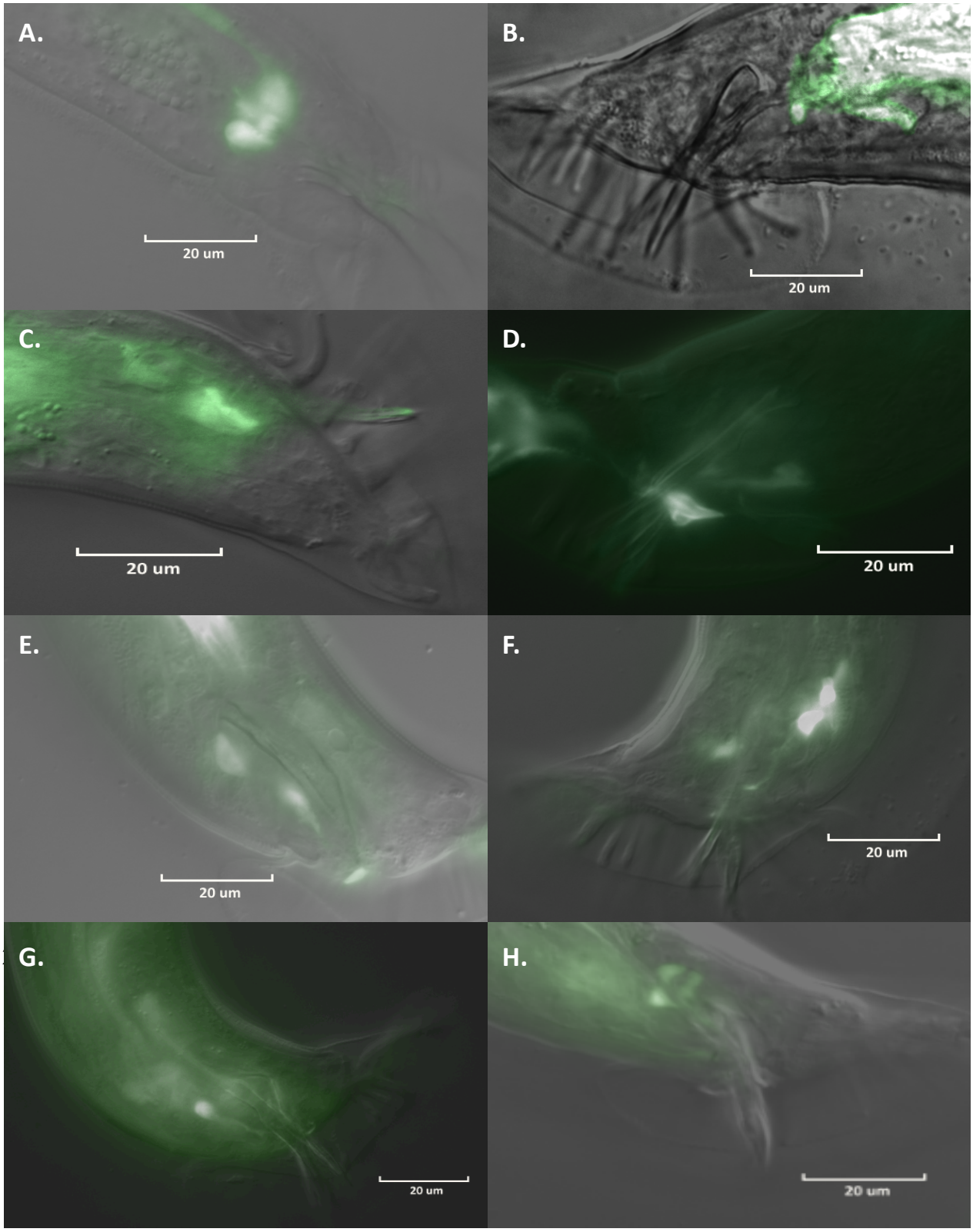

A.
20 um
B.
20 um
C.
20 um
D.
20 um
E.
20 um
F.
20 um
G.
20 um
H.
20 um

Figure 2.2 (A) GFP expression pattern of transcriptional fusion from region 1 of Figure 1.1, an area capturing 2kb upstream of the *pezo-1* gene. Only some intestinal glands are visible. (B) GFP expression pattern of transcriptional fusion from region 2 of Figure 1.1, capturing most of the central intron, minus ~100bp. Only intestinal expression is visible. (C, D) GFP expression pattern of transcriptional fusion from region 3 of Figure 1.1, capturing the front half of the central intron. The PCS neurons and HOB are visible, respectively. (E, F) GFP expression pattern of transcriptional fusion from region 4 of Figure 1.1, capturing the back half of the central intron. An interneuron and the PCS are visible, respectively. (G, H) GFP expression pattern of transcriptional fusion from regions 5a and 5b of Figure 1.1 fused together, capturing a non-coding region 4kb in front of the gene and the entire central intron. HOB, SPD, and SPC are visible.

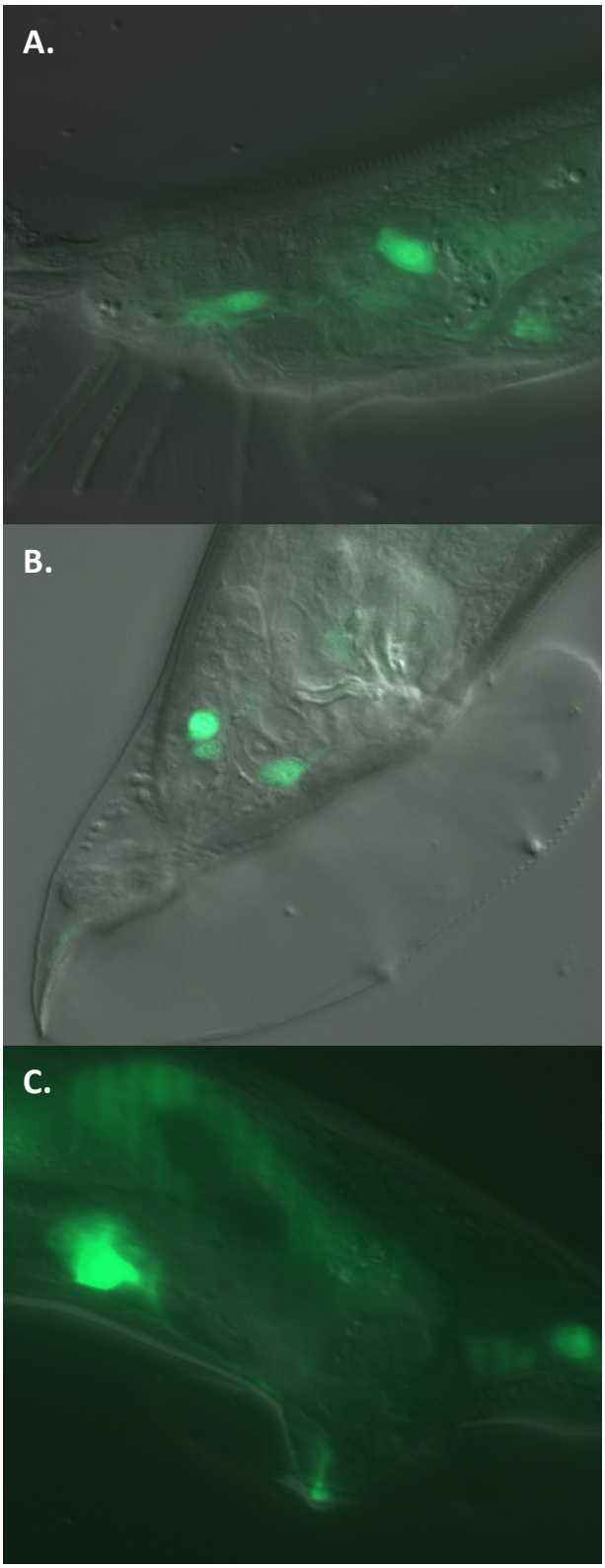

A.
B.
C.

Figure 2.3 GFP expression pattern for *trp-4*, from strain OH10235 bxIs19 [*trp-4*p::GFP]. (A) Visible are several processes into the ray neurons and an interneuron. (B) Visible here are more ray neurons and faint traces near where one would expect the PCS neurons. (C) Here, HOB is clearly visible, with process.

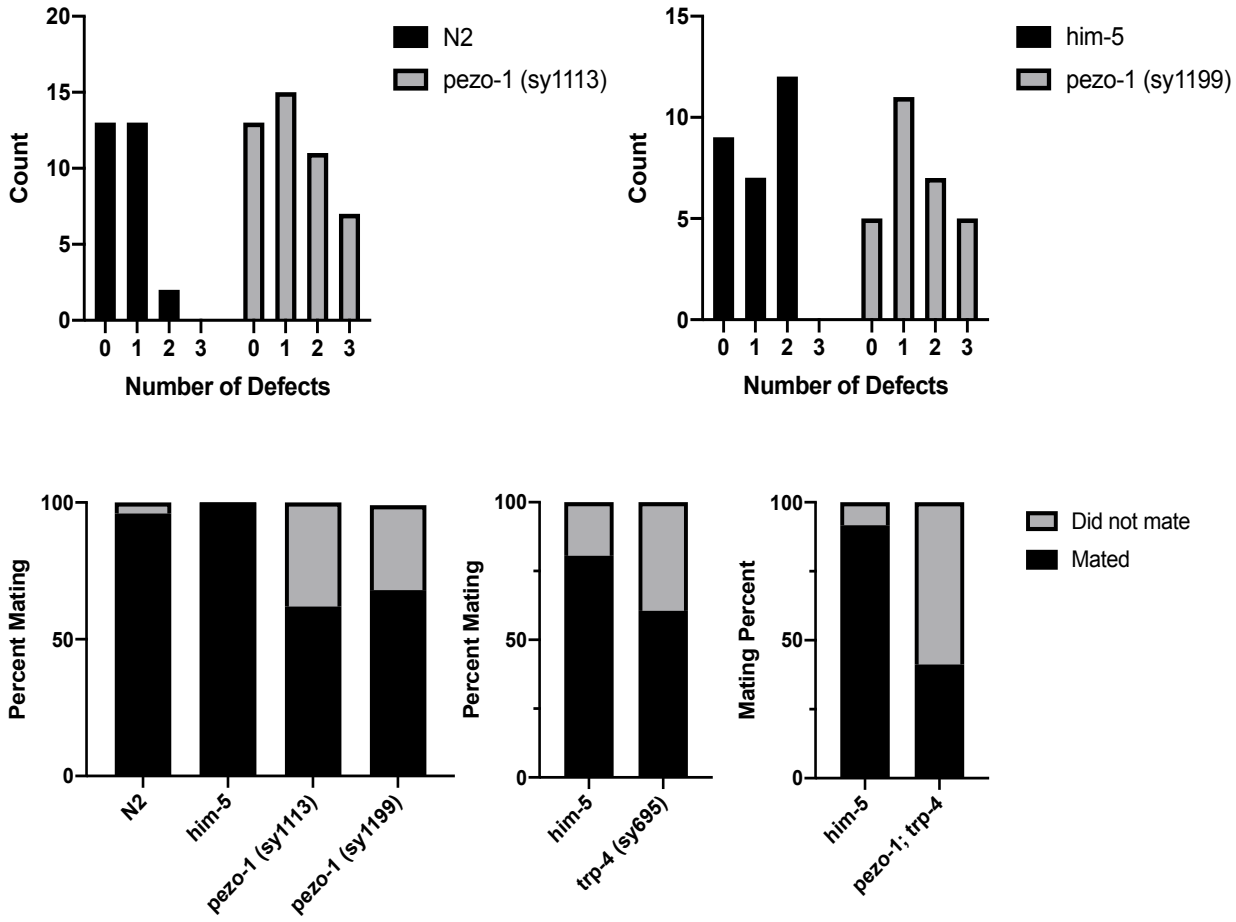


Figure 2.4 (A) Here we show the number of phenotypic defects per worm. While there is some variance between N2 and *him-5* animals, with *him-5* showing notably more defects (possibly due to not requiring mating to propagate), neither control animal group regularly possessed as many as the *pezo-1* mutants, who had significant numbers of animals who presented with two and three phenotypic defects. (B) Here we show the percent of animals that mated within 300 seconds, a time selected because most N2 and *him-5* animals were able to complete a mating in this period of time (the middle trial nonwithstanding). Completion of

a mating event here was scored as inserting of spicules. Roughly a third of *pezo-1* mutants were unable to complete a mating event. When the *trp-4* data is weighted vs its control, a similar number of *trp-4* mutant animals also failed. But the *pezo-1 (sy1199); trp-4 (sy695)* double mutant (abbreviated in the figure) demonstrated an enormous drop in the ability to mate quickly, with only 41% of males able to mate in under 300 seconds.

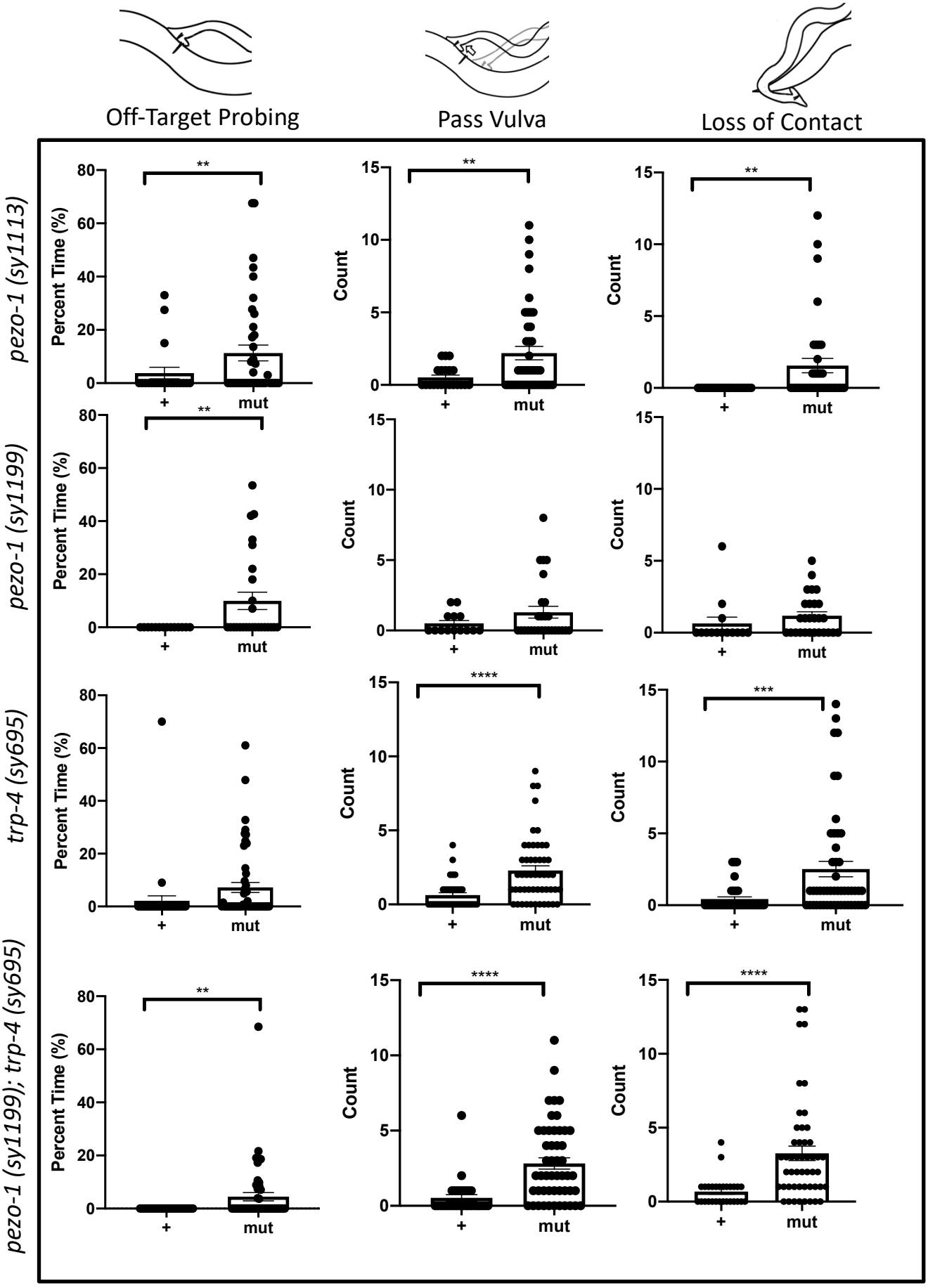

Off-Target Probing
Pass Vulva
Loss of Contact
pezo-1 (sy1113)
pezo-1 (sy1199)
trp-4 (sy695)
pezo-1 (sy1199); trp-4 (sy695)
Percent Time (%)
Count
+
mut
**
****
***

Figure 2.5 The *pezo-1* deletion mutant, *pezo-1 (sy1113)* and the *pezo-1* STOP-IN mutant, *pezo-1 (sy1199)* exhibit similar phenotypes to different degrees as the *trp-4* deletion mutant, *trp-4 (sy695)*. However, when a double mutation containing both the *pezo-1* STOP-IN and the *trp-4* deletion was generated, off-target probing decreased significantly, while passing the vulva and loss of contact phenotypes remained unchanged from the stronger *trp-4* deletion phenotype. All controls in these experiments were *him-5* except for the controls for *pezo-1 (sy1113)*, which were N2 males. All mutant strains were additionally crossed with *him-5*.

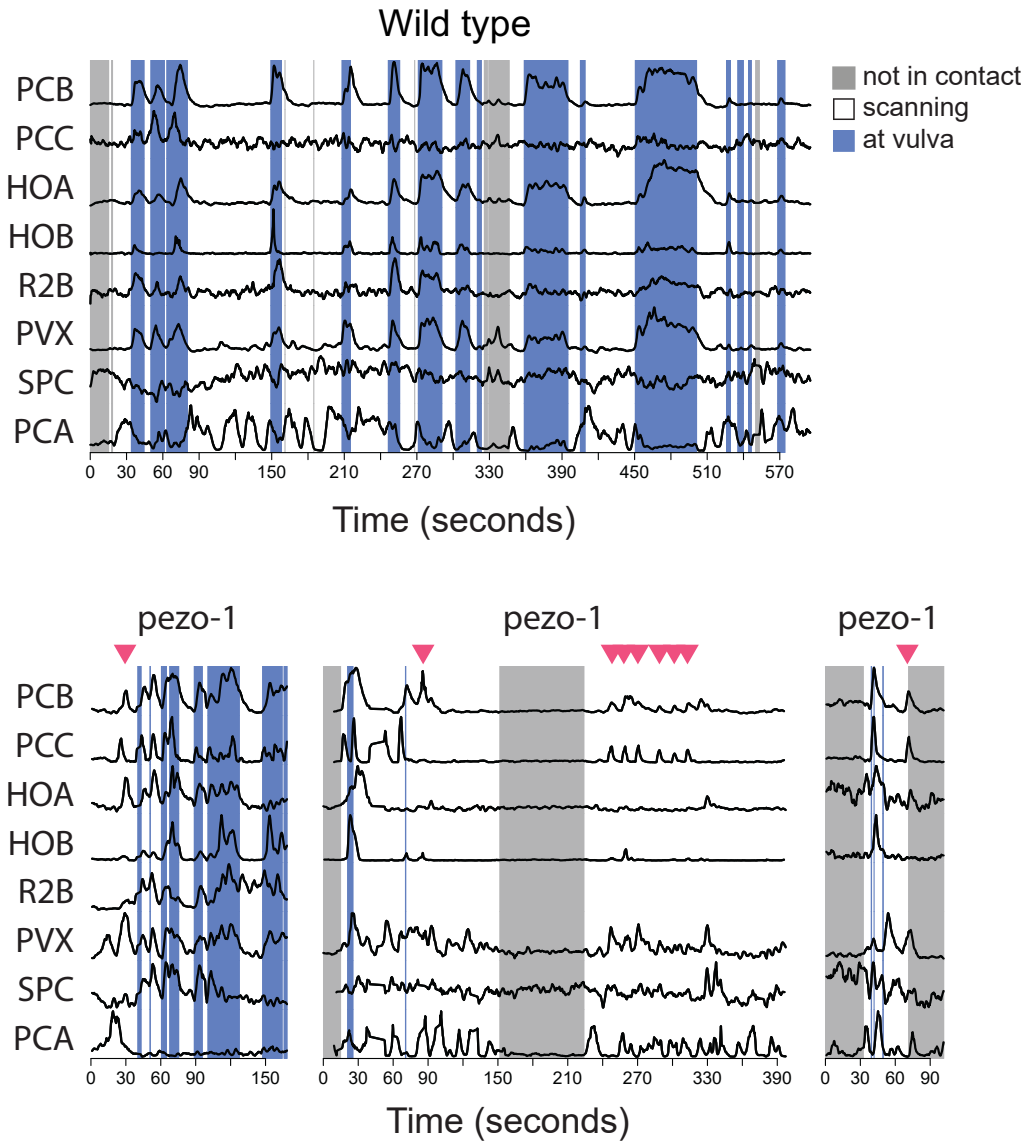


Figure 2.6 Full-brain GCamP imaging data of active mating events. Here, we can see that in the wild type, PCB and PCC neurons activate when the male passes over the vulva, while PCA activates when the male is not at the vulva. In the mutant animals below (penetrance at 21% of tested animals), PCB and PCC also activate when the male is not at the vulva.

2.7 References

1. J. Hodgkin, Male Phenotypes and Mating Efficiency in Caenorhabditis Elegans. *Genetics* **103**, 43–64 (1983).

2. J. Hodgkin, H. R. Horvitz, S. Brenner, Nondisjunction mutants of the nematode Caenorhabditis elegans. *Genetics* **91** (1979).

3. J. Kimble, D. Hirsh, The Postembryonic Cell Lineages of the Hermaphrodite and Male Gonads in Caenorhabditis elegans. *Dev. Biol.*, 396–417 (1979).

4. J. L. Anderson, L. T. Morran, P. C. Phillips, Outcrossing and the maintenance of males within C. elegans populations in *Journal of Heredity*, (2010), pp. 62–74.

5. K. S. Liu, P. W. Sternberg, Sensory regulation of male mating behavior in caenorhabditis elegans. *Neuron* **14**, 79–89 (1995).

6. S. Brenner, The Genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).

7. J. E. Sulston, D. G. Albertson, J. N. Thomson, The Caenorhabditis elegans male: Postembryonic development of nongonadal structures. *Dev. Biol.* **78**, 542–576 (1980).

8. Y. K. Bae, M. M. Barr, Sensory roles of neuronal cilia: Cilia development, morphogenesis, and function in C. elegans. *Front. Biosci.* **13**, 5959–5974 (2008).

9. T. Liu, K. Kim, C. Li, M. M. Barr, FMRFamide-Like Neuropeptides and Mechanosensory Touch Receptor Neurons Regulate Male Sexual Turning Behavior in Caenorhabditis elegans (2007) https:/doi.org/10.1523/JNEUROSCI.1405-07.2007 (August 26, 2019).

10. B. Coste, *et al.*, Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science (80-.).* **330**, 55–60 (2010).

11. S. E. Kim, B. Coste, A. Chadha, B. Cook, A. Patapoutian, The role of Drosophila Piezo in mechanical nociception. *Nature* **483**, 209–212 (2012).

12. A. Kamajaya, J. T. Kaiser, J. Lee, M. Reid, D. C. Rees, The structure of a conserved piezo channel domain reveals a topologically distinct β sandwich fold. *Structure* **22**, 1520–1527 (2014).

13. D. L. Prole, C. W. Taylor, Identification and Analysis of Putative Homologues of Mechanosensitive Channels in Pathogenic Protozoa. *PLoS One* **8**, 66068 (2013).

14. A. Faucherre, J. Nargeot, M. E. Mangoni, C. Jopling, piezo2b regulates vertebrate light touch response. *J. Neurosci.* **33**, 17089–17094 (2013).

15. Z. Zhang, *et al.*, Genetic analysis of a Piezo-like protein suppressing systemic movement of plant viruses in Arabidopsis thaliana. *Sci. Rep.* **9** (2019).

16. J. Ge, *et al.*, Architecture of the mammalian mechanosensitive Piezo1 channel. *Nature* **527**, 64–69 (2015).

17. K. Saotome, *et al.*, Structure of the mechanically activated ion channel Piezo1. *Nature* **554**, 481–486 (2018).

18. L. Wang, *et al.*, Structure and mechanogating of the mammalian tactile channel PIEZO2. *Nature* **573**, 225–229 (2019).

19. A. Faucherre, K. Kissa, J. Nargeot, M. E. Mangoni, C. Jopling, Piezo1 plays a role in erythrocyte volume homeostasis. *Haematologica* **99**, 70–75 (2014).

20. S. S. Ranade, *et al.*, Piezo1, a mechanically activated ion channel, is required for vascular development in mice. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10347–10352 (2014).

21. S. S. Ranade, *et al.*, Piezo2 is the major transducer of mechanical forces for touch sensation in mice. *Nature* **516**, 121–125 (2014).

22. N. Eijkelkamp, *et al.*, A role for Piezo2 in EPAC1-dependent mechanical allodynia. *Nat. Commun.* **4** (2013).

23. C. Bae, R. Gnanasambandam, C. Nicolai, F. Sachs, P. A. Gottlieb, Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1.

Proc. Natl. Acad. Sci. U. S. A. **110** (2013).

24. C. Bae, P. A. Gottlieb, F. Sachs, Human PIEZO1: Removing inactivation. *Biophys. J.* **105**, 880–886 (2013).

25. J. Albuisson, *et al.*, Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels. *Nat. Commun.* **4**, 1884 (2013).

26. R. Zarychanski, *et al.*, Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis. *Blood* **120**, 1908–1915 (2012).

27. C. Montell, The TRP Superfamily of Cation Channels. *Sci. STKE* **2005**, re3 (2005).

28. R. Gaudet, TRP channels entering the structural era. *J. Physiol.* **586**, 3565–3575 (2008).

29. L. Kang, J. Gao, W. R. Schafer, Z. Xie, X. Z. S. Xu, C. elegans TRP Family Protein TRP-4 Is a Pore-Forming Subunit of a Native Mechanotransduction Channel. *Neuron* **67**, 381–391 (2010).

30. M. J. Caterina, *et al.*, The capsaicin receptor: A heat-activated ion channel in the pain pathway. *Nature* **389**, 816–824 (1997).

31. D. D. McKemy, W. M. Neuhausser, D. Julius, Identification of a cold receptor reveals

a general role for TRP channels in thermosensation. *Nature* **416**, 52–58 (2002).

32. G. M. Story, *et al.*, ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. *Cell* **112**, 819–829 (2003).

33. D. P. Corey, *et al.*, TRPA1 is a candidate for the mechanosensitive transduction channel of vertebrate hair cells. *Nature* **432**, 723–730 (2004).

34. E. R. Liman, D. P. Corey, C. Dulac, TRP2: A candidate transduction channel for mammalian pheromone sensory signaling. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5791–5796 (1999).

35. W. Li, Z. Feng, P. W. Sternberg, X. Z. S. Xu, A C. elegans stretch receptor neuron revealed by a mechanosensitive TRP channel homologue. *Nature* **440**, 684–687 (2006).

36. A. E. Friedland, Y. B. Tzur, M. P. Colaiacovo, G. M. Church, J. A. Calarco, Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat. Commun.* **10**, 741–743 (2013).

37. J. A. Arribere, *et al.*, Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in caenorhabditis elegans. *Genetics* **198**, 837–846 (2014).

38. H. Wang, H. Park, J. Liu, P. W. Sternberg, An efficient genome editing strategy to

generate putative null mutants in Caenorhabditis elegans using CRISPR/Cas9. *G3 Genes, Genomes, Genet.* **8**, 3607–3616 (2018).

39. V. Venkatachalam, *et al.*, Pan-neuronal imaging in roaming Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1082–E1088 (2016).

40. C. Wolff, *et al.*, Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb. *Elife* **7** (2018).

41. J. Schindelin, *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

42. S. J. Orfanidis, *Introduction to Signal Processing* (Prentice Hall, 1995) https:/doi.org/10.1201/9781420042047.ch3.

43. P. A. Correa, T. Gruninger, L. R. García, DOP-2 D2-like receptor regulates UNC-7 innexins to attenuate recurrent sensory motor neurons during C. elegans copulation. *J. Neurosci.* **35**, 9990–10004 (2015).

C h a p t e r 3

PEZO-1 AFFECTS FECUNDITY IN *CAENORHABDITIS ELEGANS*

Katherine I. Brugman

3.1 Abstract

Piezo proteins, a class of homotrimeric mechanosensitive non-selective cation channels, function in a number of roles across the species they are extant in, the same piezo often involved in regulating multiple functions within the same organism. In this chapter, we examine how *Caenorhabditis elegans* is no different in this regard, with *pezo-1* mutants also demonstrating unique piezo function in fecundity in addition to the previously categorized mating defects. To properly investigate this response, we used several early-stop codon strains, three Million Mutation Project and one generated through Crispr, and two deletion strains, each mutation placed at different points along the *pezo-1* gene. With these mutants, we have been able to both characterize the extent of the fecundity defect and demonstrate that at least two *pezo-1* isoforms are not involved in this behavior.

3.2 Introduction

Piezo proteins are a newly identified class of mechanosensitive non-selective cation channels that have been observed sensing sheer stress, static pressure, and membrane force in a variety of different model organisms including mice, zebrafish, and drosophila (1–6). The protein structure has been identified to be homotrimeric—consisting of a large central pore with a cap and three N-terminal propeller regions—and vertebrates have been identified to possess two piezo channels, while invertebrates typically only have one (1, 7, 8). In the human piezo proteins, *PIEZO1* and *PIEZO2*, mutations cause a number of neuropathies including dehydrated hereditary xerocytosis, generalized lymphatic dysplasia, and distal arthogryposis type 5 (9–13). In *Caenorhabditis elegans*, we have shown in Chapter One that this protein exhibits some light touch sensory function in male mating behavior, but there has previously been no piezo function described for the hermaphrodite animal.

The *Caenorhabditis elegans* hermaphrodite is an androdioecious nematode with a two-armed gonad structure (14, 15). While males do exist in the species, nearly all progeny are produced through self-fertilization, a process that typically occurs in five basic steps. First, oocytes develop in the proximal arms of the gonad, and continue to grow as they move

around the loop region to the vulva side(16). Second, they are passed into the appropriate spermatheca, anterior for oocytes coming from the anterior gonad arm, posterior for the oocytes coming from the posterior gonad arm, and fertilized by the sperm, held in place by the distal and spermathecal-uterine valve to ensure that only one oocyte exists in the spermatheca at a time and that no sperm are lost into the uterus when the fertilized egg is released. Third, when the oocyte has been fertilized, it enters the uterus where it develops to a 30-cell stage. Last, the fertilized egg is expelled via the vulva, a muscular organ comprised of eight muscles organized into a “v” shape, which contracts to lay a fertilized egg into the environment.

3.3 Methods

General methods

Worms were cultured at 20˚C as described by Brenner and the experiments were conducted at the same temperature.

CRISPR-Cas9 deletion mutants

As described in Chapter One, the C-terminal deletion mutant, *pezo-1 (sy1113)*, was generated by injecting 46168 CRISPR-Cas9 plasmid (peft-3::cas9-sv40_NLS::ttb-2 UTR) and four pRB1017 plasmids (U6::gRNA construct) containing our guide RNA sequences (CAGAAGCTCGTAAGCCAGG, AGGTCGAGGTCGTGAGCGG, CCACCACTTTACGAGATGG, ATAGGCAGCTTCGAACTGG) into young adult hermaphrodites. A *dpy-10* guide RNA plasmid, also inserted into pRB1017, and a repair oligo with the *dpy-10* mutation were also included in the mixture as a marker and were later crossed out of the strain prior to use in the assay. Worm lysate was split between a control PCR with one internal and external primer (GGCGACCGTTGGTTGGGCTGGTTT, AATACGAGAGCCTTCACATCATC) and a detection PCR with two primers external to the

proposed deletion (GGCGACCGTTGGTTGGGCTGGTTT, ATCCTGTGTCCGATCCTGACG) to verify homozygous mutants. The deletion was also sequenced and verified to be in frame and at the location, IV:9347253...9345244 (Figure 3.1).

The short isoform deletion, *pezo-1 (sy1398)*, was also generated through this method, using four guides (GAGAACTTGAATTCAATGG, AAGCTTCTTCCGTCTCCGG, GCAGTATTTGACCAACTGG, ATAAAACAAGGCAACCAGG) along with the same *dpy-10* guide and repair oligo. The deletion was verified with two external primers (CTCTCGCCTATCCACTTGAGCTTA, GGAAACAATTGAGCCGAGAATGGA) and sequenced to place it at IV: 9356086… 9354777 (Figure 3.1).

CRISPR-Cas9 STOP-IN mutant

As in the previous chapter, this mutant was generated using a purified guide RNA near the mutation location (CCAGAAGCTCGTAAGCCAGG), a single-stranded DNA STOP-IN repair oligo (cttatcgctgtttctgaaccagaagctcgtaagccGGGAAGTTTGTCCAGAGCAGAGGTGACTAAGTG ATAAgctagcaggaggcactgaagaaacggatggtgatgaag), purified Cas9 protein, and both a guide RNA and repair oligo for a *dpy-10* marker mutation. These reagents were injected into N2

hermaphrodites and screened via the *dpy-10* marker to identify successful injections and PCR (GACAGGACTTTCCCGCCAACTTAA, ATCATTCGCCGATTGCACAAGTTG) and an internal primer (GCTTATCACTTAGTCACCTCTGCTC) to identify mutant animals (Figure 3.1).

Million Mutation Project strains

Three stop codon mutation strains from the Million Mutation Project (*pezo-1 (gk364409)*, *pezo-1 (gk208807)*, *pezo-1 (gk892838)*) were acquired from the CGC for use in this study (Figure 3.1).

GFP::pezo-1 transcriptional fusion

The same transcriptional fusion as described in Chapter One was used to identify potentially relevant structures for this study, inserting non-coding regions from 4kb in front of the N-terminal ATG and the entire central intron region (TCCTACCTTGTGCCTGTCTACGTA…TCTGTTGTTGCTGTGTTACGACCA and GCACATGCTGCAGGTTAGTTTATT...AGGGTATCGCCTGAAAAAACAATG) into pPD95_75. This was then injected with the pBX *pha-1* rescue plasmid into *pha-1; him-5*

animals to create a stable expression line (Figure 3.1).

Fecundity Assay

The fecundity assay was performed as described in Wang 2019 (17). Both N2 and mutant animals were isolated at the L4 stage and singled onto plates in batches of 20 per strain per trial. The animals were then passed to fresh plates every day until they stopped laying eggs. One day after the animal was removed from a plate, all the surviving progeny were counted from this plate as well as the dead eggs. At the end of the assay, the number of live progeny was totaled up per worm (though the daily data was also retained).

3.4 Results

Expression pattern data shows pezo-1 expression in the vulva and spermatheca.

Data from the expression pattern constructs in Chapter One reveals *pezo-1* expression in both the vulva muscle and spermatheca (spermatheca in collaborator's data) (Figure 3.2).

Most pezo-1 mutations yielded reduced brood sizes.

The *pezo-1* C-terminal deletion mutation, the STOP-IN mutation, and all three Million Mutation Project early stop codon mutations yielded significantly reduced brood sizes compared to N2, as can be seen in Figure 3.2. The data, distributed by day, shows that day 1 animals had roughly similar numbers of progeny survive as wildtype, and that number of viable progeny as compared to wildtype decreased as the animals aged (Figure 3.3).

The small isoform deletion unique to isoforms i and j does not yield a reduced brood size.

The small isoform deletion, which deletes the first exon of isoforms i and j and should be unique to those isoforms, does not demonstrate the phenotype of reduced broodsize, instead showing wild type progeny numbers for all days of the assay.

Strains with the fecundity phenotype produced no significant increase of dead eggs compared to wild type.

Strains with the fecundity phenotype did not produce significantly more dead eggs than their wild type counterparts as can be seen in Figure 3.4.

3.5 Discussion

We see a correlation between expression pattern data and mutant phenotype: We showed that the *pezo-1* mutants have a severe reduction in brood size compared with wildtype animals. We also showed that *pezo-1* is expressed in the spermatheca and vulva muscle, suggesting that *pezo-1* has some function in the spermatheca and/or vulva muscle that impacts the proper fertilization and laying of eggs by the hermaphrodite, possibly in determining the presence of an egg in order to instruct the proper function of these structures.

We do not see a phenotype in the deletion in the first exon of the small isoform: We have shown that isoforms i and j do not produce any brood size reduction when the first exon of this isoform is deleted, which would point to this isoform not having a function in fecundity. However, it is important to note that our collaborator, Xiaofei Bai, has observed in a parallel experiment that there was a slight brood size reduction ($p=.0001$), so it's possible these isoforms may have limited function.

We note no increase of dead eggs to account for the reduction of brood size: We have

shown by counting the dead eggs on the plates that the *pezo-1* mutants do not exhibit an increase in the number of dead eggs present, so the number of dead eggs seen cannot account for the "missing" progeny of the total brood size. So, where have the missing progeny gone? The answer to this likely lies in *pezo-1*'s spermathecal valve and the vulva muscle. A malfunctioning mechanosensory ion channel present in either of these two structures could result improper movement of the egg through these passages, resulting in damaged or destroyed eggs, the former of which would appear as a few dead eggs, the latter of which might not be visible under the imaging techniques used in this study.

Implications for piezo: This is the first instance of a piezo protein showing a function in a self-fertilization mechanism in an organism. While the precise methods by which piezo functions in this process are unclear, it may resemble the vasodilatory function seen in mice. Our collaborator, Xiaofei Bai, was able to provide more insight into this problem in his paper (currently in review) which can be summarized by the spermathecal valves spasming as oocytes pass through them, causing these oocytes to be destroyed. This also accounts for the lack of dead eggs seen as most destroyed oocytes would not be visible under a standard dissecting scope. As egg-laying and passage through the spermatheca still occur to some

degree, however, it is possible there are other mechanosensors involved in this process or that chemosensation plays a part in determining the presence of an egg in these structures. Also of note is how the age of the animal plays a role in the severity of the phenotype, with young animals displaying near wild type egg laying and older hermaphrodites exhibiting severely reduced broods. This could point to the mechanosensory nature of the associated structures being age-linked. A possible mechanism for this to occur would involve unfertilized oocytes being pushed evenly into the spermatheca and out the vulva by upstream pressure when the worm is young, but as the worm ages and the gonads become more jumbled with multiple rows of oocytes, the “feeding” process is less structured and relies more on the mechanosensory apparatus for guidance. At this point, lack of mechanosensory input results in destroyed or damaged oocytes and a reduced brood size.

3.6 Figures

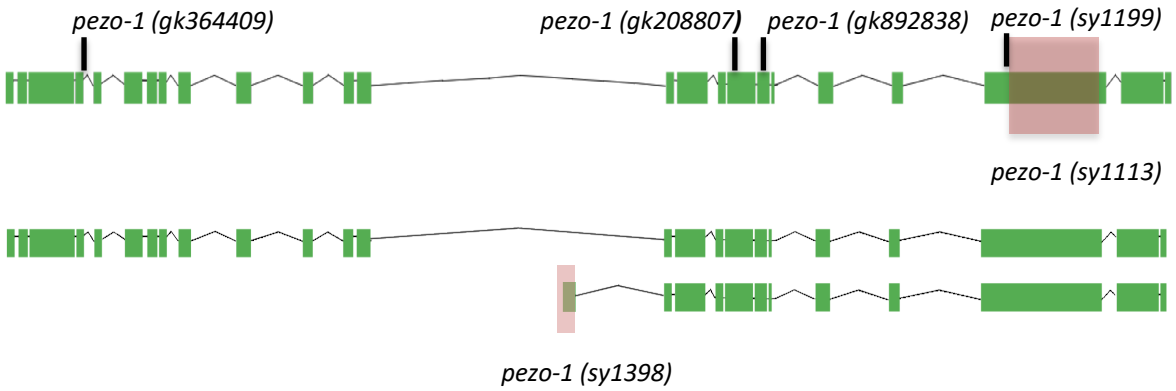


Figure 3.1 Gene diagram of all mutations used in this study. VC20634, VC20227, and VC40944 are Million Mutation Project strains containing a stop codon at these locations. *pezo-1 (sy1199)* is a STOP-IN mutation generated by protein CRISPR/Cas9, and *pezo-1 (sy1113)* is a deletion generated by plasmid CRISPR/Cas9. *pezo-1(sy1398)* is a deletion generated at the first exon that only exists in two small isoforms, one shown here as an example.

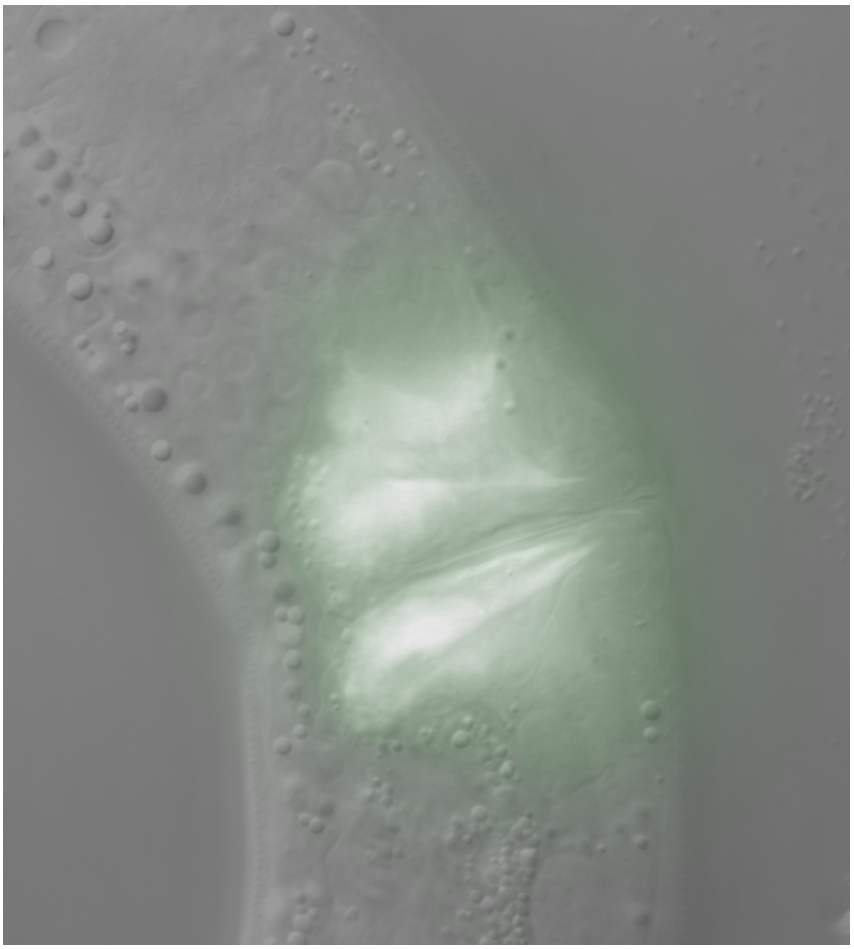


Figure 3.2 *pezo-1* GFP expression pattern of vulva muscle, potentially involved in part in the destruction of the oocyte.

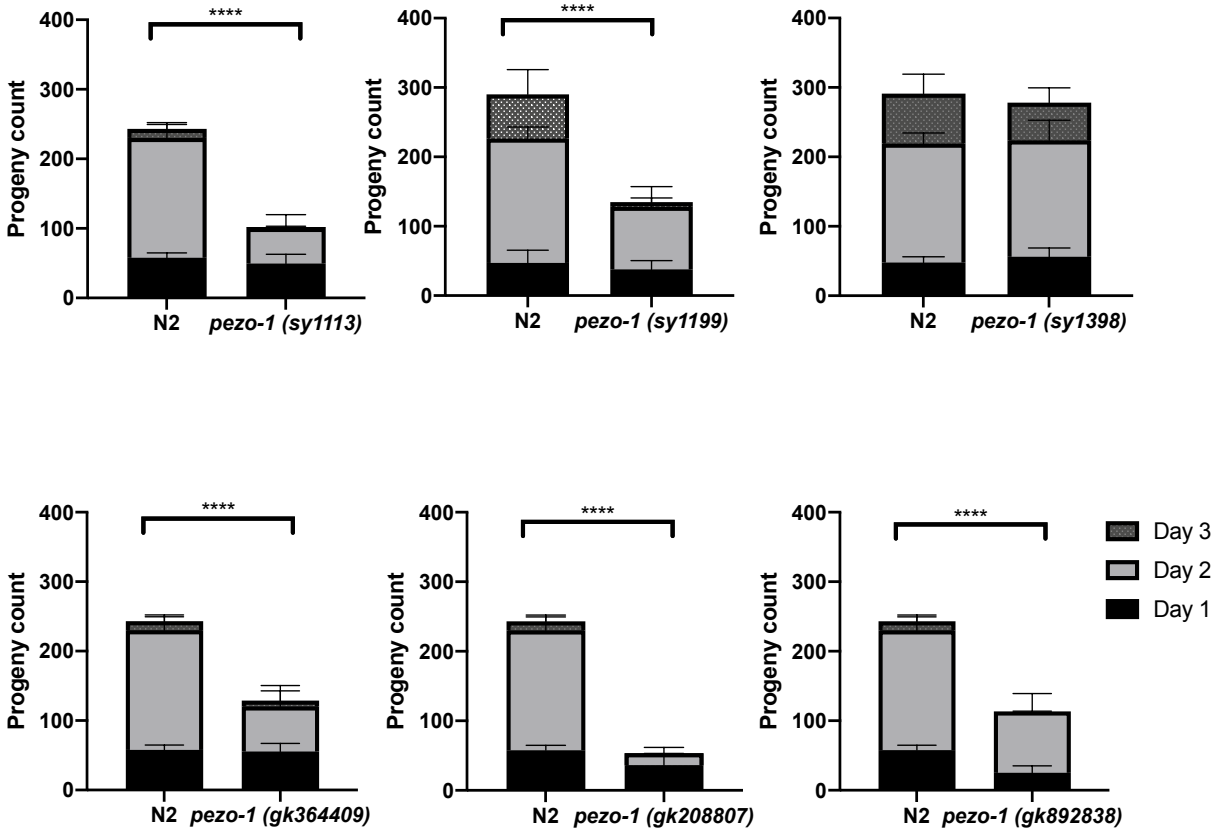


Figure 3.3 Fecundity assay data for all mutants tested. As seen here, mutant animals had wildtype-like progeny numbers on day 1, but on subsequent days, their progeny dwindled rapidly, with a marked decrease on day 2 compared to wildtype. The exception here is *pezo-1 (sy1398)*, which seemed to have minimal effect, though in another study by a collaborator, it was shown that this mutant has a slight decrease in progeny (data shown in Appendix – B).

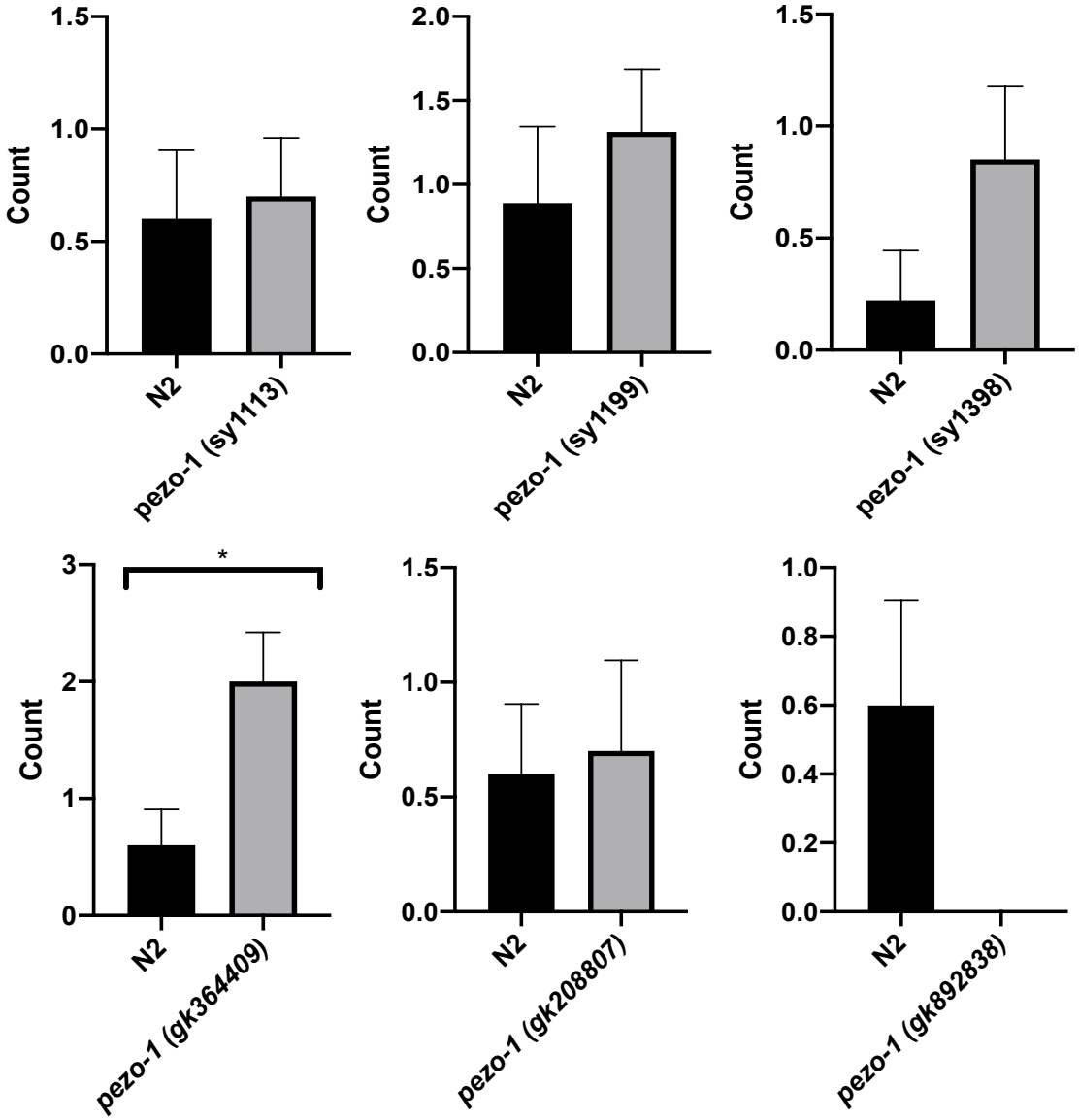


Figure 3.4 Cumulative dead eggs observed per plate assayed for fecundity over all days assayed. As can be seen here, the number of dead eggs is not significantly different from that observed in N2 animals, nor does it account for the large number of missing progeny shown in Figure 3.3.

3.7 References

1. B. Coste, *et al.*, Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science (80-.).* **330**, 55–60 (2010).

2. B. Coste, *et al.*, Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* **483**, 176–181 (2012).

3. A. Faucherre, J. Nargeot, M. E. Mangoni, C. Jopling, piezo2b regulates vertebrate light touch response. *J. Neurosci.* **33**, 17089–17094 (2013).

4. A. Faucherre, K. Kissa, J. Nargeot, M. E. Mangoni, C. Jopling, Piezo1 plays a role in erythrocyte volume homeostasis. *Haematologica* **99**, 70–75 (2014).

5. S. S. Ranade, *et al.*, Piezo2 is the major transducer of mechanical forces for touch sensation in mice. *Nature* **516**, 121–125 (2014).

6. S. S. Ranade, *et al.*, Piezo1, a mechanically activated ion channel, is required for vascular development in mice. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10347–10352 (2014).

7. J. Ge, *et al.*, Architecture of the mammalian mechanosensitive Piezo1 channel. *Nature* **527**, 64–69 (2015).

8. K. Saotome, *et al.*, Structure of the mechanically activated ion channel Piezo1. *Nature* **554**, 481–486 (2018).

9. C. Bae, R. Gnanasambandam, C. Nicolai, F. Sachs, P. A. Gottlieb, Xerocytosis is caused by mutations that alter the kinetics of the mechanosensitive channel PIEZO1. *Proc. Natl. Acad. Sci. U. S. A.* **110** (2013).

10. J. Albuisson, *et al.*, Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels. *Nat. Commun.* **4**, 1884 (2013).

11. R. Zarychanski, *et al.*, Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis. *Blood* **120**, 1908–1915 (2012).

12. B. Coste, *et al.*, Gain-of-function mutations in the mechanically activated ion channel PIEZO2 cause a subtype of Distal Arthrogryposis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4667–4672 (2013).

13. C. Bae, P. A. Gottlieb, F. Sachs, Human PIEZO1: Removing inactivation. *Biophys. J.* **105**, 880–886 (2013).

14. S. Brenner, The Genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).

15. J. Kimble, D. Hirsh, The Postembryonic Cell Lineages of the Hermaphrodite and Male

Gonads in Caenorhabditis elegans. *Dev. Biol.*, 396–417 (1979).

16. J. McCarter, B. Bartlett, T. Dang, T. Schedl, On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans. *Dev. Biol.* **205**, 111–128 (1999).

17. W. R. Wong, *et al.*, Autism-Associated missense genetic variants impact locomotion and neurodevelopment in Caenorhabditis elegans. *Hum. Mol. Genet.* **28**, 2271–2281 (2019).

C h a p t e r 4

OTHER OBSERVATIONS AND FUTURE DIRECTIONS

4.1 Observations

Summary

I have shown in this thesis that the *C. elegans* piezo protein has some function in various aspects of male mating and fecundity. While a significant portion of the relationship between piezo and fecundity has been solved by my collaborator, Xiaofei Bai, there is still quite a lot of ground to cover in the subject of male mating as well as some other observations that I have made in the *pezo-1* defective worms. In this section, I'm going to discuss other potential *pezo-1* phenotypes in *Caenorhabditis elegans* that didn't merit an entire chapter for themselves and future directions for *pezo-1* research in male mating.

pezo-1 is expressed in the CAN neurons

The CAN neurons are a pair of neurons in the midbody of *C. elegans* with processes that run along the excretory canal. These neurons currently have an unknown function as ablating them causes the animal to die, a discovery first noted by Sulston and Hodgkin in personal communication, later confirmed by Forrester and Garriga (1). In fact, of all the neurons in *C. elegans*, only the CANL, CANR, and M3 neurons are essential in this way, as

reported by Chalfie and White (2). So, while the function of these neurons is unknown, the presence of a mechanoreceptor in them is intriguing, and the fact that a null mutant of this mechanoreceptor doesn't cause the death of the animal provides the opportunity to discover a function for these neurons (Figure 4.1).

One potential hypothesis here would work in conjunction with the data published in the thesis by James Lee which showed that *pezo-1* animals showed decreased dauer entry (an alternate lifecycle a larvae enters instead of proceeding directly to adulthood) by two fold, which seemed like a peculiar role for a mechanotransducer to be involved in (3). Animals typically enter a dauer state due to a change in external conditions, such as heat, crowding, dessication, or lack of food (4, 5). As one important function of the excretory duct is osmoregulation, a defect in *pezo-1* could result in a mutant animal not getting the signal from the CAN neurons to go into dauer in a dessicating environment (1).

One way I attempted to test this was via a split-CGaL construct to create a cell-specific GCamP for the CAN neurons in a *pezo-1* mutant animal and test different osmotic conditions. However, none of the split-CGaL promoters currently available for the CAN neurons were strong enough to see the fluorescence clearly enough to proceed with the

experiment (6). In the future, with a stronger promoter or a brighter GCamP, this could a promising opportunity.

pezo-1 **mutants demonstrate a defect in sperm release during mating**

This is an isolated result taken from an incomplete mating assay performed before directly before quarantine. In this assay, *pezo-1 (sy1113)* mutants demonstrated a significant delay of a minute or more before sperm release in 41% of mating events (n=17), while N2 worms showed this delay in 0% of mating events (n=8). This could be linked to *pezo-1* expression in the SPCL/R neurons.

pezo-1 **mutants show an acceleration phenotype linked to** ***trp-4*** **and** ***mec-4***

This is an isolated result taken from a series of exploratory locomotion assays I did to test the effects of *pezo-1 (sy1199)* and various combinations of mutants with *mec-4 (sy1124)* and *trp-4 (sy695)* (Figure 4.2). As can be seen in the figure below (program analysis designed by John Brugman) *pezo-1* showed a slight acceleration increase compared to wild type

animals, but, when paired with *mec-4* and *trp-4* mutants, caused a significant decrease in acceleration compared to the *mec-4; trp-4* double. The *mec-4; pezo-1* double mutant showed very little change compared to the *mec-4* single mutant. The numbers here are unfortunately too small to make any great claims, but the data does suggest further interactions between *pezo-1* and *trp-4*, and, now that a *pezo-1; trp-4* mutant is extant, the next step in this investigation would be to test and analyze this mutant for aberrant acceleration.

Future directions with *pezo-1* fecundity

While most of the main mechanisms of *pezo-1* fecundity have been elaborated on extensively, the phenotype is incredibly robust and is stable enough that even defects caused by point mutations can be detected. This presents an opportunity for studying the effects of targeted mutations on the function of *pezo-1 in vivo*. *In vitro* studies of membrane proteins like piezo are challenging, both due to their oligomeric nature and the requirement of a membrane to fold correctly. *In vivo,* neither is a serious consideration. With CRISPR/Cas9, suspected interaction sites can now be targeted with great precision, so investigations into

specific domains, should future structural analysis reveal good targets, would probably yield good results.

Future directions with *pezo-1* mating

While *pezo-1* and *trp-4* have been shown to have function in the PCS neurons, ray neurons, and HOB, the exhibited phenotype has been incompletely penetrant, which suggests the possibility of another associated mechanotransducer. In mice, an ortholog of *mec-4* has been found to associate with *pezo-1* in some fashion, so it's feasible that some similar arrangement occurs in *C. elegans*. In coordination with James Lee, a *pezo-1; mec-4; trp-4* triple mutant was designed earlier for use in locomotion and dauer assays, though some work is necessary to make this mutant suitable for a mating assay (namely the addition of a *him-5* mutation). I would expect that the observed mating defects would be far more pronounced in this mutant, possibly to the point of creating a mutant that would be unable or nearly unable to mate.

In addition, the full-brain imaging data would be useful to have for *trp-4* mutants as well as *pezo-1* so as to further elucidate the interplay between these two mechanotransducers.

I would expect to see the *trp-4* mutant create a defect primarily in PCA to counter the defects of *pezo-1*. Creation of a version *trp-4* mutant with the pan-neuronal GCamP is currently in progress.

Conclusion

In conclusion, we have shown that *pezo-1* is involved in male mating and some processes involved hermaphrodite self-fertilization. We have also shown that *trp-4* is involved in male mating and is potentially involved in some of the same pathways as *pezo-1* both redundantly and opposed. In addition, we have highlighted some potential further areas in *C. elegans* that *pezo-1* may be involved in, though further research will need to be done to confirm the full extent of its involvement.

4.2 Figures

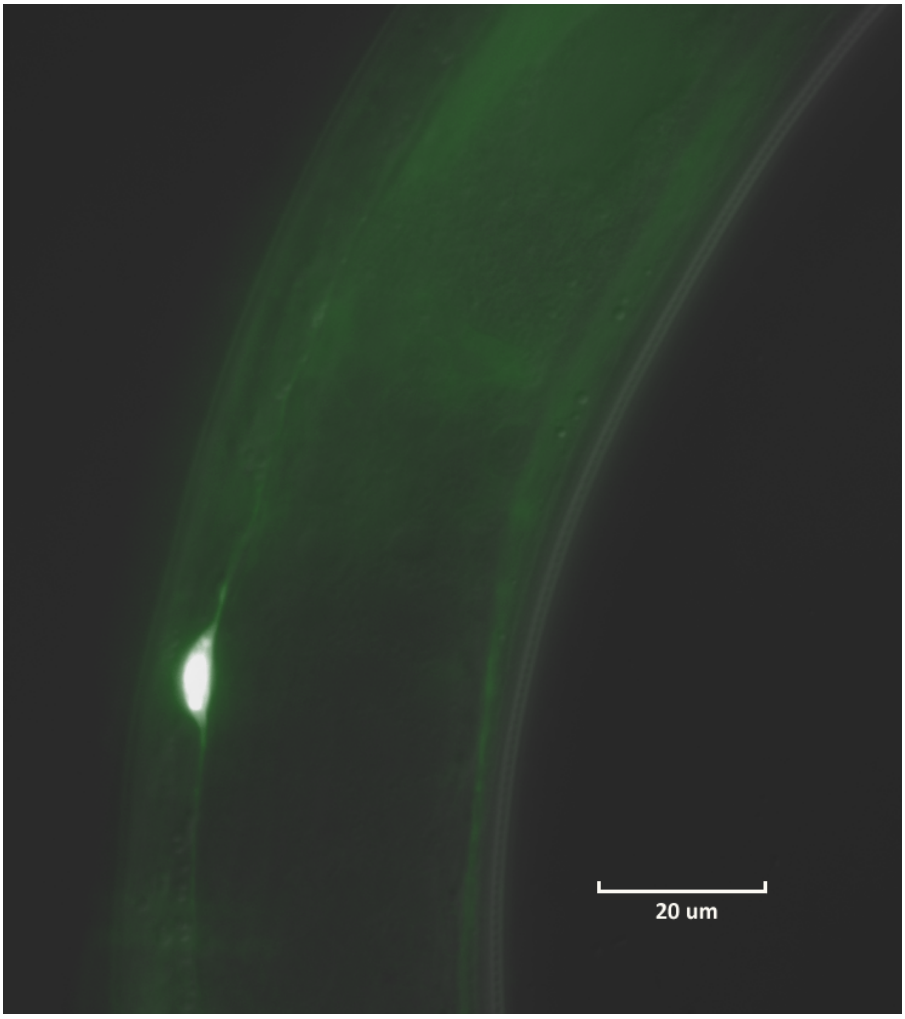


Figure 4.1 *pezo-1* promoter GFP transcriptional fusion showing *pezo-1* expression in the CAN neuron. The CAN neuron is a neuron of unknown function, and previous studies have shown that ablating it kills the worm, making *pezo-1*'s presence in it an intriguing line of study.

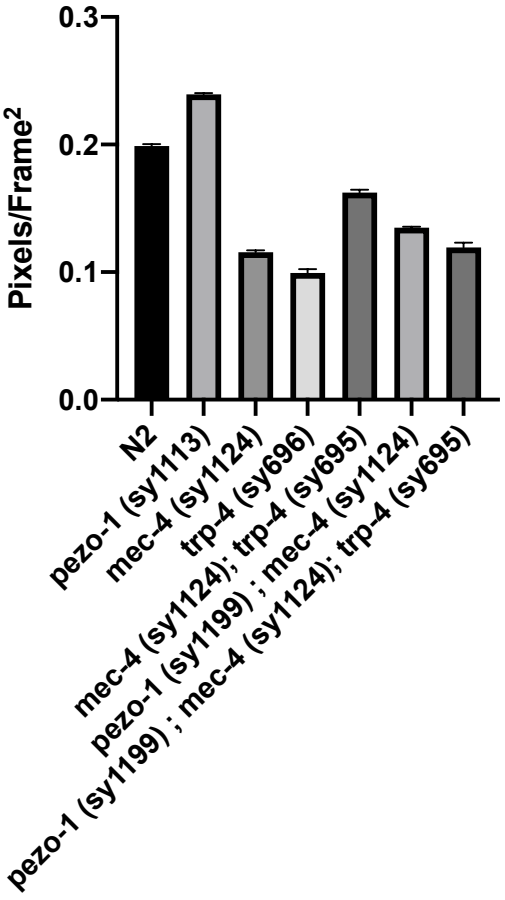


Figure 4.2 Acceleration of the worm strains tested measured in pixels traveled per frames^2 (n=10 per strain). The sample size is small, but the data is very tightly clustered and there is some indication here that while *mec-4 (sy1124)* and *trp-4 (sy695)* have significantly slower accelerations than N2, the double mutant has a statistically significantly higher acceleration than either signal mutant put together ($p<.0001$). But when the *pezo-1 (sy1199)* mutant is added to the strain, despite the *pezo*-1 mutant typically not having an effect on acceleration, the triple mutant's acceleration drops back down significantly ($p<.0001$).

4.3 References

1. W. C. Forrester, G. Garriga, Genes necessary for C. elegans cell and growth cone migrations. *Development* **124**, 1831–1843 (1997).

2. M. Koga, Y. Ohshima, Mosaic analysis of the let-23 gene function in vulval induction of Caenorhabditis elegans. *Development* **121**, 2655–2666 (1995).

3. J. S. Lee, "The genomics of stress-induced life cycle decisions in nematodes." (2018).

4. J. W. Golden, D. L. Riddle, A pheromone influences larval development in the nematode Caenorhabditis elegans. *Science (80-.).* **218**, 578–580 (1982).

5. J. W. Golden, D. L. Riddle, The Caenorhabditis elegans dauer larva: Developmental effects of pheromone, food, and temperature. *Dev. Biol.* **102**, 368–378 (1984).

6. H. Wang, J. Liu, K. P. Yuet, A. J. Hill, P. W. Sternberg, Split cGAL, an intersectional strategy using a split intein for refined spatiotemporal transgene control in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3900–3905 (2018).

APPENDIX-A: Additional Pezo-1 Mating Data

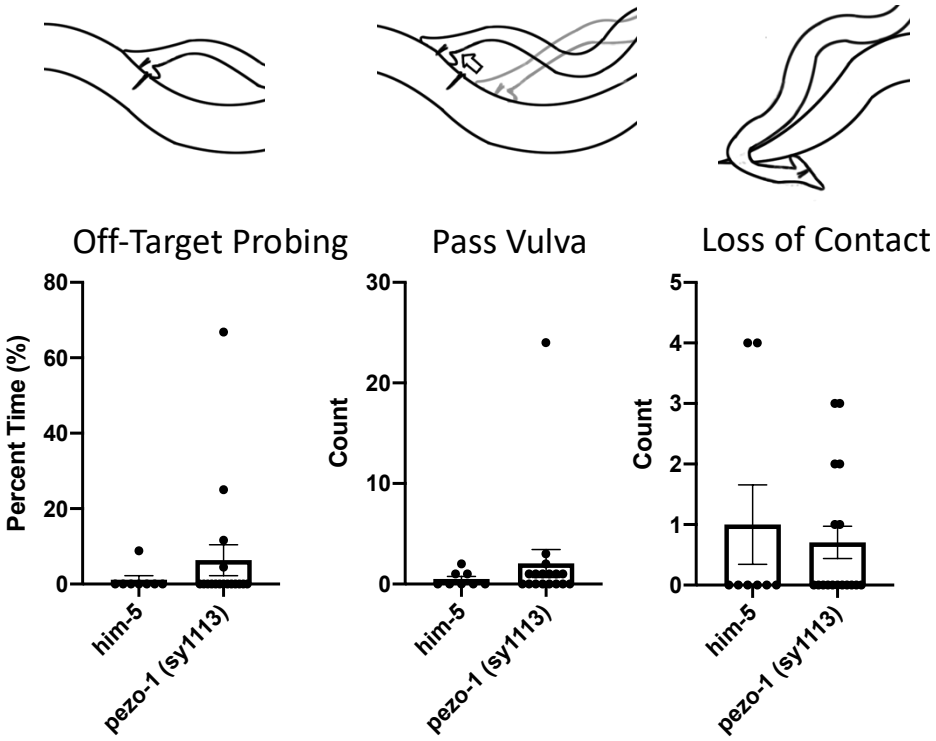


Figure A.1 Data from incomplete secondary assay for *pezo-1 (sy1113)*. As can be seen here there is a lesser degree of the phenotype and more spread in the individual data points than the first assay, however, some of this may be attributed to the far smaller sample size, particularly in the controls (n=8). The loss of contact phenotype in particular is particularly shaky here due to two wild type males that experienced a large number of contact

interruptions during mating. In comparison to the original assay, we see more along the lines of 20% phenotypic animals for probing, which is the most obvious phenotype.

APPENDIX-B: ADDITIONAL PEZO-1 FECUNDITY DATA

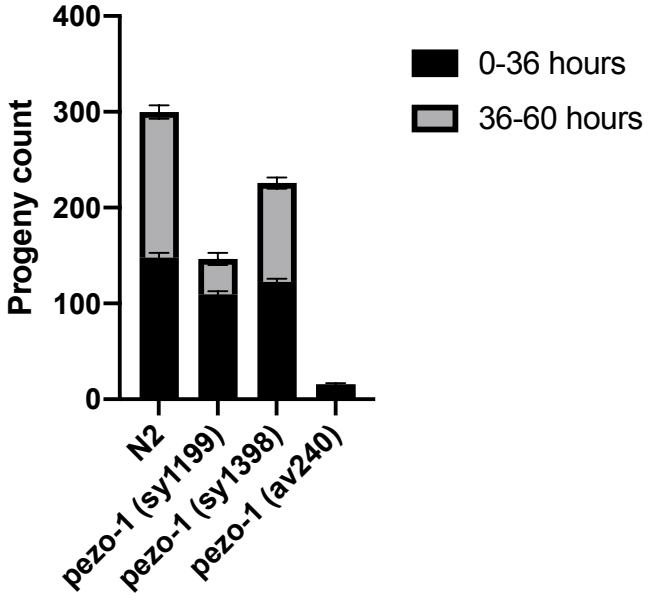


Figure B.1 Data collection on fecundity by Xiaofei Bai, plotted by me. The assay here is structured slightly differently, with two data collection points at 36 and 60 hours instead of every 24 hours, however a similar trend of early progeny levels similar to wild type followed by a sharp drop off in offspring can be seen, save for *pezo-1 (av240)* a full-deletion mutant I did not possess, which had minimal progeny overall. The other primary difference seen here is with *pezo-1 (sy1398)*, which has a notable decrease in progeny in the second half of the assay with a $p < .0001$, which runs contrary to my own analysis. This could be due to varying conditions and represent a marginal effect of this particular isoform on *C. elegans* fecundity that I did not detect in my assay but was detectable under stricter temperature controls.

APPENDIX-C: ADDITIONAL RESEARCH

Autism-associated missense genetic variants impact locomotion and neurodevelopment in *Caenorhabditis elegans*

Wan-Rong Wong[1], Katherine I. Brugman[1], Shayda Maher[1], Jun Young Oh[1], Kevin Howe[2], Mihoko Kato[1], and Paul W. Sternberg[1,*]

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States of America

[2]European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, United Kingdom

[*]Correspondence: pws@caltech.edu

1200 E. California Blvd, MC 156-29, Pasadena, CA 91125, USA

TEL: 626-395-2181, FAX: 626-568-8012

C.1 Abstract

Autism spectrum disorder (ASD) involves thousands of alleles in over 850 genes. The causal relationship of most of these genetic variants in the pathogenesis of ASD has not yet been demonstrated. Current functional inference tools used to predict damaging effects are not sufficient to inform clinical cases. Thus, an experimental method prioritizing missense alleles for further intensive analysis is crucial. Here, we describe a pipeline that uses *C. elegans* as a genetic model to screen for phenotype-changing missense alleles inferred from human ASD studies. We identified these highly conserved human ASD-associated missense variants in their *C. elegans* orthologs, and used a CRISPR/Cas9-mediated homology-directed knock-in strategy to generate missense mutants and analyze their impact on behavioral and development. All tested missense alleles are predicted to be phenotype altering, but we found only 70% of them showed detectable phenotypic changes in morphology, locomotion and fecundity as measured in the related assays. Our findings indicate that missense variants in the *C. elegans* orthologs of human *CHD7*, *CHD8*, *GLRA2*, *NAA15*, and *TPH2* impact neurodevelopment and movement functions, elevating these genes as candidates for future study into ASD. Our approach will help prioritize functionally important missense variants for detailed studies in vertebrate models and human cells.

C.2 Introduction

Many psychiatric disorders such as autism spectrum disorder (ASD, OMIM: 209850) have been linked to genetic variants that disrupt but not necessarily eliminate protein functions. Most studies focus on identifying likely gene-disruptive (LGD) mutations (e.g., nonsense, frameshift, or splice-site) instead of missense mutations because only 13% of identified *de novo* missense variants are suspected to contribute to the risk of ASD (1, 2). Missense mutations (i.e., point mutations in which a single nucleotide change results in a codon that encodes a different amino acid) account for approximately half of the genetic changes known to cause disease (3). The severity of ASD is thought to be correlated with the average contribution of familial influences and *de novo* mutations (4): individuals with ASD are more likely to carry a *de novo* missense mutation (2). The pathogenesis of autism-risk alleles has been supported by functional studies of CTNND2 (5) and GLRA2 (6). Given that some missense alleles have been validated, one challenge is to identify the subset of ASD-associated mutations that are deleterious.

Missense mutations are numerous, but the functional inference tools used to predict damaging effects are not accurate enough to be used as the sole basis for a conclusion. Most current software relies heavily on sequence conservation to predict the potency of missense

variants, as conserved regions are considered more likely to be affected by purifying selection (7). However, interpretation of these data is inadequate due to variable penetrance, dosage sensitivity, and functional redundancy of mutated proteins, and can result in a high false-positive rate of prediction (3, 8). For example, only 27% of missense mutations predicted by sequence conservation showed disruption of protein function in a recent rodent study (9). Evaluating such mutations *in vivo* is essential for generating an accurate interpretation of available data.

Due to the large number of missense variants, an efficient pipeline is needed to evaluate the functional consequence of all residues. Given that any gene carries a certain chance of containing a missense mutation and every individual will have a different subset of missense mutations in their genome, computational analyses are insufficient for predicting the functional importance of such mutations (9). There are few studies conducted to validate the functional consequence of missense variants *in vivo*. One study evaluated the disruptiveness of a mutation exclusively on its capacity to disrupt protein interactions using the yeast two-hybrid method (10). Another study compared the deleterious effect from computational prediction to the actual ENU mutant rodent models (9). Despite such exciting findings, a comprehensively targeted screen for function-disrupting ASD-associated missense mutations

in a multi-cellular model organism has not yet been done. The short life cycle and easily accessible genome in *Caenorhabditis elegans* (*C. elegans*) make it an useful tool to rapidly evaluate whether a particular disease-associated missense variant shows phenotypic consequences (11–13).

In this study, we established a pipeline for identifying ASD-associated protein-disrupting missense residues in the orthologous *C. elegans* proteins (Figure C.1A). First, the *C. elegans* residues corresponding to human missense variants were identified based on sequence conservation. The *C. elegans* equivalents of human missense mutants were generated using Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-Cas9 and homology-directed genome editing ("knock-in"). We then analyzed the effects of these autism-associated missense alleles by comparing observable phenotypes from these missense mutants to the wild type and to known loss-of-function mutant controls. Missense mutants with phenotypic changes reflect alteration in protein function, indicating the importance of these alleles. We found that 19% of the ASD-associated missense variants are conserved in *C. elegans*. We evaluated the effects of 20 missense alleles that were predicted to be phenotype altering and found that only 70% of them displayed phenotypic changes in morphology, locomotion, and fecundity. Our method demonstrates our ability to screen for

subtle phenotypic changes and in doing so, illustrate the functional importance of the effect of missense mutations on human disease.

C.3 Results

Identifying C. elegans analogs of ASD-associated missense mutations

To identify the functionally important missense variants implicated in complex human diseases, we established a pipeline to screen for functional changes in orthologous proteins in *C. elegans* (Figure C.1A). Of 1811 human missense variants related to ASD, 778 alleles (43%) from 221 human genes were located in *C. elegans* orthologs. Most of the human genes were aligned to one *C. elegans* ortholog, but a few of them (47 of 221) had more than one orthologous protein in *C. elegans* (Figure C.2A). The challenge was to identify each orthologous protein and corresponding equivalent residue based on sequence conservation. To this end, our software utilized comparative genomics and multiple alignments to ensure that the detected residue reflects conservation across the evolutionary tree and gene family. We found that 345 (19%) of missense loci from 157 human genes not only have orthologs in *C. elegans* but also had at least one conserved amino acid residue between human and *C. elegans* (Figure C.2). Sometimes, one human residue could be matched to multiple orthologs in *C. elegans* (37 of the 345 conserved residues). In these cases, we picked the worm residue candidate that had the sgRNA sequence most likely to produce an efficient CRISPR-Cas9 double strand break based on an online sgRNA prediction tool (22). For each allele, we

identified the corresponding *C. elegans* ortholog, assessed the residues affected by missense mutations for evolutionary conservation, and selected genes with a known phenotype for their loss-of-function mutation in *C. elegans* (from existing mutants or RNAi). To prioritize genes for functional screenings, we focused on those genes with multiple missense variants as the chance of one causing a phenotypic defect increases when multiple missense mutations are observed in a single gene (25). We also prioritized genes involved in multiple biological pathways (26) or genes with other mutations resulting in a stop codon.

To capture the impacts of missense mutations in diverse physiological functions, we sampled 20 ASD-associated missense changes in residues conserved in the *C. elegans* orthologs of 11 human genes (Table C.1; Figure S1). These ASD-associated missense mutations were identified in genes that were known to have a role in synaptic function (i.e., *DLG4*, *SYNGAP1, CACNA1D,* and *GLRA2*), gene expression regulation (i.e., *CHD7*, *CHD8, and CUL3)* or neuronal signaling and cytoskeleton functions (i.e., *PTEN*, *MAPK3*, *TPH2*, and *NAA15*). Multiple aspects of physiological functions were examined, including morphology, locomotion, and fecundity. These well-established quantitative assays enabled us to detect subtle changes in morphology, movement, and coordination, as well as reproduction and complete of embryonic development (13, 27).

Morphology of missense mutants

To examine changes in morphology, we utilized a quantitative tracking system to measure the length, width, and body area of these missense mutants under freely moving condition (Table C.2). Alterations in size were detected in *egl-19/CACNA1D*, *avr-15/GLR2*, *chd-7/CHD7 or CHD8*, *hpo-29/NAA15*, and *tph-1/TPH2*. Every *chd-7* mutant tested showed a significant decrease in body width and area. A null mutant, *chd-7(sy956),* displayed the most severe defects. Other missense alleles, *chd-7(L1220P)*, *chd-7(L1487R)*, *chd-7(G1225S)*, and *chd-7(P253L)* showed a milder degree of defects. One of the *egl-19* missense mutants, *egl-19(Y333S),* displayed a smaller decrease in body length, width, and areas compared to the semidominant allele, *egl-19(n2368)* (18). Another *egl-19* mutant, *egl-19(V331M)*, showed a similar body size as the N2 wildtype strain. Similarly, the *tph-1(R259Q)* mutant showed a decrease in body length and area, and the change was milder in the missense mutant as compared to the null mutant *tph-1(mg280)*. One missense mutation in *avr-15, avr-15(R364Q),* caused a decrease in body length, width, and area, similar to its null mutant, *avr-15(ad1051)* (17). Another *avr-15* missense mutant, *avr-15(N347S)*, showed no morphological changes. Missense mutant *hpo-29(L575S)* exhibited shorter body length. Missense mutants

cul-3(H728R), *daf-18(H168Q)*, and *gap-2(C417Y)* displayed increased body width and area. Missense mutants of *dlg-1/DLG4* and *mpk-1/MAPK3* did not show morphological changes.

Movement and coordination of missense mutants

To examine movement and coordination in these missense mutants, a quantitative tracking system was used to measure moving speed, reversal rate, and sinusoidal wavelength and amplitude (Figure C.3; Table S2). Locomotion defects were found in missense mutants of *chd-7/CHD7* or *CHD8*, *daf-18/PTEN*, and *gap-2/GLRA2*. Less severe than the null mutant, all missense mutants in *chd-7*, except *chd-7(P253L)*, exhibited decreased speed. Missense mutant *hpo-29(L575S)* also showed a significant decrease in speed. In terms of reversal rate, most *chd-7* mutants, except *chd-7(P253L)*, showed decreased reversal rates. Missense mutants *daf-18(H138R)* and *gap-2(C417Y)* displayed an increased reversal rate. Missense mutations in *avr-15/GLRA2*, *cul-3/CUL3*, *dlg-1/DLG4*, *egl-19/CACNA1D*, *mpk-1/MAPK3*, and *tph-1/TPH2* did not result in differences in speed and reversal rate.

Locomotion in *C. elegans* is typically expressed as the wavelength and amplitude of a sinusoidal wave (28). Motor coordination defects have been associated with ASD (29) and were found in missense mutants of *avr-15/GLRA2*, *chd-7/CHD7* or *CHD8*, *daf-18/PTEN*, and

tph-1/TPH2 (Figure C.3; Table S2; Supplemental video). Similar but less severe than its null mutant, *tph-1(R259Q)* in particular exhibited significantly lower wavelength and higher amplitude, indicating a more curvy sinusoidal wave. One of the *avr-15* missense mutants, *avr-15(R364Q)*, showed a decrease in wavelength, slightly milder than the null mutant, *avr-15(ad1051)*. Another *avr-15* mutant, *avr-15(N347S)*, showed normal sinusoidal shape. All missense mutants in *chd-7*, except *chd-7(P253L)*, displayed a decreased wavelength and/or amplitude. The mutation in *chd-7(L1220P)* resulted in a decrease in both wavelength and amplitude. Mutations in *chd-7(G1225S)* and *chd-7(L1487R)* led to a decrease in wavelength and amplitude respectively. The null mutant of *chd-7* also displayed a decreased wavelength. Two of the *daf-18* missense mutants, *daf-18(H138R)* and *daf-18(H168Q)*, exhibited an increase in amplitude and wavelength respectively. Missense mutations in *cul-3/CUL3*, *dlg-1/DLG4*, *egl-19/CACNA1D*, *gap-2/SYNGAP1, hpo-29/NAA15*, and *mpk-1/MAPK3* did not lead to differences in the sinusoidal wave.

Fecundity of missense mutants

We used the fecundity assay to examine larvae viability in genes with reported sterile or lethal phenotypes in null mutants (Figure C.4). Three of four missense mutations in the

chromatin modifier gene *chd-7* displayed a reduced fecundity phenotype compared to the wild type control strain N2. Specifically, the *chd-7(L1220P)* allele had a median fecundity of 119 ($p < 10^{-6}$); *chd-7(G1225S)* had a median fecundity of 176 ($p = 1.2*10^{-5}$); *chd-7(L1487R)* had a median fecundity of 168 ($p = 4.4*10^{-5}$); and *chd-7(P253L)* had a median fecundity of 254.5 compared to a median fecundity of 228 for N2 control. These missense alleles showed weaker fecundity defects compared to its deletion (*chd-7(tm6139)*) or frameshift (*chd-7(sy956)*) controls, which had median fecundity of 38 and 47 respectively ($p < 10^{-6}$). Missense variants in the DNA replication gene, *cul-3(H728R)*, also displayed a decreased fecundity of 168.5 ($p = 10^{-6}$). In additions, the missense variant *dlg-1(V964I)* showed a reduced median fecundity of 154.5 ($p < 10^{-6}$), which is slightly less severe than the 67% reduction in a previous RNAi study (30). We did not observe changes in fecundity in missense mutants in other genes, namely *daf-18/PTEN*, *egl-19/CACNA1D*, *gap-2/SYNGAP1*, *hpo-29/NAA15*, *mpk-1/MAPK3*, and *tph-1/TPH2*.

Comparison with phenotype-predicting software

To examine the accuracy of our biological platform, we compared our results to the existing prediction software Sorting Intolerant From Tolerant (SIFT) and Polymorphism

Phenotyping v.2 (PolyPhen-2) (Table C.3). SIFT emphasizes sequence conservation and the physical properties of amino acids (31), whereas PolyPhen-2 considers both the analysis of multiple sequence alignments and protein 3D-structures (32). Both software programs are commonly used to predict the effects of non-synonymous amino acid changes. For SIFT, all the alleles tested were predicted to be damaging due to having a similar approach of analyzing sequence conservation as our program. Our phenotypic assays identified six residues (among the 20 predictions) that did not agree with the prediction. As compared to PolyPhen-2's prediction, 35% (7/20) of the phenotypic results do not agree with the predictions. Among the 7 strains that did not match, 5 of them were predicted to have damaging effects but had no phenotypic change in our functional assays (false positive), and 2 were predicted to be benign but displayed phenotypic changes (false negative). Overall, our results demonstrated that 70% (14 of 20) missense alleles predicted to be damaging by at least one functional inference tool and actually showed detectable phenotypic changes in morphology, locomotion, and fecundity.

C.4 Discussion

In this study, we have developed a fast and tractable pipeline to comprehensively screen for ASD-associated missense mutations. Our analysis finds that 43% of the human disease-associated alleles have an ortholog in the genome of *C. elegans*, which is consistent with previous estimates (33). Among the 19% conserved loci, we evaluated 20 missense alleles that were predicted to be damaging and found 70% of them actually cause detectable phenotypic changes. We have successfully prioritized 14 missense variants that are functionally significant in *C. elegans* orthologs of human genes. These are the first animal models with deliberately engineered missense mutations in these loci. Our approach is useful for characterizing novel missense alleles that are potentially relevant to human disease and be used as a tool to identify functionally consequential alleles.

Compared to null mutants, most of the phenotypically altered missense alleles displayed milder phenotypes, indicating that our assays can detect relatively subtle changes in protein functions. For example, the *chd-7* missense mutants and *tph-1(R259Q)* displayed hypomorphic phenotypes less severe than their null mutants (34). The *cul-3(H728R)* and *dlg-1(V964I)* missense mutants displayed a smaller reduction in fecundity compared to previous RNAi studies (30, 35, 36). *avr-15(R364Q)* showed defects in morphology and locomotion

similar to its null mutant, *avr-15(ad1051)*, even though it did not recapitulate the spontaneous reversal rate defect documented in a RNAi study (37). Interestingly, the functional consequences of missense alleles may differ in different assays. For instance, the *egl-19(Y333S)* showed milder morphological changes similar to its null mutant (34) but displayed normal functions in locomotion and fecundity. Missense mutants *hpo-29(L575S)* and *gap-2(C417Y)* displayed defects in morphology and locomotion, but they did not show the fecundity defects reported in RNAi studies (36, 38). The *daf-18(H138R)* and *daf-18(H168Q)* missense mutants showed defects in morphology and locomotion, which were not documented before. These results suggested a role for our biological screening platforms to detect subtle phenotypic changes in different physiological functions.

As pointed out in previous literature, computational inference tends to have a higher false-positive rate of identifying protein function-disrupting missense alleles (3). This study demonstrated that predictions based solely on sequence conservation did not effectively distinguish missense mutations that cause phenotypic changes from ones that exhibit no observable phenotype. Only 70% of our behavioral results agreed with the predictions from two commonly used computational programs, PolyPhen-2 and SIFT. Most of the discrepancies are false-positive predictions. Absence of a phenotype *in vivo* may occur due to

genetic redundancy and robust gene networks compensating for the inhibition of a single component, especially in tightly regulated cellular networks involved in signaling, metabolic, and transcriptional pathways (39). Or, we did not observe every possible phenotype. More pointedly, our study validated two phenotypic changing missense alleles predicted by PolyPhen-2 as benign. The fecundity defect found in *dlg-1(V964I)* can be recapitulated by RNAi, whereas the *tph-1(R259Q)* displayed hypomorphic phenotypes similar to its null mutant (30). The false negatives predicted by the software indicate a void in current prediction algorithms, suggesting a need for a screening platform in a multicellular model organism such as our own. Our *in vivo* screening platform not only selects for genes that display sequence conservation across evolution, it also reflects the complex nature in biological system, such as redundancy and compensation. Our phenotypic results can also provide feedback to improve the accuracy of prediction algorithm.

In contrast with previous studies on the phenotypic consequences of missense mutations, our platform examines gene functions in their endogenous multicellular context. Compared to a previous study using yeast two-hybrid to verify the effects of missense mutations in protein interaction experimentally and computationally (10), our strategy captures the overall readout of mutation effects and intercellular interaction. Furthermore, our

use of endogenous proteins allows all other molecular interactions to remain intact, and thus avoids potential confounding factors, such as intron disruption and isoform imbalance (40, 41). As a result, our approach may more accurately reflect the consequence of a variant than a "humanized" model organism does (42–44), and CRISPR knocking-in a DNA missense template is more efficient. Our high throughput screening strategy occupies an unusual niche in primary screening for the consequence of missense mutations *in vivo*.

Using *C. elegans* as a model for psychiatric disorders has some limitations, including a lack of complicated behaviors and some neurotransmitter systems (e.g., norepinephrine). However, *C. elegans* and humans share essential physiological pathways (e.g., insulin signaling, Ras/Notch signaling, p53, and many miRNAs), neurotransmitter systems, and receptor pharmacology (13, 33). The transparency and easy access genetic tools make *C. elegans* a powerful model for dissecting the mechanisms of pathological conditions and drug target identification. The short generation time of *C. elegans* enables high-throughput screening for numerous targets (such as missense variants) before embarking on less efficient and more costly animal models (33). In addition, with the tissue-specific promoters and conditional knockout techniques available in *C. elegans*, it is possible to decipher the effects of these disease-associated missense mutations spatially and temporally (45–47). For genetic

candidates that show correlated expression, our platform also can be used to investigate the interaction between missense variants by generating double/multiple missense mutations model.

The discovery rate of novel genetic variants associated with human diseases has increased due to technical improvements and decreasing costs of next-generation sequencing. However, it is difficult to assess the impact of single missense mutations due to the complexity of human genetic backgrounds. One solution is to test variants in a model organism with an isogenic background to quickly identify variants producing changes in protein function. Here, we developed an experimental pipeline to investigate the functional consequences of ASD-associated missense variants in *C. elegans*. Our approach will help prioritize consequential missense variants for detailed studies in vertebrate models or human cells. This pipeline will serve as a stepping stone for defining molecular mechanisms in complex human diseases such as ASD.

C.5 Materials and Methods

Mapping locations of human residues to the C. elegans *genome*

ASD missense variants were obtained from the SFARI Gene–Human Gene Module (14) (Table S1). We used the comparative genomics resources provided by Ensembl (release 90), which integrates in-house annotation for nearly 100 vertebrate genomes (e.g., human, mouse, and zebrafish) with reference annotation for selected invertebrate model organisms (e.g., *C. elegans*, with genome and annotation provided by WormBase). Ensembl provides a protein multiple alignment and evolutionary trees for each gene family, and asserts orthology and paralogy relationships between pairs of genes (15). These data were organized with a custom automated pipeline (16) : for a given human genome coordinate, (a) identify which human protein-coding gene (if any) coincided with the provided coordinate; (b) obtain the amino acid coordinates in that protein; (c) check if the human gene has a *C. elegans* ortholog; (d) if so, use the multiple alignment associated with the orthology assertion to identify the orthologous amino acid in the *C. elegans* protein; (e) from the protein coordinates, obtain the corresponding position in the *C. elegans* reference genome.

Strains

The Bristol N2 *C. elegans* strain was used as the wild-type control and background for all CRISPR experiments(12). The control strains for functional assays were obtained from lab stock, the Caenorhabditis Genetics Center (CGC), and the National BioResource Project- *C. elegans* (NBRP). Loss-of-function mutant controls were: JD105 *avr-15(ad1051)*(17), FX17094 *chd-7(tm6139)*, PS3071 *egl-19(n2368sd)*(18), SD464 *mpk-1(ga117)*(19), and PS3156 *tph-1(mg280)*(20). All strains were maintained on nematode growth medium (NGM) agar plates seeded with *Escherichia coli* OP50 at room temperature (20-22 °C).

Generation of missense mutant strains

The Cas9 protein-based CRISPR knock-in protocol was adapted from Paix et al. (21). The sgRNA sequences were selected using the *C. elegans* CRISPR guide RNA tool (22). Single-stranded donor oligonucleotides contained 35 bp of flanking homology on both sides of the mutated region. An online tool for restriction analysis, WatCut, was used to assist in designing restriction sites that did not affect protein sequence. The crRNA, tracrRNA and donor oligonucleotides were commercially synthesized and dissolved in Nuclease Free

Duplex Buffer (Integrated DNA Technologies Inc., Coralville, IA). Purified Cas9 protein was a kind gift from Dr. Tsui-Fen Chou (LA BioMed). gRNA duplexes were generated by mixing crRNA and tracrRNA at 1:1 ratio and incubating at 94°C for 2 minutes. The Cas9 protein (25 μM final concentration) and gRNA duplex (27 μM final concentration) were mixed and incubated at room temperature for 5 minutes before adding donor oligonucleotides (0.6 μM final concentration). To facilitate screening, *dpy-10(cn64)* or *unc-58(e665)* was used as a co-conversion marker and made up part of the crRNA and donor oligo used (23) (Figure C.1B). A crRNA ratio [marker: target gene] of 1:4 and 2:3 was used for *dpy-10* and *unc-58,* respectively. A donor ratio [marker: target gene] of 1:2 was used for both *dpy-10* and *unc-58*.

The F1 offspring displaying the co-conversion phenotype were genotyped as follows: About 5 worms were picked into 10 μL lysis buffer (10 mM Tris, 50 mM KCl, 2 mM MgCl2, pH 8.0) with proteinase K (500 ng/mL; Invitrogen, Carlsbad, CA) and incubated at 65°C for an hour to extract genomic DNA. The genomic prep was amplified in a PCR reaction and then treated with a restriction enzyme (NEB, Ipswich, MA) to check the presence of targeted missense mutation. Mutants with the correct length were confirmed by sequencing (Laragen, Culver City, CA). When available, we saved two independent missense mutant lines. While

there are little to no off-targets effects of Cas9 (24), *C. elegans* N2 suffers approximately one mutation per generation, so it was useful to have more than one strain for each locus.

Fecundity assay

Well-fed *C. elegans* were synchronized at the L4 stage. Individual L4 hermaphrodites were placed on separate NGM plates seeded with OP50, and these animals were subsequently transferred to a new plate every day. The number of newly hatched larvae progeny was counted for every plate one day after the adult was transferred. The total fecundity consisted of the sum of progeny produced for three days per animal.

Locomotion tracking

Well-fed L4 hermaphrodites were picked at about 16 hours before the experiment to provide synchronized young adults. On the day of the experiment, 8 young adults were picked onto NGM plates freshly seeded with a 50-μL drop of a saturation-phase culture of OP50. The worms were given 30 min for habituation and then tracked for 4 min. Strains were tracked

between 1 p.m. and 6 p.m. across several days. WormLab (MBF Bioscience, Williston, VT) equipment and software were used for tracking and analyses. The camera was a Nikon AF Micro 60/2.8D with zoom magnification. A 2456×2052–resolution, 7.5-f.p.s. camera with a magnification that results in 8.2 μm per pixel and an FOV of roughly 2×2 cm^2 was used. Approximately 8-10 plates were tracked per experimental strain. The mean of each plate was first calculated and then the total mean of all plates of the same genotype was computed.

Statistical analysis

The fecundity assay was analyzed using a non-parametric bootstrap analysis (D. Angeles-Albores & P.W.S., unpublished). Initially, the two datasets were mixed, samples were selected at random with replacements from the mixed population into two new datasets, and then the difference in the averages of these new datasets were calculated; this process was iterated 10^6 times. We reported the *p*-value as the probability when the difference in the average of simulated datasets was greater than the difference in the average of the original datasets. If $p < 0.01/(\text{total testing number})$, we rejected the null hypothesis that the average values of the two datasets were not equal to each other. Morphology and locomotion were

analyzed by one-way ANOVA using GraphPad Prism version 6 (GraphPad, La Jolla, CA). Dunnett multiple comparisons were performed between wild-type and mutant strains. The significant level was defined as $p < 0.01$.

Acknowledgements

We thank WormBase for genome information. We thank CGC (funded by NIH Office of Research Infrastructure Programs, P40 OD010440) and NBRP for providing strains. We thank Tsui-Fen Chou for Cas9 protein. We thank Shahla Gharib for the assistance in strain generation. We thank David Angeles-Albores for sharing his data analysis software. We thank Hillel Schwartz and Han Wang for comments on the manuscript. This work was supported by Simons Foundation (SFARI award # 367560) to PWS. KB was supported by National Institute of Health (NIH) pre-doctoral training grant T32GM007616. PWS was an investigator with the Howard Hughes Medical Institute during part of this study.

Conflicts of Interest

The authors have no conflicts of interest to disclose.

Web Resources

SFARI Gene–Human Gene Module, https://gene.sfari.org/database/human-gene/

Ensembl, www.ensembl.org

WormBase, www.wormbase.org

C. elegans CRISPR guide RNA tool, http://genome.sfu.ca/crispr/

WatCut, http://watcut.uwaterloo.ca

PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/

SIFT, http://sift.bii.a-star.edu.sg

C.6 References

1. Iossifov,I., Ronemus,M., Levy,D., Wang,Z., Hakker,I., Rosenbaum,J., Yamrom,B., Lee,Y.H., Narzisi,G., Leotta,A., *et al.* (2012) De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron*, **74**, 285–299.

2. Iossifov,I., O'Roak,B.J., Sanders,S.J., Ronemus,M., Krumm,N., Levy,D., Stessman,H.A., Witherspoon,K.T., Vives,L., Patterson,K.E., *et al.* (2014) The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, **515**, 216–221.

3. Andrews,T.D., Sjollema,G. and Goodnow,C.C. (2013) Understanding the immunological impact of the human mutation explosion. *Trends Immunol.*, **34**, 99–106.

4. Robinson,E.B., Samocha,K.E., Kosmicki,J.A., McGrath,L., Neale,B.M., Perlis,R.H. and Daly,M.J. (2014) Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 15161–15165.

5. Turner,T.N., Sharma,K., Oh,E.C., Liu,Y.P., Collins,R.L., Sosa,M.X., Auer,D.R., Brand,H., Sanders,S.J., Moreno-De-Luca,D., *et al.* (2015) Loss of delta catenin function in severe autism. *Nature*, **520**, 51–56.

6. Pilorge,M., Fassier,C., Le Corronc,H., Potey,A., Bai,J., De Gois,S., Delaby,E., Assouline,B., Guinchat,V., Devillard,F., *et al.* (2016) Genetic and functional analyses demonstrate a role for abnormal glycinergic signaling in autism. *Mol. Psychiatry*, **21**, 936–945.

7. Alfoldi,J. and Lindblad-toh,K. (2013) Comparative genomics as a tool to understand evolution and disease. *Genome Res.*, **23**, 1063–1068.

8. Tennessen,J.A., Bigham,A.W., O'Connor,T.D., Fu,W., Kenny,E.E., Gravel,S., McGee,S., Do,R., Liu,X., Jun,G., *et al.* (2012) Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science (80-.).*, **337**, 64–69.

9. Miosge,L.A., Field,M.A., Sontani,Y., Cho,V., Johnson,S., Palkova,A., Balakishnan,B., Liang,R., Zhang,Y., Lyon,S., *et al.* (2015) Comparison of predicted and actual consequences of missense mutations. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, E5189–E5198.

10. Chen,S., Fragoza,R., Klei,L., Liu,Y., Wang,J., Roeder,K., Devlin,B. and Yu,H. (2018) An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. *Nat. Genet.*, **50**, 1032–1040.

11. Kim,S., Twigg,S.R.F., Scanlon,V.A., Chandra,A., Hansen,T.J., Alsubait,A., Fenwick,A.L., McGowan,S.J., Lord,H., Lester,T., *et al.* (2017) Localized TWIST1 and TWIST2 basic domain substitutions cause four distinct human diseases that can be modeled in Caenorhabditis elegans. *Hum. Mol. Genet.*, **26**, 2118–2132.

12. Brenner,S. (1974) The genetics of Caenorhabditis elegans. *Genetics*, **77**, 71–94.

13. Engleman,E.A., Katner,S.N. and Neal-beliveau,B.S. (2016) Caenorhabditis elegans as a Model to Study the Molecular and Genetic Mechanisms of Drug Addiction. *Prog Mol Biol Transl Sci*, **137**, 229–252.

14. Fischbach,G.D. and Lord,C. (2010) The simons simplex collection: A resource for identification of autism genetic risk factors. *Neuron*, **68**, 192–195.

15. Herrero,J., Muffato,M., Beal,K., Fitzgerald,S., Gordon,L., Pignatelli,M., Vilella,A.J., Searle,S.M.J., Amode,R., Brent,S., *et al.* (2016) Ensembl comparative genomics resources. *Database*, **2016**, bav096.

16. Yates,A., Beal,K., Keenan,S., McLaren,W., Pignatelli,M., Ritchie,G.R.S., Ruffier,M., Taylor,K., Vullo,A. and Flicek,P. (2015) The Ensembl REST API: Ensembl Data for Any Language. *Bioinformatics*, **31**, 143–145.

17. Dent,J.A., Davis,M.W. and Avery,L. (1997) avr-15 encodes a chloride channel subunit that mediates inhibitory glutamatergic neurotransmission and ivermectin sensitivity in Caenorhabditis elegans. *EMBO J.*, **16**, 5867–5879.

18. Lee,R.Y., Lobel,L., Hengartner,M., Horvitz,H.R. and Avery,L. (1997) Mutations in the alpha1 subunit of an L-type voltage-activated Ca2+ channel cause myotonia in Caenorhabditis elegans. *Embo J*, **16**, 6066–6076.

19. Lackner,M.R., Kornfeld,K., Miller,L.M., Robert Horvitz,H. and Kim,S.K. (1994) A MAP kinase homolog, mpk-1, is involved in ras-mediated induction of vulval cell fates in Caenorhabditis elegans. *Genes Dev.*, **8**, 160–173.

20. Sze,J.Y., Victor,M., Loer,C., Shi,Y. and Ruvkun,G. (2000) Food and metabolic signalling defects in a Caenorhabditis elegans serotonin-synthesis mutant. *Nature*, **403**, 560–564.

21. Paix,A., Folkmann,A., Rasoloson,D. and Seydoux,G. (2015) High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR/Cas9 Ribonucleoprotein Complexes. *Genetics*, **201**, 47–54.

22. Au,V., Li-Leger,E., Raymant,G., Flibotte,S., Chen,G., Martin,K., Fernando,L., Doell,C., Rosell,F., Wang,S., *et al.* (2018) Optimizing guide RNA selection and CRISPR/Cas9

methodology for efficient generation of deletions in C. elegans. *bioRxiv*, 10.1101/359588.

23. Arribere,J., Bell,R., Fu,B., Artiles,K., Hartman,P. and Fire,A. (2014) Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in Caenorhabditis elegans. *Genetics*, **198**, 837–846.

24. Chiu,H., Schwartz,H.T., Antoshechkin,I. and Sternberg,P.W. (2013) Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. *Genetics*, **195**, 1167–1171.

25. Geisheker,M.R., Heymann,G., Wang,T., Coe,B.P., Turner,T.N., Stessman,H.A.F., Hoekzema,K., Kvarnung,M., Shaw,M., Friend,K., *et al.* (2017) Hotspots of missense mutation identify novel neurodevelopmental disorder genes and functional domains. *Nat. Neurosci.*, **20**, 1043–1051.

26. Krumm,N., O'Roak,B.J., Shendure,J. and Eichler,E.E. (2014) A de novo convergence of autism genetics and molecular neuroscience. *Trends Neurosci.*, **37**, 95–105.

27. Bono,M. de and Villu Maricq,A. (2005) Neuronal substrates of complex behaviors in C. elegans. *Annu. Rev. Neurosci.*, **28**, 451–501.

28. Cronin,C.J., Mendel,J.E., Mukhtar,S., Kim,Y.M., Stirbl,R.C., Bruck,J. and Sternberg,P.W. (2005) An automated system for measuring parameters of nematode sinusoidal movement. *BMC Genet.*, **6**, 1–19.

29. Fournier,K.A., Hass,C.J., Naik,S.K., Lodha,N. and Cauraugh,J.H. (2010) Motor coordination in autism spectrum disorders: A synthesis and meta-analysis. *J. Autism Dev. Disord.*, **40**, 1227–1240.

30. Pilipiuk,J., Lefebvre,C., Wiesenfahrt,T., Legouis,R. and Bossinger,O. (2009) Increased IP3/Ca2+ signaling compensates depletion of LET-413/DLG-1 in C. elegans epithelial junction assembly. *Dev. Biol.*, **327**, 34–47.

31. Ng,P.C. and Henikoff,S. (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.*, **31**, 3812–3814.

32. Adzhubei,I.A., Schmidt,S., Peshkin,L., Ramensky,V.E., Gerasimova,A., Bork,P., Kondrashov,A.S. and Sunyaev,S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.

33. Markaki,M. and Tavernarakis,N. (2010) Modeling human diseases in Caenorhabditis elegans. *Biotechnol. J.*, **5**, 1261–1276.

34. Yemini,E., Jucikas,T., Grundy,L.J., Brown,A.E.X. and Schafer,W.R. (2013) A database of Caenorhabditis elegans behavioral phenotypes. *Nat. Methods*, **10**, 877–879.

35. Sonnichsen,B., Koski,L.B., Walsh,A., Marschall,P., Neumann,B., Brehm,M., Alleaume,A.-M.M., Artelt,J., Bettencourt,P., Cassin,E., *et al.* (2005) Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature*, **434**, 462–469.

36. Maeda,I., Kohara,Y., Yamamoto,M. and Sugimoto,A. (2001) Large-scale analysis of gene function in Caenorhabditis elegans by high-throughput RNAi. *Curr. Biol.*, **11**, 171–176.

37. Cook,A., Aptel,N., Portillo,V., Siney,E., Sihota,R., Holden-Dye,L. and Wolstenholme,A. (2006) Caenorhabditis elegans ivermectin receptors regulate locomotor behaviour and are functional orthologues of Haemonchus contortus receptors. *Mol. Biochem. Parasitol.*, **147**, 118–125.

38. Rual,J.-F., Ceron,J., Koreth,J., Hao,T., Nicot,A., Hirozane-kishikawa,T., Vandenhaute,J., Orkin,S.H., Hill,D.E. and Vidal,M. (2004) Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome Res.*, **14**, 2162–2168.

39. El-Brolosy,M.A. and Stainier,D.Y.R. (2017) Genetic compensation : A phenomenon in search of mechanisms. *PLoS Genet.*, **13**, e1006780.

40. Reble,E., Dineen,A. and Barr,C.L. (2018) The contribution of alternative splicing to genetic risk for psychiatric disorders. *Genes, Brain Behav.*, **17**, 1–12.

41. Robison,A.J. (2014) Emerging role of CaMKII in neuropsychiatric disease. *Trends Neurosci.*, **37**, 653–662.

42. McDiarmid,T.A., Au,V., Loewen,A.D., Liang,J., Mizumoto,K., Moerman,D.G. and Rankin,C.H. (2018) CRISPR-Cas9 human gene replacement and phenomic characterization in *Caenorhabditis elegans* to understand the functional conservation of human genes and decipher variants of uncertain significance. *Dis. Model. Mech.*, 10.1242/dmm.036517.

43. Baruah,P.S., Beauchemin,M., Parker,J.A. and Bertrand,R. (2017) Expression of human Bcl-xL (Ser49) and (Ser62) mutants in Caenorhabditis elegans causes germline defects and aneuploidy. *PLoS One*, **12**, e0177413.

44. Walsh,N., Kenney,L., Jangalwe,S., Aryee,K., Greiner,D.L., Brehm,M.A., Shultz,L.D. and Harbor,B. (2017) Humanized mouse models of clinical disease. *Annu. Rev. Pathol.*, **24**, 187–215.

45. Shen,Z., Zhang,X., Chai,Y., Zhu,Z., Yi,P., Feng,G., Li,W. and Ou,G. (2014) Conditional knockouts generated by engineered CRISPR-Cas9 endonuclease reveal the roles of coronin in c.elegans neural development. *Dev. Cell*, **30**, 625–636.

46. Hubbard,E.J.A. (2014) FLP/FRT and Cre/lox recombination technology in C. elegans. *Methods*, **68**, 417–424.

47. Voutev,R. and Hubbard,E.J.A. (2008) A 'FLP-out' system for controlled gene expression in Caenorhabditis elegans. *Genetics*, **180**, 103–119.

C. 7 Figures and Tables

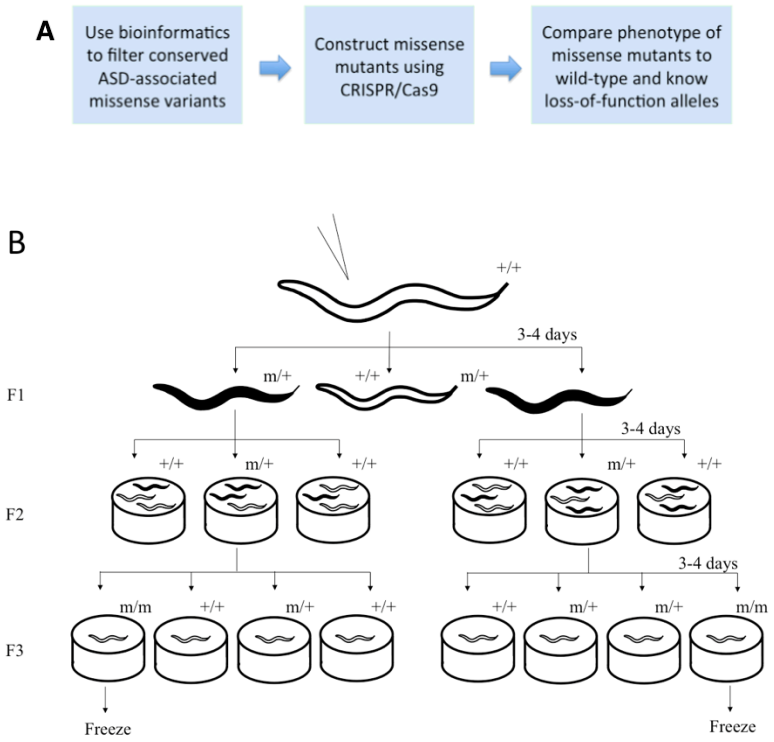


Figure C.1. Generation of ASD-associated missense variants in *C. elegans*. (A) Experimental pipeline. First, we used bioinformatics to identify human ASD-associated missense variants conserved in *C. elegans* protein sequences. We generated the *C. elegans* missense mutants using CRISPR/Cas9 and homology-directed knock-in technique. We then analyzed the effects of these autism-associated alleles by comparing observable phenotypes from these alleles to both wildtype and known loss-of-function mutant controls. (B) Mutant strain screening process. Three days after injection (P0), F1 offspring expressing co-conversion marker, such as dumpy (black), were selected into individual plates. We identified the heterozygous plates (F2) via

genotype. For each heterozygous plate containing the successful knock-in target missense residue, we randomly selected 16-20 F2 offspring and separated them into individual plates (F3). A second round of genotyping process was then conducted to identify homozygous plates, which were cryo-preserved for future studies.

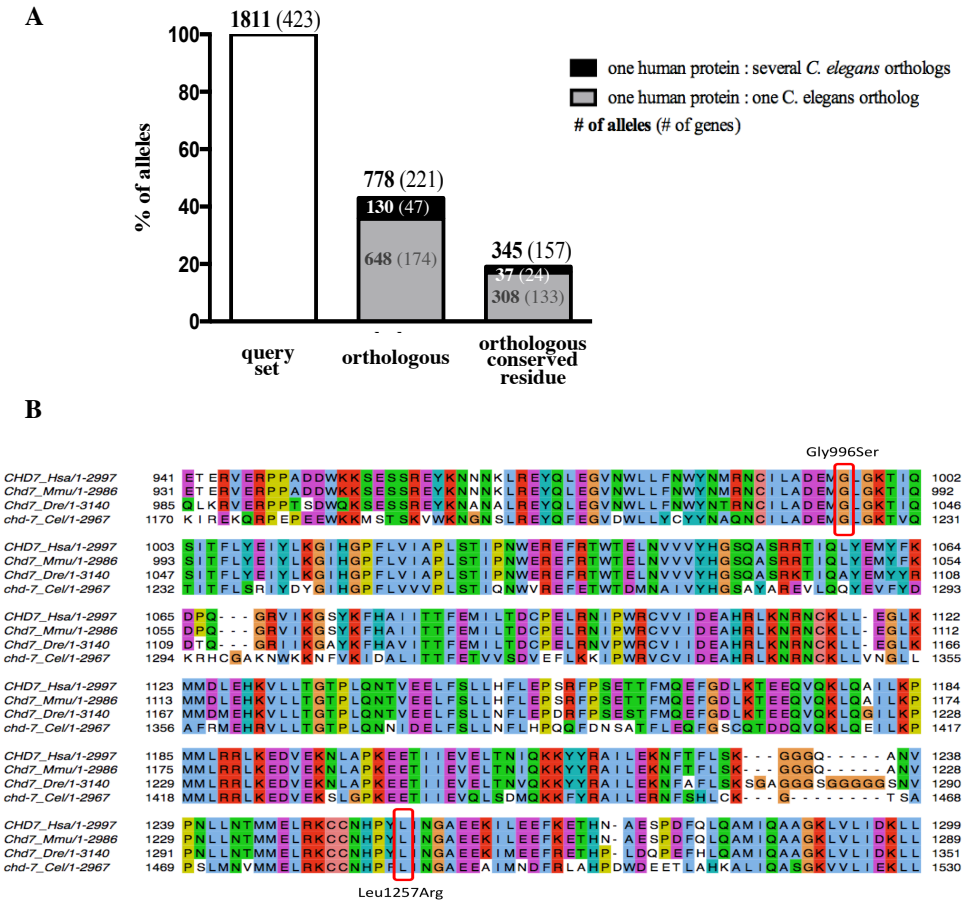


Figure C.2. Detection of conserved ASD-associated missense residues in *C. elegans*. (A) For each missense allele, the *C. elegans* ortholog of the corresponding human gene was identified using the Ensembl Compara method. Each Ensembl ortholog pair is underpinned by a protein multiple-sequence alignment which can be used to identify the putative orthologous *C. elegans* amino acid for a given human amino acid. 43% of the human missense variants have at least one *C. elegans* ortholog. 130 of the 778 orthologous missense variants have more than one *C. elegans*

ortholog (black: 1-many). Overall, only 19% of the missense residues are orthologous and conserved in *C. elegans*. 37 of the 345 conserved residues have more than one orthologs in *C. elegans*. (B) Section of the protein multiple alignment for human (*Hsa*) *CHD7* and its orthologs in mouse (*Mmu*), zebrafish (*Dre*), and *C. elegans (Cel)*. Residues in the alignment have been colored by JalView (PMID:19151095) using the Clustal X coloring scheme. Circled are two ASD-linked missense variants in *CHD7* that occur in a highly conserved region across all the species.

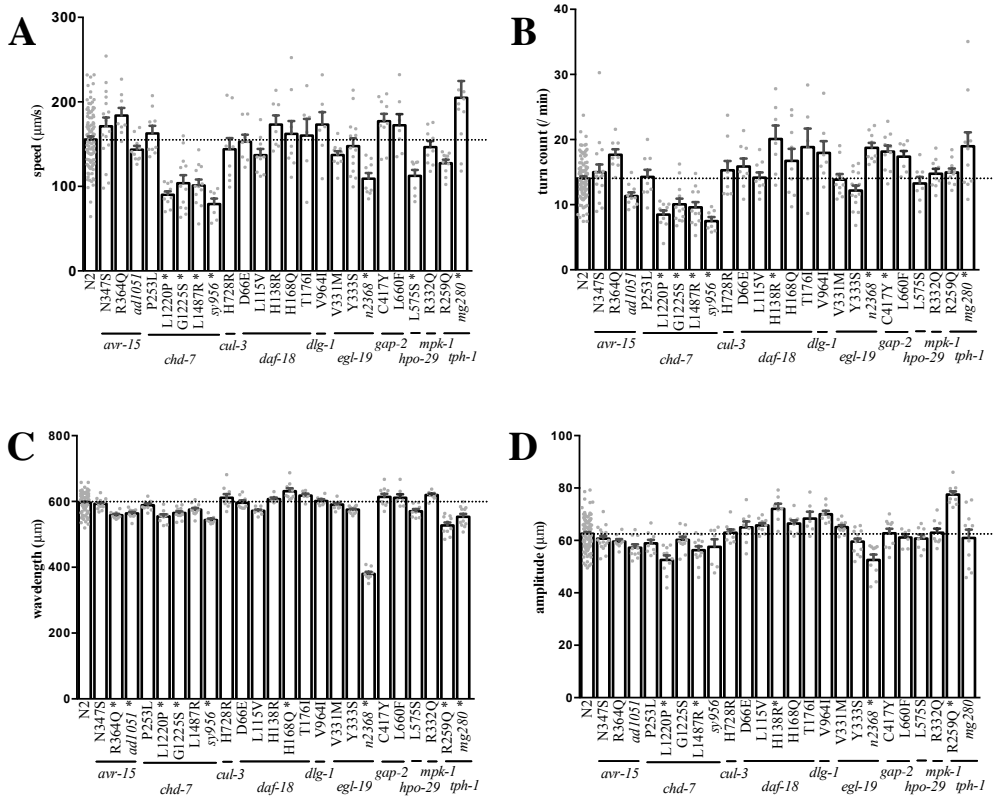


Figure C.3. Locomotion phenotypes of missense alleles and their controls. Speed (A) and reversal rate (B) are measurements of locomotion while wavelength (C) and amplitude (D) represent sinusoidal shape of movement. *avr-15(R364Q)* mutant shows decreased wavelength. Most *chd-7* missense mutants display a weaker version of the null mutant phenotypes in all measurements. *daf-18(H138R)* and *daf-18(H168Q)* exhibit larger sinusoidal wavelength and amplitude. *gap-2(C417Y)* shows higher reversal rate. *hpo-29(L575S)* displays slower speed. *tph-1(R259Q)* exhibits changes in sinusoidal movement. Bars are presented as the mean ± SEM. Each dot represents the

average of one plate containing 8-10 worms. *: $p < 0.01$ via one-way ANOVA and multiple comparison to wild type. Horizontal line indicates the mean of wild type.

Figure C.4. Fecundity phenotype of missense alleles and their controls. Most *chd-7* mutants displayed decreased fecundity. The defects were more subtle than its deletion*(tm6139)* or frameshift*(sy956)* null controls. *cul-3(H728R)* and *dlg-1(V964I)* also show decrease in fecundity. Each dot represents total number of living larvae from one animal. Approximately 20 animals were tested in each strain. Wild type

(N2) and its median value (dotted vertical line) was shown in blue. Mutants significantly different (sig) from wild type were shown in red. $p < 0.01$/(total test number) via non-parametric bootstrap analysis.

Table C.1. Strain information. 20 ASD-associated missense alleles and their corresponding mutant strains in *C. elegans*.

Human gene	Human cDNA change*	Human protein change	Inheritance pattern	*C. elegans* gene (allele)	*C. elegans* protein change	Strain name
CACNA1D	c.1105G>A	V369M	unknown	*egl-19(sy849)*	V331M	PS7085
CACNA1D	c.1112A>C	Y371S	unknown	*egl-19(sy850)*	Y333S	PS7156
CHD7	c.2986G>A	G996S	*de novo*	*chd-7(sy861)*	G1225S	PS7293
CHD7	c.3770T>G	L1257R	*de novo*	*chd-7(sy855)*	L1487R	PS7317
CHD8	c.2501T>C	L834P	*de novo*	*chd-7(sy859)*	L1220P	PS7318
CHD8	c.494C>T	P165L	unknown	*chd-7(sy1049)*	P253L	PS7267
CUL3	c.2156A>G	H719R	de novo	*cul-3(sy874)*	H728R	PS7387
DLG4	c.2281G>A	V761I	unknown	*dlg-1(sy872)*	V964I	PS7343
GLRA2	c.407A>G	N136S	*de novo*	*avr-15(sy873)*	N347S	PS7384
GLRA2	c.458G>A	R153Q	*de novo*	*avr-15(sy851)*	R364Q	PS7257
MAPK3	c.833G>A	R278Q	*de novo*	*mpk-1(sy870)*	R332Q	PS7382
NAA15	c.1319T>C	L440S	familial	*hpo-29(sy877)*	L575S	PS7394
PTEN	c.66C>G	D22E	familial	*daf-18(sy879)*	D66E	PS7439
PTEN	c.208C>G	L70V	unknown	*daf-18(sy887)*	L115V	PS7432
PTEN	c.278A>G	H93R	*de novo*	*daf-18(sy881)*	H138R	PS7436
PTEN	c.369C>G	H123Q	unknown	*daf-18(sy885)*	H168Q	PS7430
PTEN	c.392C>T	T131I	*de novo*	*daf-18(sy882)*	T176I	PS7434
SYNGAP1	c.698G>A	C233Y	*de novo*	*gap-2 (sy889)*	C417Y	PS7433
SYNGAP1	c.1288C>T	L430F	familial	*gap-2(sy886)*	L660F	PS7457
TPH2	c.674G>A	R225Q	familial	*tph-1(sy878)*	R259Q	PS7395

Table C.2. Morphological phenotypes of missense alleles and their controls

Gene	Length (µm)	Width (µm)	Area (µm^2)
N2	1105 ± 5	87.8 ± 0.7	98798 ± 1147
avr-15(N347S)	1101 ± 7	87.9 ± 1.5	98214 ± 1273
avr-15(R364Q)	989 ± 7 *	77.3 ± 0.9 *	77860 ± 1457 *
avr-15(ad1051)	1033 ± 10 *	78.5 ± 0.8 *	82468 ± 1410 *
chd-7(P253L)	1110 ± 9	78.9 ± 0.9 *	86572 ± 1815 *
chd-7(L1220P)	978 ± 10 *	76.2 ± 1.0 *	75909 ± 1432 *
chd-7(G1225S)	1038 ± 11 *	80.8 ± 1.3 *	85400 ± 1977 *
chd-7(L1487R)	1033 ± 13 *	81.7 ± 1.3 *	86060 ± 2277 *
chd-7(sy956)	954 ± 6 *	75.4 ± 0.9 *	73201 ± 1254 *
cul-3(H728R)	1111 ± 9	96.1 ± 2.8 *	108771 ± 3865 *
daf-18(D66E)	1125 ± 14	91.5 ± 1.1	104804 ± 2496
daf-18(L115V)	1128 ± 9	94.2 ± 1.1	108208 ± 1902
daf-18(H138R)	1130 ± 6	94.6 ± 1.2	108706 ± 1366
daf-18(H168Q)	1148 ± 6	103.3 ± 2.7 *	120936 ± 3535 *
daf-18(T176I)	1135 ± 9	93.1 ± 1.4	107395 ± 1711
dlg-1(V964I)	1129 ± 4	83.16 ± 0.9	95502 ± 1210
egl-19(V331M)	1082 ± 4	83.1 ± 1.1	91504 ± 1448
egl-19(Y333S)	1052 ± 8 *	81.2 ± 0.7 *	86896 ± 1310 *
egl-19(n2368sd)	639 ± 10 *	68.1 ± 0.8 *	44402 ± 1104 *
gap-2(C417Y)	1128 ± 14	98.7 ± 2.6 *	113351 ± 3713 *
gap-2(L660F)	1120 ± 13	89.0 ± 1.7	101634 ± 3020
hpo-29(L575S)	1044 ± 12 *	94.0 ± 2.4	99849 ± 3111
mpk-1(R332Q)	1145 ± 7	88.0 ± 1.7	104188 ± 1889
tph-1(R259Q)	1026 ± 9 *	84.5 ± 1.1	88238 ± 1151 *
tph-1(mg280)	993 ± 20 *	80.9 ± 1.3 *	81918 ± 2765 *

All values are presented as mean ± SEM.

*$p < 0.01$ via one-way ANOVA and multiple comparison.

Table C.3. Comparison of behavioral results to software prediction

Human gene	*C. elegans* gene	PolyPhen-2[a]	(1-SIFT)[a]	Phenotype
CACNA1D(V369M)	*egl-19(V331M)*	0.995	0.99	no
CACNA1D(Y371S)	*egl-19(Y333S)*	1	1	morphology changes
CHD7(G996S)	*chd-7(G1225S)*	0.998	1	morphology changes, locomotion variants, reduced fecundity
CHD7(L1257R)	*chd-7(L1487R)*	1	1	morphology changes, locomotion variants, reduced fecundity
CHD8(L834P)	*chd-7(L1220P)*	1	1	morphology changes, locomotion variants, reduced fecundity
CHD8(P165L)	*chd-7(P253L)*	0.996	0.86	morphology changes
CUL3(H719R)	*cul-3(H728R)*	1	0.96	morphology changes, reduced fecundity
DLG4(V761I)	*dlg-1(V964I)*	0.001	0.84	reduced fecundity
GLRA2(N136S)	*avr-15(N347S)*	0.979	1	locomotion variants
GLRA2(R153Q)	*avr-15(R364Q)*	0.997	1	locomotion variants
MAPK3(R278Q)	*mpk-1(R332Q)*	0.997	1	no
NAA15(L440S)	*hpo-29(L575S)*	0.999	0.96	morphology changes, locomotion variants
PTEN(D22E)	*daf-18(D66E)*	0.297	0.93	no
PTEN(L70V)	*daf-18(L115V)*	0.999	1	no
PTEN(H93R)	*daf-18(H138R)*	1	0.97	locomotion variants
PTEN(H123Q)	*daf-18(H168Q)*	1	1	morphology changes, locomotion variants
PTEN(T131I)	*daf-18(T176I)*	1	0.82	no
SYNGAP1(C233Y)	*gap-2(C417Y)*	0.94	1	no
SYNGAP1(L430F)	*gap-2(L660F)*	1	1	morphology changes, locomotion variants
TPH2(R225Q)	*tph-1(R259Q)*	0.162	0.92	morphology changes, locomotion variants

[a] PolyPhen-2 and SIFT prediction score were based on human sequence (*: 1=probably damaging)

Supplemental Tables

Table S1. The list of 347 missense alleles that have *C. elegans* orthologs

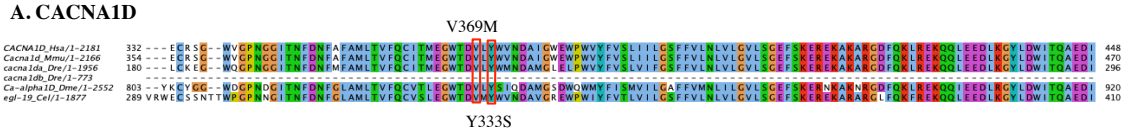
Human_geno	Human_geno	Humam_genc	Human_gene	cDNA_chang	Protein_change	Relationship	Human_gene	Human_transi	Human_CDS	Human_CDS	Human_prote	Human_dom	Human_aa_5(	Human_aa	Human_aa_5(	Worm_aa_50	Worm_aa	Worm_
19	1045110	1045110	ABCA7	c.1324G>A	p.Gly442Arg	ortholog_one2	ENSG000000	ENST000002	1324	1324	442	IPR026082;IF	DVGHLVGT	G	PEDSSDPTE	E------IANKY	I	DS---N/
19	1046318	1046318	ABCA7	c.1534C>G	p.Arg512Gly	ortholog_one2	ENSG000000	ENST000002	1534	1534	512	IPR026082;IF	IKIRMDIDV	R	AAVRVLSG	YKIRHYPEN	R	AIMSEI
19	1050997	1050997	ABCA7	c.2629G>A	p.Ala877Thr	ortholog_one2	ENSG000000	ENST000002	2629	2629	877	IPR003593;IF	LSLDFYQGI	A	AIRPHLGVC	LNLRLYEGI	R	RVRDV
19	1051482	1051482	ABCA7	c.2858C>A	p.Ala953Asp	ortholog_one2	ENSG000000	ENST000002	2858	2858	953	IPR003593;IF	WFYGRLKC	A	FVGGSQVV	RLFAALKG	A	FIGGSF
19	1057344	1057344	ABCA7	c.4795G>A	p.Val1599Met	ortholog_one2	ENSG000000	ENST000002	4795	4795	1599	IPR026082;IF	VPASFTLVL	V	LIFLAFQQR	PAGFTVYL	M	LIYLCI
19	1058884	1058884	ABCA7	c.5344C>T	p.Arg1782Trp	ortholog_one2	ENSG000000	ENST000002	5344	5344	1782	IPR026082;IF	LFTLLLQH-	R	VV----QGAT	IFSQMDEFC	R	VDAIP
19	1059057	1059057	ABCA7	c.5435G>A	p.Arg1812His	ortholog_one2	ENSG000000	ENST000002	5435	5435	1812	IPR026082;IF	-GEEDEDVA	R	LCLGIPPGE	C--DDEDVV	G	ISFAVI
19	1062249	1062249	ABCA7	c.5648C>T	p.Thr1883Met	ortholog_one2	ENSG000000	ENST000002	5648	5648	1883	IPR003593;IF	FRMVTGDT	T	GREHLELL/	FAMLTAKIF	S	TRENL
19	1063614	1063614	ABCA7	c.5783G>A	p.Gly1928Glu	ortholog_one2	ENSG000000	ENST000002	5783	5783	1928	IPR003593;IF	IFELLTGRE	G	GNKRKLAT	LNMKLSTR	G	GNRRK
19	1065306	1065306	ABCA7	c.6322G>A	p.Glu2108Lys	ortholog_one2	ENSG000000	ENST000002	6322	6322	2108	IPR026082;IF	PPGGRCAL/	E	DTEE--------(	EQ-GTASVA	S	EEGHH
19	1043789	1043789	ABCA7	c.995G>A	p.Gly332Glu	ortholog_one2	ENSG000000	ENST000002	995	995	332	IPR026082;IF	LLWRRLKP	G	PRIFTFMND	-----VREALS	A	LEFQD
7	100891249	100891249	ACHE	c.1643C>T	p.Pro548Leu	ortholog_man	ENSG000000	ENST000003	1643	1643	548	IPR029058;IF	KIFAQRLMI	P	L-----EVRRC	RELSNRFMI	S	SYPSM
7	100891249	100891249	ACHE	c.1643C>T	p.Pro548Leu	ortholog_man	ENSG000000	ENST000003	1643	1643	548	IPR029058;IF	EKIFAQRLM	P	L-EVRRGLR	EQIFSRKIM	S	NIRPIK
7	100891249	100891249	ACHE	c.1643C>T	p.Pro548Leu	ortholog_man	ENSG000000	ENST000003	1643	1643	548	IPR029058;IF	FAQRLMRY	P	L-EVRRGLR	VSEKVIDFV	S	-FRMIC
7	100891249	100891249	ACHE	c.1643C>T	p.Pro548Leu	ortholog_man	ENSG000000	ENST000003	1643	1643	548	IPR029058;IF	KIFAQRLMI	P	L----EV-------	KRLSSIIMQ	T	LQGKI
7	100893782	100893782	ACHE	c.451G>A	p.Gly151Arg	ortholog_man	ENSG000000	ENST000003	451	451	151	IPR029058;IF	QYVDTLYP	G	GGFYSGASS	QSEDTYFGI	G	GGFWS
7	100893782	100893782	ACHE	c.451G>A	p.Gly151Arg	ortholog_man	ENSG000000	ENST000003	451	451	151	IPR029058;IF	LYPGFEGTE	G	GGFYSGASS	YNTSFWGS	G	GGFYS
7	100893782	100893782	ACHE	c.451G>A	p.Gly151Arg	ortholog_man	ENSG000000	ENST000003	451	451	151	IPR029058;IF	LYPGFEGTE	G	GGFYSGASS	YDPTFWGS	G	GGFWY
7	100893782	100893782	ACHE	c.451G>A	p.Gly151Arg	ortholog_man	ENSG000000	ENST000003	451	451	151	IPR029058;IF	TLYPGFEGT	G	GGFYSGASS	SAFPQFPGA	G	GGFFS
7	100893445	100893445	ACHE	c.788C>T	p.Pro263Leu	ortholog_man	ENSG000000	ENST000003	788	788	263	IPR029058;IF	WVQENVA/	P	NGPWATVC	WVHKNIDL	A	TSPWA
7	100893445	100893445	ACHE	c.788C>T	p.Pro263Leu	ortholog_man	ENSG000000	ENST000003	788	788	263	IPR029058;IF	WVQENVA/	P	NGPWATVC	WVRQNIVS	L	ENTW/
7	100893445	100893445	ACHE	c.788C>T	p.Pro263Leu	ortholog_man	ENSG000000	ENST000003	788	788	263	IPR029058;IF	WVQENVA/	P	NGPWATVC	WIRDHIFSF	L	DNKW
7	100893445	100893445	ACHE	c.788C>T	p.Pro263Leu	ortholog_man	ENSG000000	ENST000003	788	788	263	IPR029058;IF	VQENVAAF	P	NGPWATVC	VHENIGAFC	I	INSWA
19	38717957	38717957	ACTN4	c.1174C>G	p.Gln392Glu	ortholog_one2	ENSG000001	ENST000002	1174	1174	392	IPR002017;IF	HKPPKVQE	Q	AEKGYEEW	EKPPRIEDK	D	SEKGF
19	38724045	38724045	ACTN4	c.1660A>G	p.Met554Val	ortholog_one2	ENSG000001	ENST000002	1660	1660	554	IPR002017;IF	ALGSLTHSF	M	FIVHTTEEIE	RLGQLSSKF	L	VIVHEI
3	51988821	51988821	ACY1	c.1057C>T	p.Arg353Cys	ortholog_man	ENSG000002	ENST000006	1057	1057	353	IPR002933;IF	LEFAQKWM	R	AVGVPALG	YEFSKHSDI	R	EAGVR
3	51988821	51988821	ACY1	c.1057C>T	p.Arg353Cys	ortholog_man	ENSG000002	ENST000006	1057	1057	353	IPR002933;IF	LEFAQKWM	R	AVGVPALG	YEFSQYSDC	R	YEGVN
3	51988821	51988821	ACY1	c.1057C>T	p.Arg353Cys	ortholog_man	ENSG000002	ENST000006	1057	1057	353	IPR002933;IF	---------VTP	R	AVGVPALG	IDFETFSNF	R	AQGIR
3	51988821	51988821	ACY1	c.1057C>T	p.Arg353Cys	ortholog_man	ENSG000002	ENST000006	1057	1057	353	IPR002933;IF	LEFAQKWM	R	AVGVPALG	YEFSQHSDY	R	KAGVF
1	111503087	111503087	ADORA3	c.268C>G	p.Leu90Val	ortholog_one2	ENSG000002	ENST000002	268	268	90	IPR001634;IF	PSLQTTTFY	L	LIFTHASIM	KHVRTVTN	L	CILCTI
1	111500396	111500396	ADORA3	c.511G>A	p.Val171Ile	ortholog_one2	ENSG000002	ENST000002	511	511	171	IPR001634;IF	LGL--CWLV	V	MRMDYMV	FLLTTSWIL	V	VDYRY
22	40349920	40349920	ADSL	c.242A>C	p.Glu80Asp	ortholog_one2	ENSG000002	ENST000006	242	242	80	IPR024083;IF	RYKF-RTWI	E	KRLRHDVM	TNKAT-LWI	E	RKLKH
22	40349941	40349941	ADSL	c.263T>G	p.Asp87Glu	ortholog_one2	ENSG000002	ENST000006	263	263	87	IPR024083;IF	WRQLWLW	D	MAHVHTFC	WRQLWIWI	D	MAHN
2	236123941	236123941	AGAP1	c.2234G>A	p.Arg745Gly	ortholog_one2	ENSG000001	ENST000003	2393	2393	798	IPR020683;IF	ILLLAHG-SI	R	DAHGNTAL	NVLLANGM	L	DNNGF
2	236120386	236120386	AGAP1	c.2309G>A	p.Arg770His	ortholog_one2	ENSG000001	ENST000003	2309	2309	770	IPR020683;IF	FLAPLPCTE	R	TALHLACR	FLKPIASGE	R	TVLHL
2	235750389	235750389	AGAP1	c.574G>A	p.Asp192Asn	ortholog_one2	ENSG000001	ENST000003	574	574	192	IPR001806;IF	FVFSLEDEI	D	ARARKLSN	FVFNVCSIQ	E	DEGRQ
2	235799383	235799383	AGAP1	c.818A>G	p.Ser83Gly	ortholog_one2	ENSG000001	ENST000003	818	818	273	NONE	PC---KSLPN	G	----------	AAASSOAR	R	RNTVH

Table S2. Locomotion data

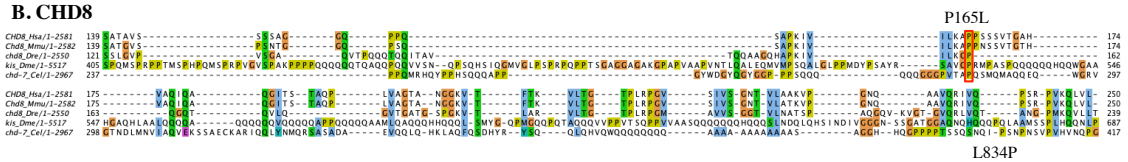
Gene	Speed (μm/s)	Wavelength (μm)	Amplitude (μm)	Reversal rate (min^{-1})
N2	155.6 ± 3.4	596.5 ± 2.7	62.5 ± 0.6	14.0 ± 0.3
avr-15(N347S)	171.2 ± 10.2	592.9 ± 3.5	60.7 ± 1.1	15.0 ± 1.2
avr-15(R364Q)	183.8 ± 9.0	558.3 ± 3.9 *	59.6 ± 0.9	17.7 ± 0.8
avr-15(ad1051)	143.4 ± 4.6	563.7 ± 4.5 *	57.2 ± 1.3	11.3 ± 0.6
chd-7(P253L)	162.7 ± 8.9	588.6 ± 5.3	58.9 ± 1.4	14.3 ± 1.1
chd-7(L1220P)	90.07 ± 4.1 *	554.7 ± 6.1 *	52.6 ± 1.7 *	8.5 ± 0.6 *
chd-7(G1225S)	103.8 ± 9.6 *	565.0 ± 4.7 *	60.2 ± 1.2	10.1 ± 0.9 *
chd-7(L1487R)	101.2 ± 7.0 *	575.4 ± 4.9	56.3 ± 1.5 *	9.6 ± 0.8 *
chd-7(sy956)	79.2 ± 6.4 *	543.6 ± 4.4 *	57.6 ± 2.9	7.5 ± 0.6 *
cul-3(H728R)	143.9 ± 13.2	611.3 ± 11.3	63.0 ± 1.2	15.3 ± 1.4
daf-18(D66E)	152.9 ± 8.3	596.2 ± 7.0	65.0 ± 2.2	15.9 ± 1.2
daf-18(L115V)	136.9 ± 7.4	571.6 ± 3.5	65.7 ± 0.9	14.2 ± 0.8
daf-18(H138R)	173.1 ± 11.0	610.3 ± 7.0	72.1 ± 2.0 *	20.1 ± 2.1 *
daf-18(H168Q)	162.2 ± 15.1	632.4 ± 9.1 *	66.5 ± 1.2	16.8 ± 1.9
daf-18(T176I)	160.1 ± 19.7	617.7 ± 6.0	68.4 ± 2.6	18.9 ± 2.8
dlg-1(V964I)	173.2 ± 14.7	601.6 ± 5.1	204.8 ± 2.8	18.0 ± 1.8
egl-19(V331M)	137.2 ± 4.5	590.7 ± 4.2	65.1 ± 0.8	13.8 ± 0.9
egl-19(Y333S)	147.6 ± 9.2	575.0 ± 2.6	59.5 ± 1.2	12.2 ± 0.8
egl-19(n2368sd)	109.1 ± 6.7 *	380.1 ± 6.0 *	52.5 ± 2.1 *	18.7 ± 0.8 *
gap-2(C417Y)	177.2 ± 8.7	613.9 ± 9.3	62.7 ± 1.8	18.2 ± 0.9 *
gap-2(L660F)	172.4 ± 13.1	611.2 ± 11.0	61.1 ± 1.2	17.4 ± 0.9
hpo-29(L575S)	112.5 ± 7.0 *	570.2 ± 6.8	60.7 ± 1.6	13.3 ± 1.0
mpk-1(R332Q)	146.5 ± 7.2	619.8 ± 4.0	194.9 ± 2.7	14.7 ± 0.8
tph-1(R259Q)	127.6 ± 4.2	527.2 ± 8.9 *	77.5 ± 1.2 *	15.0 ± 0.6
tph-1(mg280)	204.9 ± 19.8 *	553.2 ± 9.3 *	61.0 ± 3.2	19.0 ± 2.1 *

All values are presented as mean ± SEM. *$p < 0.01$ via one-way ANOVA and multiple comparison.

A. CACNA1D



B. CHD8



C. CUL3

D. DLG4

E. GLRA2

F. MAPK3

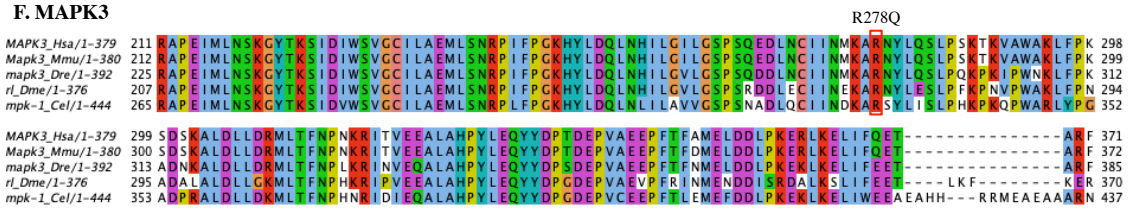


Figure S1. Protein alignments for the missense alleles tested in this study. Sections of the protein multiple alignments for human (*Hsa*) protein and its orthologs in mouse (*Mmu*), zebrafish (*Dre*), Drosophila (*Dme*), and *C. elegans (Cel)*. Circled are ASD-linked missense variants in the following human genes: (A) *CACNA1D*, (B) *CHD8*, (C) *CUL3*, (D) *DLG4*, (E) *GLRA2*, (F) *MAPK3*, (G) *PTEN*, (H) *NAA15*, (I) *SYNGAP1*, and (J) *TPH2*. Residues in the alignment have been colored by JalView (PMID:19151095) using the Clustal X coloring scheme.