A forward genetic screen for *Caenorhabditis elegans* pharyngeal gland cell defects and characterization of mutagenized strains with a homozygous viable gland cell under migration

by

Stephanie Tkachuk

A thesis submitted to the Faculty of Graduate Studies of The University of Manitoba

in partial fulfilment of the requirements for the degree of

Master of Science

Department of Biological Sciences

University of Manitoba

Winnipeg, Manitoba

Abstract

Caenorhabditis elegans is a powerful genetic tool to study development. The pharynx is a neuromuscular organ of the upper digestive tract used to study how cells regulate their shape and movements during development. In the embryo, a receptor tyrosine kinase-like receptor (ROR), *cam-1*, is necessary for a single dorsal gland cell to migrate through the pharynx. A mutagenesis screen to isolate factors that may work with *cam-1* examined 4,986 haploid genomes and isolated 60 strains with gland cell defects. Snip-SNP mapping placed one strain on the first chromosome, seven strains on the fourth, and one strain on the X chromosome. Complementation crosses and whole genome sequencing identified the LG IV mutations as *cwn-2*, a Wnt ligand, and *ham-1*, a storkhead box factor. The mutations on LG I and X require further testing to confirm the causative lesion. *ham-1* RNAi in the *cam-1* null background suggests that *ham-1* and *cam-1* function in different genetic pathways.

Acknowledgements

I thank Dr. Jay Kormish for her mentorship and for the opportunities to learn and grow as a scientist. I thank my committee members Dr. Jeffrey Marcus and Dr. David Merz for their support and guidance during my graduate studies. I would also like to thank Stephane Flibotte for his assistance in bioinformatic analysis and Andre Dufresne for technical microscopic support.

Thank you to Kormish lab colleagues Shinhye Kim, Maxwell Burg, Sasha Loewen, Emma Bennici Clendinnen, Mackenzie Sato, Monty Singh, Patrycja Sroga, Stacey Line, Tinsley Douglas, and Cheryl Camia for their camaraderie and technical assistance.

Finally, I would like to express my gratitude to my friends and family, who provided unwavering support and encouragement throughout my university career. Thank you especially to Andrew Yorski, Austein McLoughlin, and Arafa Khan for the coffee and fellowship that powered me through the challenging times.

Dedication

To my parents Steve and Kathie

Table of Contents

Abstract i
Acknowledgements ii
Dedication iii
Table of Contents iv
List of Tables vii
List of Figures viii
1 Introduction 1
1.1 *C. elegans* is a powerful genetic tool 2
1.2 The *C. elegans* Research Community 3
1.3 *C. elegans* transgenic transformation 5
1.4 Mature pharynx structure 7
1.5 Early pharyngeal development 8
1.6 Basic specification of the pharynx and its gland cells 8
1.7 Gland cell function 9
1.8 Overview of pharyngeal precursor cell movement and morphogenesis 10
1.9 Retrograde extension of the g1P pharyngeal gland cell 13
1.10 Gland retrograde extension is affected by the Ror receptor *cam-1* 14
1.11 Rors are conserved factors in embryonic organ development 14
1.12 Ror structure 16
1.12.1 Ror extracellular N-terminus 16
1.12.2 Ror intracellular C-terminus 17
1.13 Rors affect a variety of tissues during development 19
1.14 *cam-1* expression and roles in *C. elegans* embryonic development 20
1.15 Preliminary Results 24
1.16 Study Aims 24
2 Materials and methods 30
2.1 Strains and Maintenance 30
2.2 Mutagenesis and forward genetic screen for pharyngeal gland cell migration 31
2.3 Identifying Cam-1-like mutations obtained from the forward genetic screen 32
2.3.1 SNP Mapping- Outcrossing 32

2.3.2 SNP Mapping - Semi-quantitative analysis using unique restriction enzyme recognition sites and recombination 33
2.3.3 Complementation testing mapped mutant strains 34
2.3.4 Preparation of mutagenized strains for whole genome sequencing 35
2.3.5 Whole genome sequencing: bioinformatic analysis 38
2.3.6 Sanger sequencing to confirm identity of other mutant strains 39
2.3.7 Interval mapping to reduce LGI and LGX candidate genes 40
2.4 RNA interference 41
2.4.1 Collections for RNA extraction 41
2.4.2 RNA extraction 42
2.4.3 cDNA synthesis 42
2.4.4 Amplification of *cam-1* and *ham-1* cDNA into dsDNA 43
2.4.5 Amplification of L4440 control dsDNA 43
2.4.6 Synthesis and recovery of *cam-1*, *ham-1* and L4440 dsRNA 44
2.4.7 dsRNA injection 45
2.5 Slide/mount preparation for microscopy 46
2.6 Quantification and image capture 46
2.7 g1P distribution analysis 46
3 Results 55
3.1 60 strains with gland cells defects were isolated from forward genetic screening 55
3.2 Candidate Cam-1-like strains mapped to chromosomes I, II, IV, V, and X 55
3.3 Complementation crosses reveal two complementation groups on LGIV, and potential alleles of *cwn-2* and *cam-1* 58
3.4 Whole genome sequencing analysis 61
3.5 LGIV group A is *cwn-2*, and LGIV group B is *ham-1* 63
3.6 Two candidate genes each for *kor42* LGI and *kor44* LGX 64
3.7 g1P quantification of novel mutations 65
3.8 RNA interference of *ham-1* and *cam-1* 66
3.8.1 *ham-1* RNAi phenocopies *ham-1* strains possessing g1P under migration 66
3.8.2 *cam-1* RNAi produces an under-migration phenotype 67
3.8.3 *cam-1* and *ham-1* genetic interaction 67
4 Discussion 99
4.1 Complementation results 99

4.2 Hawaiian mapping-by-sequencing troubleshooting 101
4.3 *kor44* LGX recombination 102
4.4 Candidate genes on LGI and LGX 102
4.5 *ham-1* is a novel factor in pharyngeal development 103
4.6 *ham-1* functions in a different genetic pathway from *cam-1* 106
4.7 Strains may be temperature sensitive 108
5 Conclusion and Future Directions 111
5.1 Summary 111
5.2 Proposed model 111
5.3 Future Directions 112
5.3.1 *srt-61* and *plpp-1.3* RNA interference 112
5.3.2 Examining CAM-1 and HAM-1 distributions in opposite genetic backgrounds . 112
5.3.3 Examining pharyngeal polarity in *cam-1* and *ham-1* backgrounds 113
5.3.4 Further testing *cam-1* kinase domain function in pharyngeal development 114
5.3.5 Identifying other genes isolated from the forward genetic screen 114
6 References 116
7 Appendix 129

List of Tables

Table 2.1 Mutagenized strains selected for characterization in this thesis. 31
Table 2.2 Oligonucleotides used for chromosomal linkage mapping. 48
Table 2.3 Oligonucleotides used in amplifying and Sanger sequencing *cam-1*C genomic sequence in *kor37*. 49
Table 2.4 Additional oligonucleotides used in LGI interval mapping of *kor42*. 50
Table 2.5 Additional oligonucleotides used in LGX interval mapping of *kor44*. 51
Table 2.6 Oligonucleotides used in amplifying and Sanger sequencing *cam-1*A genomic sequence in *kor37*. 52
Table 2.7 Oligonucleotides used in amplifying and Sanger sequencing *cwn-2* genomic sequence in *kor32*, *kor43*, and *kor52*. 53
Table 2.8 Oligonucleotides used in amplifying *ham-1* and *cam-1* cDNA. 54
Table 3.1 Summary of chromosomal linkage analysis for sixteen alleles. 72
Table 3.2 Results of complementation crosses between strains mapping to LG IV 76
Table 3.3 Results of complementation crosses between strains mapping to LG IV and *cwn-2(ok895)*. 78
Table 3.4 Results of complementation cross between strains mapping to LGV 78
Table 3.5 Quality of genomic DNA preparations submitted to Genome Quebec for Illumina paired-end whole genome sequencing. 78
Table 3.6 Candidate gene list for *kor5* 83
Table 3.7 Candidate gene list for *kor15* 83
Table 3.8 Candidate gene list for *kor42* 84
Table 3.9 Candidate gene list for *kor44* 84
Table 3.10 Candidate gene list for *kor45* 85
Table 3.11 Candidate gene list for *kor49* 85
Table 3.12 Under-migration penetrance and additional phenotypes observed in *kor5*, *kor15*, *kor42*, and *kor44*. 91

List of Figures

Figure 1.1 Stages of *C. elegans* embryonic development 27
Figure 1.2 The *C. elegans* life cycle at 22˚C. 28
Figure 1.3 Cartoon representation of the mature *C. elegans* pharynx 29
Figure 3.1 Summary of mutant strain characterization completed in this thesis 71
Figure 3.2 JDK99 shows linkage to chromosome IV 73
Figure 3.3 JDK98 shows linkage to the X 74
Figure 3.4 Linkage group analysis for JDK56 75
Figure 3.5 Summary of complementation crosses between strains mapping to LGIV 77
Figure 3.6 Sample output of "Per base sequence quality" from a FastQC report on raw reads. .. 79
Figure 3.7 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool 80
Figure 3.8 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool 81
Figure 3.9 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool 82
Figure 3.10 SNP mapping strategy 86
Figure 3.11 snip-SNP interval mapping results of (A) *kor42* and (B) *kor44* 87
Figure 3.12 Schematic of genes indicating location and nature of mutations sequenced 88
Figure 3.13 The identified lesions for *kor5* and *kor15* in HAM-1 and *kor37* in CAM-1 89
Figure 3.14 Quantitation of g1P gland cell position 90
Figure 3.15 Tail defects found in strains JDK114 *ham-1*(*kor15*); *ivIs17* and JDK119 *ham-1*(*kor5*); *ivIs17* 92
Figure 3.16 Bar graph and boxplots comparing the g1P distributions of *ham-1*(*kor5*), *ham-1*(*kor15*); L4440 RNAi, and *ivIs17*; *ham-1*(*RNAi*) 93
Figure 3.17 Bar graph and boxplots comparing the g1P distributions of two GD282 treatments *ivIs17*; *cam-1*(*RNAi*) and control *ivIs17*; L4440 RNAi. 94
Figure 3.18 Electrophoretic gel of L4440, *cam-1*, and *ham-1* dsRNA. 95
Figure 3.19 Quantitation of g1P gland cell position in (A) *ham-1*(*kor15*); *cam-1*(*RNAi*) double mutant and (B) *cam-1*(*gm122*); *ham-1*(*RNAi*) double mutant 96
Figure 3.20 Examples of pharyngeal defects observed in *cam-1*(*gm122*); *ham-1*(*RNAi*) mutants 97
Figure 3.21 Examples of pharyngeal defects observed in *ham-1*(*kor15*); *cam-1*(*RNAi*) mutants. 98
Figure 4.1 Boxplots comparing g1P distributions between different temperature treatments ... 110

List of material used under Creative Commons licence in this study

Section	Reference	Licence
Figure 3.10	Doitsidou et al., 2010	CC-BY

1 Introduction

For organs to form within a developing embryo, cells must migrate to new locations to assemble with other cells destined to become part of the organ. Generally, a cell must become polarized to the direction it will move, initiate membrane protrusion with its cytoskeleton, and stabilize its advancement by adhering to its surroundings. These steps are accomplished by incorporating signals and cues from both inside and outside the cell to navigate its three-dimensional environment; some of these underlying genetic and molecular pathways are thought to be conserved across all eukaryotic organisms (Kurosaka and Kashina, 2008; Mak et al., 2016). Although there are various cells and cell types within a single organ, they all must be coordinated to some degree in their processes to take on the correct shapes and functions. Mis-regulated migratory cells can lead to congenital defects and disease, such as cleft palate and cancer in humans (Friedl and Alexander, 2012; He et al., 2008).

The pharynx of the nematode *Caenorhabditis elegans* is a tool to study organ morphogenesis. Functionally, the pharynx is the worm foregut, which is a muscular pumping tube that draws bacterial food from the mouth into the intestine. The relatively small number of cells in the pharynx, combined with other properties of the worm, make it a powerful tool for studying organ development. There is a “simple to complex” duality that comes from studying pharynx development. By understanding how a simple ball of cells changes shape to form an elongated tube, it becomes possible to apply these principles to more complex organs and organisms.

1.1 *C. elegans* is a powerful genetic tool

C. elegans is an ideal developmental and genetic tool due to its well-studied cellular biology and relative ease of genetic manipulation. It is microscopic, about 1mm in length, and has a rapid life cycle of about three days when kept under ideal culturing conditions (Byerly et al., 1976). Most of a *C. elegans* population is self-fertilizing hermaphrodites, and a single worm can give rise to hundreds of progeny. About 0.1% of worms develop into males through spontaneous non-disjunction of the X chromosome (Emmons, 2005; Hodgkin et al., 1979). Males only produce sperm and therefore cannot give rise to a new generation on their own, however they are useful for genetic outcrossing in the laboratory.

The *C. elegans* embryo develops inside the hermaphrodite uterus for 150 minutes after fertilization before it is laid outside the body. At 22˚C, it takes approximately 840 minutes to progress from a single-celled zygote to a hatched larva (Figure 1.1). Larvae proceed through the L1-L4 stages of growth before becoming an adult capable of reproduction. These stages are separated by molts where the animal sheds its cuticle and can subsequently increase in size (Figure 1.2). An alternate stage after L2, called the dauer stage, can be entered when the worm senses extreme temperatures, high population density, and deficient food supply. Dauer form allows the worm to remain in a dormant stage until environmental conditions improve (Fielenbach and Antebi, 2008).

Cell movements in both embryos and larvae are well-characterized. The cellular lineage of *C. elegans* is invariant – that is, all somatic cells in the body plan undergo the same divisions, movements, and changes within every individual. This predictability, coupled with the transparency of the embryo and worm body, has allowed for incredibly detailed diagramming and anatomical description of the entire worm life cycle (Sulston, 1977; Sulston et al., 1983).

1.2 The *C. elegans* Research Community

The *C. elegans* research community has collaborated to create extensive resources for studying the worm. For example, online resources such as WormBook and WormAtlas are free for public access and contain detailed reviews and illustrations by experts in the many fields of study conducted with nematodes. The *Caenorhabditis* Genetics Center at the University of Minnesota maintains a repository of worm strains generated by various labs and makes them available to other researchers.

C. elegans was the first multicellular organism to have its genome sequenced. The six pairs of chromosomes (five autosomal and one X) contain approximately 100 megabases with about 20,000 genes (Consortium*, 1998). Re-sequencing and additional experiments have resulted in a refined, well-annotated reference genome. For example, sequencing strains sampled from across the world has revealed tracts of single nucleotide polymorphisms (SNPs) between natural isolates (Hodgkin and Doniach, 1997; Koch et al., 2000). Expressed sequence tag (EST) libraries are partial cDNA sequences that have been cloned from large-scale RNA extractions. These were instrumental in finding alternative splice sites and new genes in *C. elegans*. ESTs can also confirm the coding sequences produced by whole genome sequencing (Mangone et al., 2010; McCombie et al., 1992; Waterston et al., 1992). The genome, as well as a wealth of other worm-related resources, are compiled and regularly updated on the Wormbase database (www.wormbase.org). Each identified gene has a page that includes characterized alleles, expression profiles, reagents, and published references. Other databases such as SNPs and ESTs have been incorporated into the site as well. This allows researchers to access virtually any information regarding *C. elegans* biology in an organized, comprehensive way.

Genomic and genetic analyses have revealed characteristics such as the unusual presence of operons in the animal. Groups of genes common to a biological function in the worm are under simultaneous control of a common promoter (Blumenthal and Gleason, 2003). Despite differences in genome organization, many protein sequences in *C. elegans* have probable homology with human proteins (Lai et al., 2000). It is estimated that 40-60% of *C. elegans* protein-coding genes share some degree of similarity with a human gene (Lai et al., 2000; Shaye and Greenwald, 2011). Some human sequences can even functionally replace the homologue in the worm (Levitan et al., 1996; Stern et al., 1993). For example, the worm gene *ceh-22* is expressed in the pharyngeal muscle, and activates expression of other pharyngeal muscle genes during development. In a *ceh-22* loss-of-function mutant, muscle development is defective. Expressing the related vertebrate gene *nkx2.5* in the worm rescues pharyngeal muscle development and activates the appropriate downstream genes (Haun et al., 1998). This demonstrates that study of conserved pathways and developmental processes in the worm can be used to understand development in other organisms.

Classical genetics can be used in conjunction with modern sequencing techniques to map and analyze mutations. Protocols to infer genetic distance analyze recombinant progeny using classic genetic markers or higher resolution physical markers such as snip-SNPs (Brenner, 1974; Davis et al., 2005). Snip-SNP mapping takes advantage of a polymorphic wild-type strain, Hawaiian CB4856, which possesses single nucleotide polymorphisms (SNPs) that differ from the N2 Bristol reference strain (Hodgkin and Doniach, 1997; Wicks et al., 2001). These Hawaiian-specific SNPs are within restriction sites, where they either generate a new recognition cut site or remove one that exists in the N2 background. By crossing an N2 mutant with the Hawaiian strain and observing the progeny restriction digest patterns, one can use these as physical markers in

recombination mapping. These snip-SNPs also occur approximately every 1,000 base pairs, so that large numbers of recombinant progeny can be used to map phenotypes to single genetic loci.

1.3 *C. elegans* transgenic transformation

C. elegans is receptive to transgenic transformation, allowing for introduction of exogenous DNA into the germline (Mello and Fire, 1995). Transgenes are foreign pieces of DNA taken up by a cell. A common use of transgenes in *C. elegans* research is to create reporters. Fusing a promoter or open reading frame to a fluorescent protein, such as the green fluorescent protein, GFP, will visualize when and where that construct is expressed. The worm's transparent body makes for straightforward screening and imaging under a microscope equipped with fluorophore excitation. Live-imaging techniques have also been developed to allow for observation of a reporter over an extended period of time (Podbilewicz and Gruenbaum, 2006).

Transgenes can be introduced by microinjecting DNA constructs into the gonad, where they will be taken up by germ cell nuclei. The DNA is usually maintained as a multi-copy extrachromosomal array (Mello et al., 1991). An array is not stable, though, and may not be inherited by all dividing cells in an embryo. Arrays can be stabilized by integrating them into the genome. By exposing transgenic worms to mutagenic chemicals or irradiation such as ultraviolet (UV) rays, double-stranded breaks will occur throughout the genome. This provides an opportunity for the array to ligate to a broken portion of chromosome prior to DNA repair (Mello et al., 1991).

The tandem copies of an array can be disadvantageous. One issue is that a transgene is prone to over-expression in a transgenic individual, meaning it would not accurately represent how the endogenous version of the gene is expressed. Repetitive sequence can also be silenced in

the germline, which would prevent transgene expression in the developing oocyte (Kelly et al., 1997).

To address issues with transgenic arrays, techniques such as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) have been developed for use in *C. elegans*. This system was discovered in bacteria and archaea as a form of defense against exogenous DNA (Mojica et al., 2005). It has since been modified for use in eukaryotes. One version of CRISPR technology uses an engineered guide strand to target a homologous sequence in the genome. The Cas9 protein (CRISPR-associated) interacts with this site and creates a precise double-stranded cut. A donor sequence that contains the desired transgene will undergo homologous repair with the cut site. This allows the donor sequence to act as a template so a copy of the transgene is incorporated directly into the genome (Tzur et al., 2013). Changes such as small or large indels can also be incorporated instead of a transgene.

In theory, CRISPR allows for precise modification of a genome. Transgenes can be inserted in low copy numbers, preventing over-expression issues that come with integrating transgenic arrays. Targeting the exact modification site also reduces the risk of disrupting/rearranging important loci that may occur with random integration. There are challenges that also come with utilizing CRISPR, however. Because some mis-matches between the guide strand and the homologous genomic sequence can be tolerated, off-target edits can occur. Varying levels of reagent activity and successful repair can still require laborious screening to find worms with the appropriate modifications (Sugi, 2016). Repetitive sequences can also be resistant to being cut by some versions of Cas9 (Harrison et al., 2014; Struhl and Segal, 2013), further lowering the success rate of obtaining recombination. Using recombinant Cas proteins with increased regulation of activity may help address these issues.

1.4 Mature pharynx structure

Mature pharynx structure resembles a tripartite tube with two lobes. It runs along the anterior-posterior axis, connecting to the anterior oral cavity and the posterior pharyngeal-intestinal valve. The valve is a series of pharyngeal cells organized into three rings at the very posterior of the pharynx. It is responsible for aligning with another ring of cells derived from the intestine. This alignment is crucial for establishing a continuous digestive tract between the pharynx and intestine (Rasmussen et al., 2013). Of the 95 cells composing the pharynx, there are six differentiated cell types: epithelia, muscle, marginal, neuron, gland, and valve (Albertson and Thomson, 1976; Altun and Hall, 2009). Nine epithelial cells encircle the anterior end of the pharynx to connect the organ to the hypodermis. The pharynx is a neuromuscular organ composed of neurons and muscle cell groups supported structurally by marginal cells. The muscle groups have distinct structures and functions along the antero-posterior axis. The sections are numbered one through eight relative to their position along the antero-posterior axis. For example, pm8 is the posterior-most segment of muscle, and is located at the back of the terminal bulb (Figure 1.3). Pharyngeal neurons are embedded in the pharynx and function to innervate the muscle, synapse with somatic neurons, or act in neurosecretion (Ramakrishnan et al., 2014). The five gland cells belong to two different classes, g1 and g2. There is a single dorsal g1P cell, while g1a and g2 cells are located ventrally and are paired as left and right counterparts. The gland cells have ducts that connect to the lumen and secrete mucins, lubricating the pharyngeal lumen to aid in passing bacteria through the foregut (Smit et al., 2008).

1.5 Early pharyngeal development

During embryogenesis, the cells of the pharynx originate from divisions of two distinct predecessor cells. Proceeding pharyngeal development uses some conserved pathways and factors, such as *pha-4*/FoxA transcription factor. The FoxA family is a group of transcription factors heavily involved in organogenesis. They are pioneer factors that bind to chromatin and allow other transcription factors to access regulatory areas of DNA (Friedman and Kaestner, 2006). This means FoxA factors are crucial for proper gene regulation during organ development. Studying *pha-4* in the worm has contributed to this understanding of the FoxA mechanism of function.

pha-4 has been identified as the pharyngeal organ identity factor. PHA-4 is present in all cells fated to become part of the pharynx; when *pha-4* is mutated and not expressed in the embryo, the pharynx does not form (Horner et al., 1998; Kalb et al., 1998; Mango et al., 1994). Ectopic expression of *pha-4* can cause a fate switch in the cells of the early embryo. When *pha-4* is expressed in a cell not normally destined to be part of the pharynx, other pharynx-specific genes are also expressed. This causes the cell to switch fates and become an ectopic pharyngeal cell.

1.6 Basic specification of the pharynx and its gland cells

PHA-4 is the top of the pharyngeal transcriptional hierarchy and activates all other genes in pharyngeal development. The best characterized example is within pharyngeal muscle development: PHA-4 activates *ceh-22*/*Nkx2.5*, which then both work to activate and enhance *myo-2*, a myosin heavy chain that is expressed only in pharyngeal muscle cells (Okkema et al., 1997; Okkema and Fire, 1994). Regarding the gland cells, PHA-4 directly targets *hlh-6*, which is a basic helix-loop-helix transcription factor (Raharjo and Gaudet, 2007; Smit et al., 2008). *hlh-6*

expression is restricted to the gland cells because it needs to be activated by PHA-4. The presence of the HRL *cis* factors, or hlh-6 regulatory elements, in the *hlh-6* promoter repress its expression in any non-gland cell (Raharjo and Gaudet, 2007). Additional transcription factors, such as LAG-1, interact with HRLs to repress expression outside of the gland cells (Ghai and Gaudet, 2008). HRLs work in combination with PHA-4 to restrict *hlh-6* expression to a specific lineage (cells derived from the MS blastomere) and a specific position (the posterior of the pharynx).

HLH-6 can then go on to activate a battery of other gland-specific genes. This type of transcription factor has two important domains. First, the helix-loop-helix domain dimerizes with other proteins to create a functional complex. Second, the basic domain interacts with DNA to affect gene expression (Jones, 2004). HLH-6 binds to the *cis* factor PGM1, or pharyngeal gland motif 1. PGM1 is apparently necessary and sufficient for the expression of some gland-specific genes (Ghai et al., 2012; Smit et al., 2008). Some genes do not require PGM1 nor HLH-6 to be properly expressed in the gland cells, however. The transcription factors activating these HLH-6-independent batteries have not yet been identified.

1.7 Gland cell function

hlh-6 loss of function worms are slow-growing, have a stuffed pharynx phenotype and a feeding defect that contributes to a low penetrance of larval arrest (Smit et al., 2008). The g2 gland cells are usually missing in the loss-of-function background. This defect has been attributed to the inability of these cells to maintain a gland cell identity and they arrest prior to their terminal cell division. Genetic ablation of the gland cells also results in a stuffed pharynx phenotype. This shows that even though the glands are present in *hlh-6* mutants, their function is compromised. A stuffed pharynx indicates that bacteria are not being passed through the pharynx

and into intestine. These issues can be resolved by switching the worm food source from OP50 *E. coli* to a less sticky HB101 strain. This suggests the gland cells support a secretory function in the foregut to lubricate the pharyngeal lumen. This is supported by the discovery of a set of gland-specific *phat* genes (pharyngeal gland toxin-related). These genes contain a sequence like the *Toxicara canis* ShK domain, including serine-threonine-rich stretches that are likely sites for O-linked glycosylation. This indicates the PHAT gene products are mucin-like proteins. Mucin carbohydrate moieties extend from a protein backbone to form a bottle-brush pattern. This arrangement attracts water to create hydrogels that lubricate and protect epithelial surfaces (Bansil and Turner, 2006; Crouzier et al., 2015). Indeed, a PHAT-5 reporter showed the protein localized in the gland cells, along the gland extensions and ducts, and along the pharyngeal lumen. When *hlh-6* is mutated, PHAT-5 expression is severely reduced (Smit et al., 2008). Thus, an important function of the pharyngeal gland cells is to secrete mucin-like proteins that lubricate the foregut.

1.8 Overview of pharyngeal precursor cell movement and morphogenesis

After *pha-4* becomes active, a cluster of pharyngeal precursor cells migrates into the embryo during gastrulation and ingression. This occurs between 150-260 minutes of embryogenesis (Portereiko and Mango, 2001; Rasmussen et al., 2012). Once internalized, the pharyngeal cells arrange themselves in a double plate. This double plate is a two-cell-wide, bilaterally symmetric structure that runs along the dorsal-ventral axis. At this stage, the pharyngeal precursor cells establish a boundary from non-pharyngeal cells and have a left-right orientation. They do not cross the plate midline nor do they intermingle with non-pharyngeal cells. In mutants that affect left-right identity, the pharyngeal cells do not maintain the double plate and form a nondescript cluster that aberrantly localizes at the ventral side of the embryo

(Rasmussen et al., 2012). Double plate formation shows that early pharyngeal development is autonomous: the pharyngeal cells have a strict spatial organization and require minimal interaction with non-pharyngeal cells.

After double plate ingression, the exterior of the double plate becomes surrounded by β laminin (LAM-1). This defines the basal surface of the organ. In response, the PAR complex localizes to the midplane, which will become the apical surface. The apical surface contracts to create wedge-shaped cells, transforming the double plate into an oblong pie shape (Horner et al., 1998; Rasmussen et al., 2012; Santella et al., 2010). This pie shape is known as the pharyngeal primordium. Correct formation of the primordium is crucial because it is a rudimentary plan of the mature pharynx. Many cells of the pharynx are roughly in their final positions at this stage. The established apical-basal polarity determines where the lumen and basement membrane will develop. Proper establishment and maintenance of the lumen is needed to form a continuous channel for food passage. The basement membrane is important because it provides structural support and signalling to fine-tune cellular polarity. Establishing apical polarity is also important for the recruitment of apical junction components (Labouesse, 2006). The apical junctions function to maintain cell-cell adhesion during morphogenesis. Keeping contact with the other pharyngeal cells will help ensure rearrangement and changes in shape occur with correct order and timing.

Near mid-embryogenesis (~330-390 minutes), the extension of the pharyngeal primordium is coordinated with the elongation of the embryo. During elongation, cells coordinate shape changes *via* actomyosin contraction to increase embryo length at the expense of circumference. The epidermal cells, which are the outermost layer of cells, change from a cuboidal to rectangular shape as they extend along the antero-posterior axis. This exerts force on

the internal cells of the embryo to also undergo elongation (Costa et al., 1998; Gally et al., 2009; Piekny et al., 2003; Priess and Hirsh, 1986).

Pharyngeal extension begins with remodelling the primordium anterior. The epithelial cells pull away from the primordium so they are also in contact with the nearby arcade cells. This means the apical and basal membranes have been shifted so they align with the dorso-ventral axis. The pharyngeal cells link to the arcade cells through apical junctions, creating a continuous connection between the pharynx, oral cavity, and exterior environment. Contraction of the epithelium pulls the pharynx anteriorly, and elongates the pharyngeal cells along the antero-posterior axis (Portereiko and Mango, 2001).

The primordium posterior undergoes a more extensive rearrangement to form the pharyngeal-intestinal valve. The valve cells are dispersed throughout the primordium and must relocate to the very posterior to organize into a series of rings that contacts the developing intestine. Contact with the intestinal primordium confers polarity to the valve, and helps the two organs align their apico-basal axes (Rasmussen et al., 2013). This will ensure apical junction formation, cell adhesion, and continuous lumen formation at the posterior end of the pharynx.

Some pharyngeal cells undergo a change from the wedge shape established in the polarized primordium. For example, pm8 takes on a thin donut shape to cover the posterior end of the pharynx. The cell detaches from its dorsal location in the primordium and slowly spreads along the posterior surface. A transient trail of laminin helps guide pm8 during this movement. pm8 then wraps around the midline and fuses with itself to create a continuous ring. Notch signalling is required for both the detachment and fusion steps of pm8 morphogenesis. No other pharyngeal cells express Notch pathway factors, lending to the idea that each cell in the pharynx

has its own unique combination of factors that influence its terminal shape (Rasmussen et al., 2013, 2008).

1.9 Retrograde extension of the g1P pharyngeal gland cell

During elongation (~390 minutes and onward), the dorsal gland cell, called g1P, moves through the pharynx using retrograde extension. The gland cell creates an anchor at the pm2/pm3 border and migrates posteriorly through the organ leaving a cytoplasmic extension in its migratory wake (Figure 1.3). This is distinct from anterograde-based cell movements, such as axonal pathfinding, where the neuronal cell body remains stationary and a growth cone at the tip of the axon extends to make contact in a different location (Lowery and Vactor, 2009). In *C. elegans*, retrograde extension was first described in the dendrites of amphid sensory neurons, and was recently shown to occur in the axon extension of the AVB neuron (Heiman and Shaham, 2009; Singhal and Shaham, 2017). The first phase of muscle arm extension in the worm may have a similar process. After a myoblast makes contact with a nearby motor neuron, the myoblast moves away and leaves a short extension behind (Dixon, 2005). A similar process has also been observed in the mouse brain, where cerebellar granule cell bodies actively migrate through the molecular layer of the cerebellum to create T-shaped cells with points of contact in multiple layers (Komuro and Yacubova, 2003).

The molecular mechanisms behind retrograde extension are poorly understood. In the worm, the amphid dendrites were shown to require *dex-1* and *dyf-7* for proper anchoring of their tips (Heiman and Shaham, 2009). These extracellular proteins contain zonadhesin and zona pellucida domains, respectively. The glial cells associated with the dendrites contribute to the anchoring process by expressing DEX-1 (Singhal and Shaham, 2017). What drives the movement of the cell body after anchoring is unknown. It was hypothesized that *sax-3*/Robo

influences retrograde extension because *sax-3* mutants have abnormal neuronal trajectories. A strong loss-of-function mutation does not affect amphid neuron cell body placement nor dendrite extension, however. *sax-3* is not likely to be a major factor in the retrograde portion of neuronal development (Singhal and Shaham, 2017).

1.10 Gland retrograde extension is affected by the Ror receptor *cam-1*

In the Kormish laboratory, the Ror (Receptor tyrosine kinase-like orphan receptor) *cam-1* has been found to play a role in pharyngeal development . In the *cam-1(gm122)* loss-of-function mutation, the g1p gland cell prematurely terminates retrograde extension and stops in a more anterior position in the pharynx. Instead of the pm7 area of the terminal bulb, the gland is commonly found in pm5 or pm6 (Figure 3.14). The defect may involve several cells in the pharynx since an asymmetry of the metacorpus is frequently observed. Developmental defects in other tissues are less penetrant and less severe when the kinase domain is mutated(Forrester et al., 1999; Kim and Forrester, 2003; Modzelewska et al., 2013). The developmental defects of multiple cell types in the *C. elegans* pharynx provide a model to study the molecular genetics of Ror and potentially the role of its kinase domain in organ development.

1.11 Rors are conserved factors in embryonic organ development

Rors are members of the RTK (Receptor Tyrosine Kinase) superfamily. Rors are conserved across many animals, from the mollusk *Aplysia californica* to mammals such as mouse and human (Minami et al., 2010; Witte et al., 2010). They received their “orphan” designation because a ligand and signalling pathway had not been initially identified, however their interaction(s) and roles have since been established (Green et al., 2008b). During embryonic development, Rors participate in tissue and organ formation through non-canonical Wnt signalling. Wnts are glycoproteins secreted as extracellular ligands, and Rors situated in the

membranes of nearby cells bind Wnts. The non-canonical pathway refers to the fact that the intracellular protein β-catenin is not involved in the downstream response (Veeman et al., 2003).

In the worm, *cam-1* can have two different functions upon binding Wnt. The first function is to act as a negative regulator or antagonist called a Wnt "sink". In this context the Ror simply binds Wnts to prevent the ligand from binding other nearby receptors and initiating a signalling cascade (Forrester et al., 2004; Green et al., 2008a, 2007; Kim and Forrester, 2003). This modulates the concentration of Wnt available in the extracellular environment. Tightly controlling ligand concentration is important for pathways that require a specific threshold of binding activity to function.

cam-1's second function is in positive regulation. Wnt binding in this context requires the intracellular portion of Ror, suggesting the receptor initiates a signalling cascade to directly influence cellular processes. The non-canonical Wnt pathways include the planar cell polarity (PCP), JNK, and calcium signalling pathways. The Wnt/calcium pathway increases the concentration of intracellular calcium to activate PKC, CaMKII, and NFAT (Komiya and Habas, 2008; Kühl et al., 2000). These downstream components can then affect processes such as cell division, morphology, and migration. The Wnt/JNK and Wnt/PCP pathways share core components. They are differentiated because they were first separately characterized in vertebrates and *Drosophila*, respectively (Li et al., 1999; Taylor et al., 1998; Wolff and Rubin, 1998). Briefly, Dishevelled activates Rho GTPases, which can then activate JNK. PCP includes additional components such as the transmembrane proteins VanGogh and Flamingo and the intracellular factors Prickle and Diego (Devenport, 2014; Kestler and Kühl, 2008). The JNK and PCP pathways polarize large sheets of cells along a body axis. This allows processes such as

convergent extension to occur, where polarized cells narrow along one axis and elongate along the perpendicular axis (Tada and Heisenberg, 2012).

1.12 Ror structure

1.12.1 Ror extracellular N-terminus

Across species, Rors have a similar N-terminus/extracellular structure. They usually have an immunoglobulin (Ig) domain, Cysteine-rich domain (CRD), Kringle (Kr) domain, and a single-pass transmembrane domain (TM) (Forrester et al., 1999; Green et al., 2008b) (Figure3.13B). In other systems, such as the immune system T-cells, the Ig domains have been shown to mediate protein interaction and complexing and control processes such as immune stimulation (Peggs and Allison, 2005). In the worm, genetic and protein interactions have shown that RIG-3/IgCAM interacts directly with the CAM-1 Ig domain at neuromuscular junctions. This pre-occupies the binding domain of CAM-1, preventing it from interacting with Wnt ligand. This allows for the proper expression and assembly of acetylcholine receptors at the neuromuscular junction, and therefore normal synapse development and activity (Francis et al., 2005; Jensen et al., 2012; Pandey et al., 2017). In the muscle-specific kinase (MuSK), Ig repeats facilitate oligomerization of an essential complex at vertebrate neuromuscular junctions (DeChiara et al., 1996; Zhang et al., 2011). In fact, *cam-1*'s role in neuromuscular junction development may originate from the function of an ancestral MuSK homologue. In support of this there is no MuSK gene in *C. elegans*, and *cam-1* functional domains are equally similar to vertebrate MuSK and Ror (Forrester et al., 1999; Sossin, 2006). It is thought that *cam-1* may represent the original function of the now diverged receptors.

Similarly, proteases of the blood clotting cascade contain Kringle (Kr) domains that function in protein interaction (Klezovitch and Scanu, 1996; Pozzi et al., 2014). This has not yet

been demonstrated with the Ror Kr domain. A recent study has shown that the Kr domain assists Rors to form oligomers in the manifestation of chronic lymphocytic leukemia (Yu et al., 2016); it may be that in normal development, the Kr domain assists in the assembly of different receptor complexes.

The cysteine-rich domain (CRD) is a widely conserved motif possessing invariable cysteines that form disulfide bridges that affect protein folding. The Ror CRD closely resembles the canonical Wnt receptor Frizzled (Dann et al., 2001; Xu and Nusse, 1998). *in vitro* binding assays demonstrated that Wnt5a can bind directly to Ror2 CRD in mice (Oishi et al., 2003)(Oishi et al., 2003)(Oishi et al., 2003)(Oishi et al., 2003)(Oishi et al., 2003). A reverse binding assay using CAM-1 confirmed the Wnt ligand CWN-2 is able to bind to the CRD domain (Green et al., 2007). Further, mutation in the *C. elegans cam-1* CRD shows that it is required for proper migration in some cells (Kim and Forrester, 2003). Genetic interactions between *cam-1* and Wnts, which indicate they function in the same genetic pathway, support this evidence (Green et al., 2007; Hayashi et al., 2009; Kennerdell et al., 2009). For example, a *cam-1 cwn-2* double mutant has the same severity as a loss-of-function mutation in either single gene when examining nerve ring displacement along the antero-posterior axis (Kennerdell et al., 2009). These pieces of data suggest that the *cam-1* CRD also binds Wnt ligands.

1.12.2 Ror intracellular C-terminus

The intracellular structure of Rors is much more divergent. Although many have a tyrosine kinase (TK) domain, it has been debated whether its kinase activity has been maintained. For example, human Ror TK functions as a pseudokinase with low or absent kinase activity and more likely functions as a scaffold for other proteins with enzymatic functions (Bainbridge et al., 2014). In *C. elegans*, the kinase domain is not required for migration of some

neurons (Forrester et al., 1999; Kim and Forrester, 2003). This is consistent with Ror function in the nervous system of other invertebrates such as flies and sea slug (McKay et al., 2001; Wilson et al., 1993). A kinase-dependent function has been established in sensitized genetic backgrounds. For instance, CAM-1 is required for polarity of the hypodermal cells during vulva development (Forrester et al., 1999; Green et al., 2008a). The retention of *cam-1* kinase activity is probably an anomaly. Its convergence of Ror and MuSK activity may have required the gene to keep its intracellular signalling function and have dual roles during development (Forrester et al., 1999; Sossin et al., 2006).

Recently, it was proposed that CAM-1 can have two simultaneous yet antagonistic functions to affect the same cell (Chien et al., 2015) .Examining putative null and hypomorphic activity, intracellular deletions, kinase domain deletions, and mosaic expression of CAM-1 shows the intracellular domain is necessary for proper ALM neuronal polarity. It is proposed that CAM-1 has two distinct functions, one that is autonomous and one that is non-autonomous. In the non-autonomous function in non-ALM cells, CAM-1 is sequestering Wnt ligand. This would modulate Wnt distribution around the ALM cell bodies, which is probably used to control signalling in a parallel canonical Wnt pathway with the MOM-5/Frizzled receptor. Within the ALM, cell-autonomous RTK signalling by CAM-1 contributes to ALM polarity. In the cell body, CAM-1 is probably signalling in a non-canonical Wnt pathway to help establish anterior polarity in the neuronal cell body. Mosaic analysis emphasizes that the multiple roles for CAM-1 can be masked when *cam-1* mutation occurs in autonomous and non-autonomous tissues at the same time. When CAM-1 is missing in the ALM cell body but present in surrounding cells, the loss of anterior polarity was enhanced. This would occur because the CAM-1 signalling pathway was abolished, but because its Wnt sequestration activity persisted, the MOM-5 pathway was also

antagonized. This scenario removes any possible redundancy between the CAM-1 and MOM-5 pathways, causing a stronger loss of polarity than when *cam-1* mutation affects both autonomous and non-autonomous cells. When CAM-1 is present only in the ALM cell body but not surrounding cells, polarity is normal. Because there is little Wnt sequestration in this mosaic, both the CAM-1 and MOM-5 pathways are able to signal and contribute to proper polarity.

Following the tyrosine kinase domain is a Serine/Threonine rich (Ser/Thr) domain. In vertebrate Ror2, the Ser/Thr domain is phosphorylated by GSK-3 upon Wnt5a stimulation, inducing cell migration (Yamamoto et al., 2007). While *C. elegans* only has one Ser/Thr domain, vertebrate Rors have two Ser/Thr domains separated by a proline-rich region (Forrester, 2002; Forrester et al., 1999; Hubbard, 1999). Proline regions can contribute to protein-protein interactions or aid in transmitting extracellular signals (Kay et al., 2000; Williamson, 1994). The helical conformation of a polyproline domain can have low-affinity but high-specificity binding of other proteins. This would allow for rapid assembly and disassembly of a signal transduction complex in response to the binding of an extracellular ligand (Adzhubei et al., 2013). The vertebrate Ror may have acquired this region to help support its pseudokinase domain in recruiting enzymatically active factors. Few inquiries have examined the functionality of the worm Ser/Thr domain but it is speculated to be required for some neuronal cell movements (Kim and Forrester, 2003).

1.13 Rors affect a variety of tissues during development

In *Drosophila melanogaster* and the sea slug, there is a single Ror homologue. Both appear to be limited to expression in the nervous system during embryonic development (Wilson et al., 1993; McKay et al., 2001). Vertebrate Rors, however, have a more complex expression. There are two Ror orthologues in mammals, referred to as Ror1 and Ror2. In the mouse embryo, Rors

are expressed in many tissues including neurons, limbs, face, heart, and lungs (Matsuda et al., 2001). Consequently, Rors are involved in events such as convergent extension (Schambony and Wedlich, 2007), gut elongation (Bakker et al., 2012; Cervantes et al., 2009; Yamada et al., 2010), palate closure (He et al., 2008), and bone and digit formation (Y. Liu et al., 2007; Yan Liu et al., 2007; Witte et al., 2010). A homozygous *Ror2* knockout mouse is perinatal lethal and has abnormal heart morphology and shortened limbs. A homozygous *Ror1* knockout is also perinatal lethal, but does not have obvious heart and skeletal abnormalities. A *Ror1 Ror2* double knockout, however, enhances the heart and limb defects observed in the *Ror2* mutant (Nomi et al., 2001). The Rors are likely acting through the planar cell polarity pathway to contribute to early convergent extension of these tissues. *Ror2* also has additional expression in the primitive streak, while *Ror1* does not (Matsuda et al., 2001). This suggests there is some redundancy between the Rors, but *Ror2* may have a more dominant role in embryonic development.

cam-1 is the sole Ror in *C. elegans* and is most closely related to Ror2 (Bainbridge et al., 2014; Forrester et al., 1999). *cam-1* is not restricted to the nervous system, however, and affects the development of other tissues such as hypodermal and vulval polarity (Forrester et al., 2004; Green et al., 2007; Green et al., 2008a). Interestingly, *cam-1(gm122)*, a putative null allele, is homozygous viable (Hayashi et al., 2009; Green et al., 2008a; Kim and Forrester 2003; Forrester et al., 1999). The viability of the adult, the absence of multiple Ror homologues and the involvement of non-neural functions make *C. elegans* an ideal genetic system to study the function of Ror receptors during organ development.

1.14 *cam-1* expression and roles in *C. elegans* embryonic development

cam-1 is broadly expressed in the embryo starting at the 200-cell stage, and persists throughout the lifecycle. A larval *cam-1Ap::*GFP transgene expresses in many neurons,

hypodermal cells, body wall muscle, pharynx, and intestine (Forrester et al., 1999; Koga et al., 1999). A variety of phenotypes have been observed in *cam-1* mutants, such as locomotion defects, aberrant neuronal polarity and migration, and hypodermal and vulval polarity defects (Chien et al., 2015; Green et al., 2008a; Kim and Forrester, 2003; Yamamoto et al., 2011).

Three isoforms of CAM-1 have been identified. CAM-1A is the longest isoform, and contains a small additional exon upstream of the Ig domain. CAM-1B begins at the Ig domain. CAM-1C is the shortest isoform and is missing the Ig domain. Little has been reported on the function or expression of these isoforms (Kennerdell et al., 2009; Koga et al., 1999; Song et al., 2010). The mRNAs of these genes have varying expression patterns, and transgenic expression of each does not equally rescue Cam-1 mutant phenotypes (Kennerdell et al., 2009). For example, *cam-1* is required to properly direct nerve ring placement along the antero-posterior axis. *cam-1* loss-of-function causes the nerve ring to be anteriorly displaced. Transgenic expression of the *cam-1* cDNAs in the mutant background resulted in different rates of rescue. The *cam-1Ap::*CAM-1A and *cam-1Bp::*CAM-1B transgenes rescued the nerve ring defect, while *cam-1Cp::*CAM-1C did not. A *cam-1Ap::*CAM-1C construct did significantly rescue the defect, however. The differing localizations and rescuing abilities of these isoforms is probably because *cam-1* is involved in multiple genetic pathways and regulated in parallel by distinct regulatory elements.

Many *cam-1* alleles, including the kinase domain deletion *ks52*, confer a weak reversal of polarity in the hypodermis and vulva (Green et al., 2008a). Genetic analysis showed that CAM-1 works in the same pathway as EGL-20/Wnt and VANG-1/Van Gogh to establish vulval ground polarity. This is similar to establishing planar cell polarity in vertebrates, where Wnt5a acts as a permissive cue and induces a Ror2 and Vangl2 complex to form. Depending on the dose of Wnt

and the position of a cell in the Wnt5a gradient, Vangl2 will become phosphorylated multiple times along its Ser-Thr rich domains. The increase in Vangl2 activity causes the field of cells in the developing mouse limb bud to polarize along the proximal-distal axis (Gao et al., 2012). The more distal a cell is, the higher the dose of Wnt it receives. Consequently, the more Ser-Thr sites on Vangl2 that will be phosphorylated. As a result, cells having distal-most identities are most likely to be stimulated to proliferate and elongate the limb.

Other processes such as QL neuroblast migration, however, do not appear to require the kinase domain or any intracellular portion of the receptor (Kennerdell et al., 2009; Kim and Forrester, 2003). As previously mentioned, two modes of action have been proposed for CAM-1. One is simply for the extracellular portion of the receptor, namely the CRD, to bind and sequester Wnts from the surrounding environment and modulate local ligand concentrations for other receptors (Green et al., 2007; Kim and Forrester, 2003; Modzelewska et al., 2013). Other functions may utilize the full receptor, where binding an extracellular ligand mediates an intracellular function of CAM-1. The intracellular function may be a kinase-dependent signalling cascade or a kinase-independent scaffolding function (Chien et al., 2015; Song et al., 2010) . In a scaffold function the types of recruited proteins may be important. It has been suggested that CAM-1 clusters receptors to specific regions of cells. For example, CAM-1 has been shown to cluster acetylcholine receptors at neuromuscular junction synapses (Francis et al., 2005). Although this function is probably derived from MuSK, it is possible that CAM-1 has adopted this function for localizing components in other cell types. CAM-1 could complex with other receptors, such as Frizzleds, which may instead provide the intracellular activity (Jensen et al., 2012; Song et al., 2010). How these modes of Ror function apply to pharynx development, particularly retrograde extension of the g1p cell, will be explored by the Kormish laboratory.

The tools and techniques developed for *C. elegans* have been invaluable to understanding these complex processes such as cell signalling and cell coordination within an organ. Genetic screening, for example, can utilize all of the previously mentioned tools. A forward genetic screen exposes worms to a mutagen, such as ethyl methanesulfonate (EMS), and are screened with various types of microscopy to find desired phenotypes or defects (Brenner, 1974). The causative mutation could then be identified by a combination of genetic mapping and sequencing. Reverse genetic screening, on the other hand, would target specific genes or sequences and examine the resulting phenotype(s). Gene targeting is most commonly accomplished with RNA interference, which delivers engineered dsRNA sequences into the worm to knock-down gene expression (Fire et al., 1998). dsRNA can be delivered through microinjection, soaking, or feeding. Feeding involves supplying worms with bacteria that produce dsRNA (Timmons et al., 2001). Libraries of bacterial strains that target individual genes have been created, allowing for systematic targeting of nearly the entire genome (Kamath et al., 2003).

Forward genetic screening can be advantageous over reverse screening in discovery-based inquiries. Forward screening allows for unbiased surveying of the entire genome. Mutagenesis can, in theory, target any nucleotide or sequence in the genome. There are a variety of mutations and changes in gene function that can be generated. Single base-pair changes, insertions, deletions, and rearrangements are all possible changes that can be directly incorporated into the genome (Kutscher and Shaham, 2014). This can result in strong loss-of-function and even null/complete loss-of-function. In reverse genetic screening targets must be pre-chosen. RNAi only knocks down gene function rather than completely removing it. In some cases, RNAi does not confer a phenotype or may not reliably and specifically target a sequence (Qu et al., 2011).

Additionally, some cells such as neurons and pharyngeal gland cells are resistant to RNAi (Asikainen et al., 2005; Timmons et al., 2001). These issues could lead to misinterpretation of a gene's contribution to a function and exclude targeting of important factors during experimental design.

1.15 Preliminary Results

To understand what factors and pathways in addition to *cam-1* are contributing to gland cell migration, a forward genetic screen was conducted. The screening work described in this thesis is a continuation of a project started in the lab by Onoruza Atta and Maxwell Berg. They previously screened 350 haploid genomes and isolated 17 strains with gland cell placement defects.

1.16 Study Aims

The overarching goal of this study is to identify the molecular mechanisms that control pharyngeal g1P gland cell retrograde extension during *C. elegans* embryonic development. The focus of my graduate work is to understand how under-migrations of g1P occur during pharyngeal development. The specific objectives of my research were the following:

1) *Completion of a forward genetic screen to identify genes required for retrograde extension of the g1p gland cell.*

   Strains isolated from the forward genetic screen were given preliminary descriptions of phenotypes. Mendelian segregation, penetrance, and lethality were described for some candidates.

2) *Characterization of Cam-1 like alleles including chromosome linkage, allele complementation, and causative mutation identification through whole genome sequencing.*

   Homozygous viable mutations with under-migration defects were selected for descriptive analysis. Preliminary snip-SNP genetic mapping was used to determine chromosome linkage of associated alleles. Alleles isolated to the same linkage group were tested for complementation. A sub selection of mapped candidate strains was prepared and sent for Illumina HiSeq 2500 paired-end whole genome sequencing. Data sets were analyzed through a combination of the Galaxy platform and a custom pipeline developed by our collaborators at UBC as a part of the gene knockout consortium.

3) *Refined physical mapping of two mono-allelic strains and confirmation of alleles in complementation groups.*

   Strains that were not sequenced but belonged to a complementation group were prepared for Sanger sequencing to confirm the nature of the lesion. Strains that were whole genome sequenced but did not belong to a complementation group had more refined snip-SNP mapping performed to narrow down their genetic interval and therefore shortlist candidate genes.

4) *RNA interference to test for genetic interactions between cam-1 and ham-1 during retrograde extension.*

   Two alleles of *ham-1* were identified from the above objectives. Because *ham-1* mutations produced a phenotype similar to *cam-1*, RNAi was used to knock down the opposing gene in each mutant background, creating *cam-1; ham-1(RNAi)* and *ham-1;*

cam-1(RNAi) double mutants. This allowed for g1P quantification and genetic analysis to determine if the two genes function in the same genetic pathway.

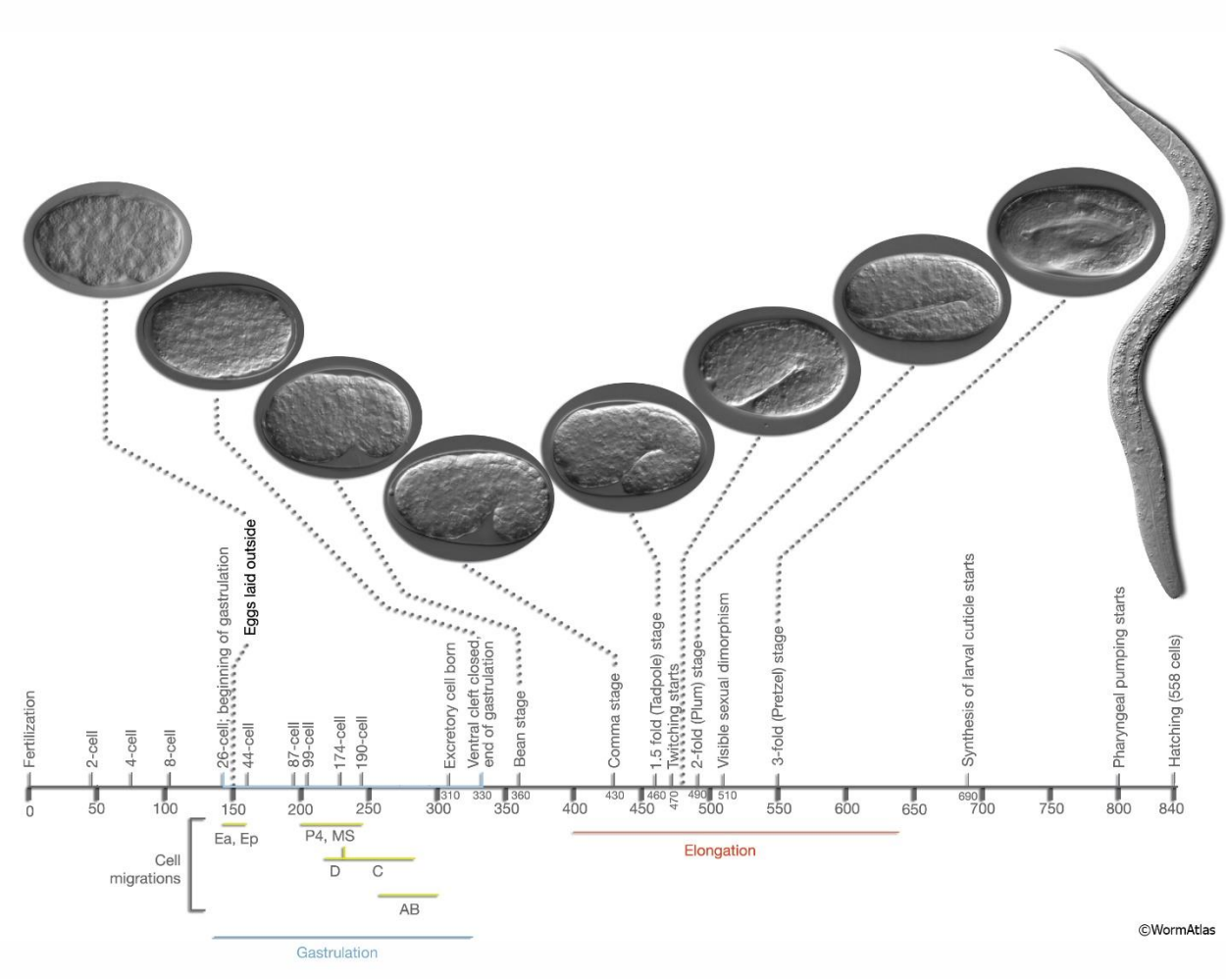

Figure 1.1 Stages of *C. elegans* embryonic development. At 22˚C, it takes approximately 840 minutes to progress from a single-cell zygote to hatched L1 larva. Numbers below the horizontal line represent time in minutes post-fertilization. Number of nuclei, stage name, and/or developmental milestones are indicated above the timeline. Gastrulation (blue bar) begins at the 26-cell stage and ends around 330 minutes. Green bars indicate when precursor cells migrate during gastrulation. Elongation of the embryo (red bar) occurs between 400 and 650 minutes. Figure based on von Ehrenstein and Schierenberg, 1980; Sulston et al., 1983; Wood, 1988; Bucher and Seydoux, 1994; Chin-Sang and Chisholm, 2000. In *WormAtlas* (Hall et al., 2009). doi:10.3908/wormatlas.1.1 http://www.wormatlas.org

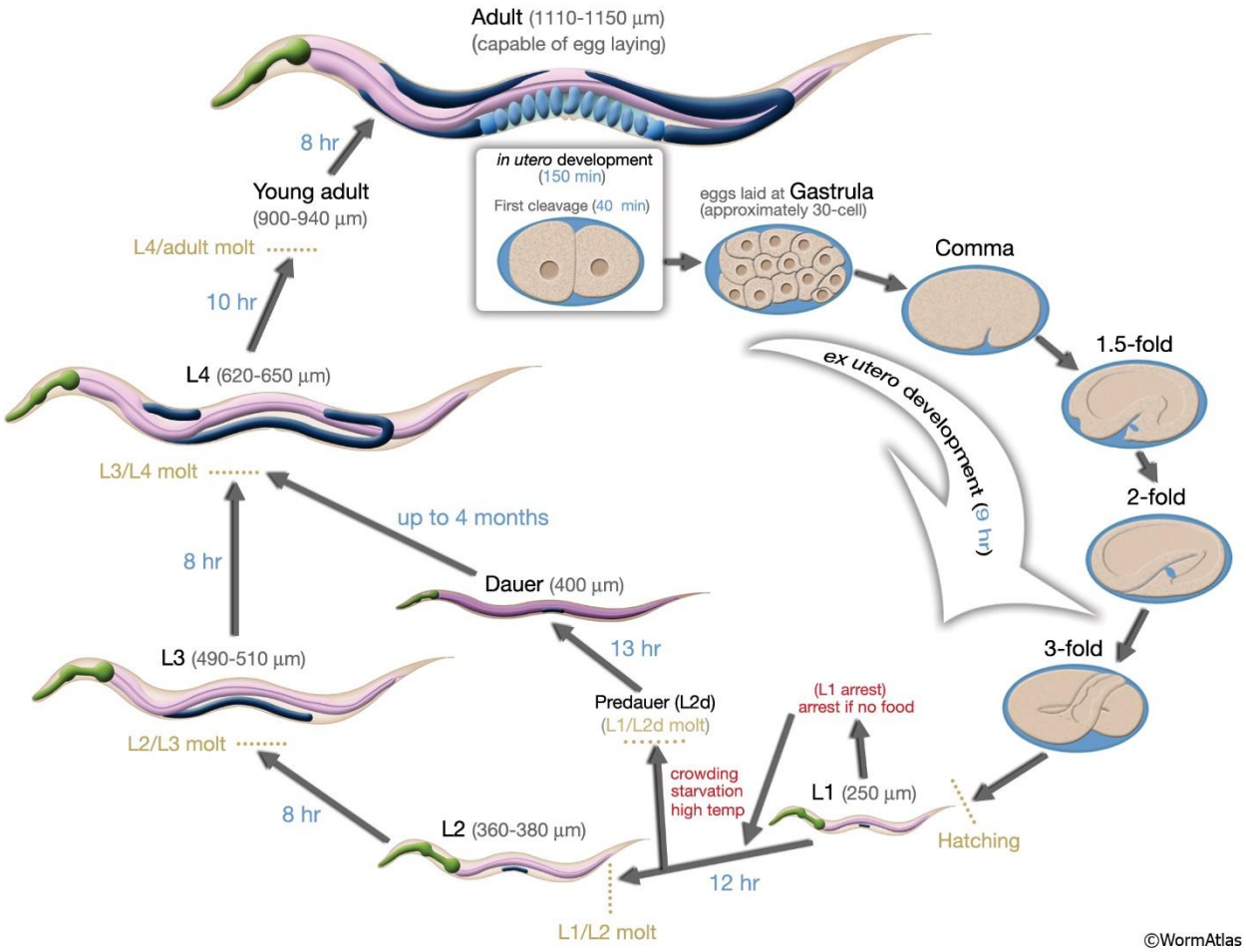

Figure 1.2 The *C. elegans* life cycle at 22˚C. Embryonic milestones are pictured in blue ovals. Blue text in brackets accompanying embryonic stages refers to minutes or hours post-fertilization. L1 through L4 are larval stages of growth. The length of the animal in µm is given beside each stage name. Blue text by arrows indicates how long is spent at each stage. Yellow text indicates the molt that occurs after each larval stage. Red text indicates alternative developmental options. If there is no food source present after hatching, L1 larvae can arrest development. If the environment becomes crowded, depleted of food, or too high in temperature, L1 larvae can enter a dormant dauer stage. Both arrested L1 and dauer larvae can return to normal development when environmental conditions improve. In *WormAtlas* (Altun and Hall, 2006). http://www.wormatlas.org

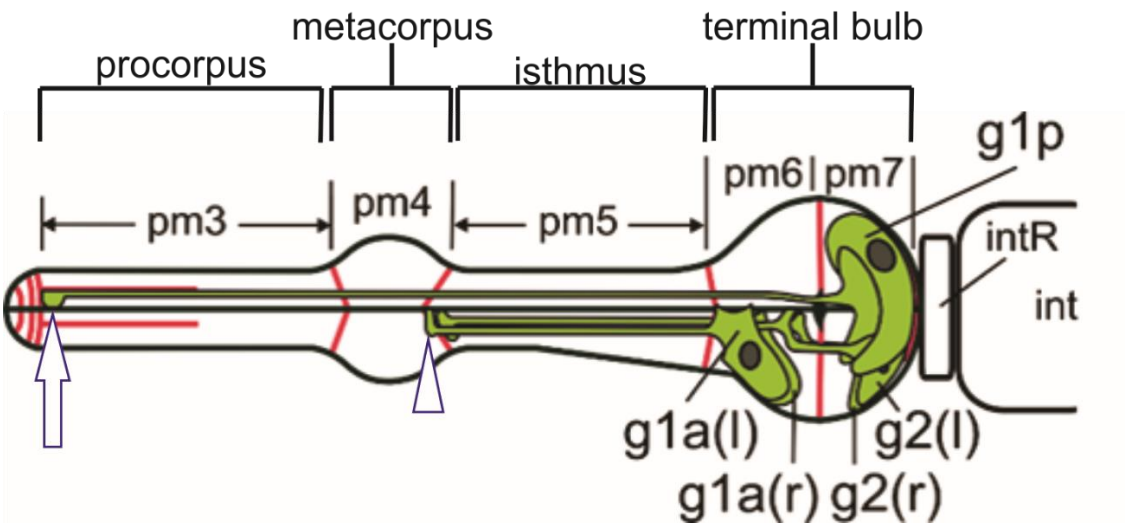

Figure 1.3 Cartoon representation of the mature *C. elegans* pharynx, with emphasis on the muscles and gland cells. Lateral view, left is anterior. The pharyngeal muscle (pm3-7) sections are in direct contact with the gland cells. Muscle boundaries are defined by red lines. The five gland cells are in green. g1P is the large dorsal gland cell that uses retrograde extension to form its long cytoplasmic projection. Its cell body is born at pm3 and migrates posteriorly to reside in pm7. The open arrow points to the g1P anchor. The g1A and g2 cells are ventral and in left-right pairs. The open arrowhead points to the g1A anchors. The g1A cell bodies lie in pm6, while the g2 cell bodies lie in pm7. intR (pharyngeal-intestinal valve), int (intestine). Image provided by Dr. Jay Kormish.

2 Materials and methods

2.1 Strains and Maintenance

Caenorhabditis elegans strains were maintained according to conditions previously published (Brenner, 1974). In brief, nematodes were cultured on Nematode Growth Medium (NGM) containing a lawn of OP50 strain of *Escherichia coli*. Plates were stored at 15˚C unless otherwise stated. The strain used in mutagenesis was GD282 *ivIs17* [*phat-1*::YFP, *elt-2*::mTomato *rol-6(su1006)*] V. A complete list of strains generated and isolated from the forward screen is available in appendix tables A1-A5. The mutagenized strains selected for further study are included in Table 2.1. CB4856 (Hawaiian) wildtype males were used for SNP mapping. Mutagenized strains were outcrossed to N2 wildtype males three times to remove background mutations. At each N2 outcrossing strains were given a unique strain name as follows: JDK117 *kor42* I; *ivIs17* V, JDK119 *kor5* IV; *ivIs17* V, JDK114 *kor15* IV; *ivIs17* V, JDK118 *ivIs17* V; *kor44* X. The strain used in *cam-1* complementation tests was NG2615 *cam-1(gm122)* II, and the strain used in *cwn-2* complementation tests was GD309 *ivIs12* [*phat-1::*YFP, *elt-2*::GFP, *rol-6 (su1006)*] II, *cwn-2(ok895)* IV.

Table 2.1 Mutagenized strains selected for characterization in this thesis.

Strain	Genotype
JDK32	*kor5* IV; *ivIs17* V
JDK56	*kor13*; *ivIs17* V
JDK58	*kor15* IV; *ivIs17* V
JDK76	*kor22* IV; *ivIs17* V
JDK77	*kor23* V; *ivIs17* V
JDK91	*kor37* II; *ivIs17* V
JDK96	*kor42* I; *ivIs17* V
JDK97	*kor43* IV; *ivIs17* V
JDK98	*ivIs17* V; *kor44* X
JDK99	*kor45* IV; *ivIs17* V
JDK100	*kor46* V; *ivIs17* V
JDK103	*kor49* IV; *ivIs17* V
JDK106	*kor52* IV; *ivIs17* V

2.2 Mutagenesis and forward genetic screen for pharyngeal gland cell migration

A forward genetic screen was conducted under conditions previously published (Brenner, 1974). Briefly, GD282 worms at the L4 stage were exposed to a solution of 25mM ethyl methanesulfonate (EMS) at 20˚C for four hours. Individual hermaphrodites (the P0 or parental generation) were picked to plates and kept at 20˚C until the appearance of the F1 generation. Six fertile F1 worms were individually plated, allowed to self-fertilize and lay eggs. The succeeding F2 progeny, grown at 20˚C, were screened for gland cell migration defects using a Leica M205C stereo microscope and XCite series photo illumination system with 120Q fluorescent light source or a Zeiss Axiozoom V16 with Illuminator HXP 200 C. Candidate strains were assigned a JDK strain number and *kor* allele designation. F2 mutants were plated for at least one more generation to confirm the mutation was heritable and to estimate penetrance. Mutations are assumed to be recessive and strains were isolated in a homozygous state where possible (Brenner, 1974).

For this project, a subset of these mutant strains were selected that displayed a phenotype similar to mutations in the *cam-1* pathway. These strains displayed an under-migration of the

g1P cell and were homozygous viable. Strains that were fertile and had no major morphological defects were preferentially studied.

2.3 Identifying *cam-1*-like mutations obtained from the forward genetic screen

2.3.1 SNP Mapping- Outcrossing

A modified genetic mapping protocol was used to determine the approximate location of mutations (Davis et al., 2005). The Hawaiian strain possesses a unique set of Single Nucleotide Polymorphisms (SNPs) when compared to the common N2 strain. This protocol differentiates between Hawaiian and N2 SNPs by analyzing DraI snip-SNP sites: whether the DraI restriction enzyme can recognize and cut that site indicates whether it is an N2 polymorphism, or a Hawaiian one. By mating Hawaiians with mutant strains originating from an N2 background, we can observe where recombination takes place in the re-isolated mutants, and compare the biallelic distribution with the non-mutant. This will indicate whether that area of the genome is linked to the GD282/N2 causative mutation. This technique can use Bulk Segregant Analysis (BSA) (Wicks et al., 2001), and makes use of populations of recombinants for single-step genetic mapping. This mapping can be done at a broad chromosomal level (picking only a few representative sites on each chromosome) and at a more refined "interval" level (analyzing eight SNP sites spread along a chromosome), depending on which primer sets are selected.

Ten mutant hermaphrodites and twenty Hawaiian males were plated and left overnight at 15˚C. These P0 hermaphrodites were transferred to individual plates, and the resulting F1s that appeared heterozygous (that is, possessing the *ivIs17* insertion but having normal gland cell migration) were also plated individually. The F2 progeny were then collected for analysis: approximately 50 worms with the mutant gland cell phenotype were resuspended into 50µL of single worm lysis buffer (SWLB; contained 1mM Tris-HCl pH 8.3, 5mM KCl, 0.25mM $MgCl_2$,

0.45% NP40, 0.45% Tween 20, 0.02% gelatin, and 0.01mg of proteinase K). 50 worms that had a non-mutant gland cell phenotype were collected the same way in a separate tube.

2.3.2 SNP Mapping - semi-quantitative analysis using unique restriction enzyme recognition sites and recombination

2.3.2.1 Polymerase Chain Reaction to amplify snip-SNP sites

Worms in SWLB were stored at -80˚C for a minimum of 1 hour. Lysis was completed by subjecting tubes to the following thermocycler program: 60˚C for 1 hour, 95˚C for 15 minutes, 4˚C hold. Lysed DNA was carried forward to PCR with the following components per 25 µL reaction, according to manufacturer protocol (Thermo Scientific #EP0404): 1x PCR Buffer with KCl, 14µM $MgCl_2$, 200nM primers, 1 unit of Taq polymerase, 2mM dNTPs, molecular grade water, and 2 µL lysed worm mixture. The thermocycler program used was 94˚C for 2 min, [94˚C for 30 seconds, 56˚C for 30 seconds, 72˚C for 1 min] repeated 30 times, 72˚C for 10 minutes, and a final 21˚C hold. Primers used to amplify SNP sites are included in Table 2.2. (Davis et al., 2005). To roughly map to a chromosome, the center-most primer from each chromosome was used. Where indicated in the results a more refined interval mapping along a specific chromosome was also attempted.

2.3.2.2 DraI restriction digest

PCR products were then carried through to an enzymatic digest (Thermo Scientific #FD0224). Each 20µL reaction contained 9.5µL molecular grade water, 1x Fast Digest buffer, 0.5 units of Fast Digest DraI, and 8µL of PCR product from the preceding reaction. Reactions were incubated in a 37˚C water bath overnight or for at least eight hours.

2.3.2.3 Resolving restriction digest patterns with gel electrophoresis

1.8% agarose gels were constructed with 1x TAE (2M Tris, 1M glacial acetic acid, and 0.05M EDTA pH 8.0) and 1x SYBR gold (Invitrogen), and covered with 1x TAE running buffer. A Bio Rad Powerpack Basic was used to apply 40V for five minutes, and then turned to 80V for approximately 90 minutes. The digested DNA products were visualized with UV light from a Bio Rad Universal Hood II gel doc. The digest patterns for the mutant DNA were analyzed according to predicted band sizes (Davis et al., 2005), and the area(s) that retained a predominantly N2 pattern were assumed to be tentatively linked to the causal mutation.

2.3.3 Complementation testing mapped mutant strains

Mutant strains that mapped to the same chromosome based on results from the SNP mapping experiments were subjected to complementation crosses to determine how many different genes existed in the Cam-1-like sub selection of strains. In the interest of time, heterozygous males were of each mutant strain were used in each cross. These were generated by mating mutant hermaphrodites with N2 males; the mature F1 males from the following generation were used in the downstream complementation crosses. This eliminated multiple time-consuming steps of backcrossing, screening, and culturing unique male strains to ensure the homozygous recessive mutations had been re-isolated. Furthermore, *cam-1* loss-of-function mutations possess tail defects and gonad migration defects. Heterozygous males ensured that defects that affected male fertility could be avoided. These heterozygous males were mated to hermaphrodites of another mutant strain. About 30 heterozygous males were plated with six young adult homozygous hermaphrodites and left overnight at 15˚C. The hermaphrodites were transferred to individual plates the next day (“day one” plates), and then transferred again every 24 hours for the next two days (“day two” and “day three” plates). The progeny on the day two

and day three plates were scored about six days after the hermaphrodite had been transferred. To make sure only plates where mating had taken place were analyzed, only plates that contained male progeny were counted. Both the number of worms with an under-migration defect and the total number of progeny on a plate were recorded. These numbers were then expressed as a ratio to estimate the occurrence of mutant phenotypes from each mating event. Because males were heterozygous, the maximum occurrence of homozygous recessive mutants on a plate would be half of the entire population. Not all strains have complete penetrance, so actual proportions were usually lower. If a cross resulted in 10-50% of progeny retaining the under-migration phenotype, it was considered a failure to complement, and the two strains were placed in the same complementation group. Results that were variable or had a penetrance lower than this were indicated and required a more involved analysis.

2.3.4 Preparation of mutagenized strains for whole genome sequencing

2.3.4.1 Strain selection for whole genome sequencing

To maximize the chances of confirming multiple novel genes, a variety of strains were selected to be sent for whole genome sequencing: two strains from LGIV complementation group A (*kor45* and *kor49*), two strains from LGIV complementation group B (*kor5* and *kor15*), one strain from chromosome I (*kor42*), and one strain from chromosome X (*kor44*). Two strains from each chromosome IV complementation group were chosen because the candidate gene list from downstream bioinformatics analysis could easily be streamlined simply by finding lesion in genes common to both strains. *kor45* was also of interest because it has a constitutive dauer phenotype. The GD282 pre-mutagenesis background and CB4856 Hawaiian wildtype strain were also sequenced to serve as controls to subtract background mutations from mutagenized and outcrossed strains.

2.3.4.2 Strain preparation for whole genome sequencing

A protocol and free automated Galaxy pipeline were developed by the Hobert lab to rapidly map and analyze mutations from *C. elegans* whole genome sequencing data. Termed CloudMap, this strategy uses a variant-based method that also takes advantage of the Hawaiian polymorphic strain and allows for simultaneous mapping and sequencing (Minevich et al., 2012). Prior to collection and preparation for whole genome sequencing, the six selected mutant strains were subjected to Hawaiian outcrossing. At least 50 F2 mutants were individually plated on 92x16mm NGM plates seeded with OP50 *E. coli*., and their subsequent progeny (F3-F4 generations) were grown at 20˚C. Plates used for collections were made with molecular grade agarose to prevent inhibition of downstream *in vitro* reactions. Crude agar extracts used in other plate preparations may interfere if carried over into sensitive molecular reactions.

Just prior to starvation, the populations of a strain were collected and pooled. Worms were washed off plates with M9 buffer and pooled into a 50mL conical tube. To remove as much remaining bacterial contamination as possible, worms were allowed to settle/pellet by gravity for 10-15 minutes. The buffer was carefully removed and replaced with a fresh aliquot; this pellet-and-wash step was repeated at least three times for each tube. The final pellet was transferred to a 1.5mL Eppendorf tube, spun down for 30 seconds, and had supernatant aspirated off. Tubes were left at -80˚C for at least 24 hours before proceeding with genomic extractions.

2.3.4.3 Genomic extraction for whole genome sequencing

Genomic extractions were completed with the Gentra Puregene Tissue Kit (Qiagen #158667). The protocol used was developed in part from a supplementary nematode protocol from the Andersen lab [http://andersenlab.org/Protocols/PurificationDNAfromnematodes.pdf]. A subsample of the previously-pooled collection (approximately 360 mg) was added to 6 mL of kit

lysis solution, and then divided into three 15 mL sonication tubes for use with a Bioruptor Pico sonication device (Diagenode). The tubes were subjected to a sonication cycle of 1s on, 30s off for 15 cycles. The contents of each tube were combined into a standard 15mL conical tube, and 30µL of kit Proteinase K was added and incubated in a 55˚C water bath moving at 100 rpm for at least three and a half hours, or until worm bodies were fully lysed.

The mixture was then divided into six 2mL screwcap vials and allowed to cool to room temperature. 5µL of RNase A was added to each tube and incubated in a thermomixer at 37˚C and 300rpm for at least one hour. Tubes were placed on ice for five minutes and 350µL of Protein Precipitation solution was added to each, returned to ice for five minutes, vortexed for 20s, and returned to ice for another five minutes. Tubes were centrifuged for 5 minutes at 16,000 xg. Supernatants were transferred to new screwcap tubes, and 750µL of isopropanol was added to each and inverted fifty times. Precipitated nucleic acids were pelleted by centrifuging for 1.5 minutes at 16,000 xg and the resulting supernatant was removed. 1mL of 70% ethanol was added and inverted 50 times, then centrifuged for 1.5 minutes at 16, 000 xg. Supernatant was removed again, and tubes were allowed to air dry at room temperature for about five minutes. The DNA pellet was re-suspended in 80µL of a low molarity hydration solution (10mM Tris and 0.1mM EDTA, pH 8.0). The kit rehydration solution was not used because it had a higher concentration of EDTA which, according to Genome Quebec specifications, could interfere with sequencing reactions.

2.3.4.4 gDNA assessment, clean-up, and submission

Completed DNA preparations were first assessed using a NanoDrop 2000 (ThermoFisher). Concentration and 260/280 OD ratio were initially satisfactory (at least 80ng/µL and between 1.7-1.8, respectively), however the 260/230 OD ratios were consistently

suboptimal, ranging from 1.5-1.7 when sequencing specifications required at least 1.8 (Genome Quebec). An extra sodium acetate and ethanol precipitation step were developed post-extraction to remove the oligosaccharides that seemed to be contaminating the gDNA. Four volumes of 3M sodium acetate (pH 5.5) were added to the DNA solution and mixed; two volumes (relative to the new volume post-acetate addition) of ice-cold 95% ethanol was then added and incubated on ice for 5-30 minutes. Tubes were centrifuged for 15 minutes at 20,000 xg, washed with 1mL of 70% ethanol and re-centrifuged for 10 minutes at 10,000 xg. Wash and spin-down was repeated twice, and the pellets were dried at room temperature for five minutes. DNA was resuspended in the low molarity hydration solution according to desired concentration (based on prior spectrophotometer estimates, this was usually between 75-100µL). Sodium acetate ethanol precipitations and NanoDrop readings were repeated until acceptable concentrations and OD ratios were obtained.

Double-stranded DNA concentrations were also estimated with a Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher #P7589). A small amount of each sample was also run on a 1% agarose gel to verify the DNA purity and quality. Each strain's preparation was then aliquoted into a 96-well plate according to the Genome Quebec Massively Parallel Sequencing Services user guide, and submitted for shotgun library preparation (NebNext Ultra II FS DNA library prep, New England Biolabs # E7103) and sequencing on a single lane of the Illumina HiSeq 2500 Paired End 125bp platform.

2.3.5 Whole genome sequencing: bioinformatic analysis

The resulting raw reads for each strain were downloaded from Genome Quebec's Nanuq project page. These were uploaded to Galaxy and entered as the "Reads" files for the "CloudMap Hw Variant Mapping with WGS and Variant Calling Workflow" (Minevich et al., 2012). While

the alignments to the reference genome and some of the Hawaiian variant interval mapping outputs worked, there were also many limitations and issues with the workflow(s) and general free instance of Galaxy. This instance of Galaxy has since been significantly modified, and because of this CloudMap is no longer supported (Oliver Hobert/Hobert lab, personal communication).

A collaboration was initiated with Stephane Flibotte and the Moerman *C. elegans* gene knockout lab (University of British Columbia) to efficiently identify candidate genes for each strain. Reads were mapped to the *C. elegans* reference genome (version WS230) using the Burrows-Wheel Aligner (Li and Durbin, 2009). SNPs were identified and filtered with the SAMtools toolbox (Li et al., 2009), including subtracting variants present in the Hawaiian or GD282 background strain. Genes that fit the filtering criteria were then annotated with a custom Perl script using gene information from WormBase (www.wormbase.org), and summarized in a candidate gene list for each strain.

2.3.6 Sanger sequencing to confirm identity of other mutant strains

To confirm the *cwn-2* lesion in the remaining strains in the LGIV complementation group A (*kor22, kor43,* and *kor52*) and a *cam-1* lesion on chromosome II (*kor37*), these genomic regions were isolated and sequenced using Sanger sequencing. Worms from each strain were collected in Single Worm Lysis Buffer and a crude lysate was obtained as described in section 2.3.2.1.

Nested primer sets were designed and tested to isolate the entire genomic sequence of each gene through PCR (Tables 2.3, 2.6, 2.7, and 2.8). The first nested reaction consisted of the previously described 25µL PCR recipe, using the crude lysate as DNA template and the "outer" primers. The thermocycler program used was 94˚C for 2 min, [94˚C for 30 seconds, 54˚C for 35

seconds, 72˚C for 3.5 minutes] – repeated 35 times, 72˚C for 10 minutes, and a 21˚C hold. These products were diluted in a two-fold series, and used as templates for the second nested reaction, using the "inner nested" primers and a Thermocycler program of 94˚C for 2 minutes, [94˚C for 30 seconds, 54˚C for 30 seconds, 72˚C for 3.5 minutes] – repeated 30 times, 72˚C for 10 minutes, and a 21˚C hold. These products were purified using the GeneJET PCR Purification Kit (ThermoFisher #K0702). The only modification made to the supplied protocol was increasing the amount of binding buffer 1.5-2 times higher than recommended.

When sample concentration or OD ratios were low, a sodium acetate ethanol precipitation was performed. An equal volume of 3M sodium acetate (pH 5.5) was added to the sample and mixed well. Two volumes of 90% ethanol were added and the tube was incubated on the bench for three to five minutes. The sample was centrifuged at 20,000 xg for 10 minutes, aspirated, washed with 250µL of 70% ethanol, inverted a few times, and centrifuged at 20,000 xg for 15 minutes. This step was repeated for a total of two wash steps. The pellet was resuspended in 30-50µL of 10mM Tris buffer (pH 8.0).

Satisfactory samples were prepared according to specifications and submitted to the Manitoba Institute of Cell Biology/Eurofins MWG Operon DNA sequencing service. The primers included in these sequencing reactions were custom and designed to sequence approximately 500 base pair fragments of the appropriate gene with some anticipated overlap (see Tables 2.3, 2.6, 2.7, and 2.8).

2.3.7 Interval mapping to reduce LGI and LGX candidate genes

Because the sequencing results for the strains mapping to chromosome I and X returned an overwhelming list of candidate genes, interval SNP mapping was performed according to the above mapping workflow (Davis et al., 2005). Instead of pooling the F2 progeny, a subset of

progeny from each F2 mutant was collected in SWLB during whole-genome sequencing preparation. Each set of progeny was used in its own reaction, revealing the location of single recombination events. From this data, approximate recombination frequencies and genetic locations for the causative mutations were calculated.

2.4 RNA interference

2.4.1 Collections for RNA extraction

cam-1 and *ham-1* double-stranded RNA (dsRNA) were formed from freshly-isolated cDNA. The N2 strain was cultured on 150x20 mm plates, and mid-to-late embryos (*ie.*, those already laid) were collected. These stages have the highest expression of *ham-1* and *cam-1* mRNA according to the modENCODE FPKM expression data available on each gene's entry on Wormbase.

Just prior to starvation, plates were gently flooded with M9 buffer to dislodge the adults and larvae. This liquid fraction was aspirated and discarded for this experiment. Another aliquot of M9 was added and the desired embryos were scraped off the media using a sterilized glass hockey stick. This fraction was transferred to a 15mL conical tube. The tube was topped up with M9 to about 13mL, and 5% reagent-grade bleach was added to make up the remaining volume. Once the solution had turned clear, the tube was briefly spun down and the liquid removed; any remaining bleach was quenched by adding a few milliliters each of M9 and Bovine Serum Albumin (20mg/mL). Three spin-and-wash steps were then completed with more M9 buffer. After a final spin-down, the embryos were transferred to a 1.5mL microtube, and centrifuged at 15,000 xg for five minutes. The supernatant was removed and 50µL of RNAlater solution was added (ThermoFisher #AM7020). Samples were stored at -80˚C until ready for extraction.

2.4.2 RNA extraction

Extraction was completed using the RNeasy Micro kit (Qiagen #74034). 50µL of buffer RLT was added to thawed samples, which were then flash-frozen in liquid nitrogen and ground with a plastic pestle. About three freeze-thaws were required to homogenize the embryo material. An additional 300µL of buffer RLT was added, and the tube was then vortexed for five minutes. One volume of 70% ethanol was added and gently pipetted up and down. This solution was transferred in its entirety to an RNeasy spin column, and then centrifuged at 8,000 xg for 30 seconds, and then for an additional minute to ensure all the liquid had passed through the column. 350µL of buffer RW1 was added to the column and centrifuged at 8,000 xg for 30 seconds. A solution of 20µL DNase I and 70µL of buffer RDD was prepared, added to the column, and then incubated on the bench top for 15 minutes. 350µL of buffer RW1 was added and centrifuged at 8,000 xg for 30 seconds. 500µL of buffer RPE was added and then also spun down at 8,000 xg for 30 seconds. Two washes of 500µL 80% ethanol were applied and centrifuged at 8,000 xg for 2 minutes each. The lid of the column was opened and then spun down at maximum speed for three minutes to ensure sample was dry. The column was transferred to a 1.5mL microcentrifuge tube, 14µL of RNase-free water was added, and then centrifuged at maximum speed for one minute to elute the RNA.

2.4.3 cDNA synthesis

The RNA was immediately carried through to a cDNA synthesis reaction using the iScript Select cDNA synthesis kit (Bio-Rad #1708896). In a 0.6mL microtube, the following reaction was mixed: 4µL of 5x iScript reaction mix, 2µL of oligo$(dT)_{20}$ primers, 3µL of nuclease-free water, 10µL of extracted RNA, and 1µL of iScript Reverse Transcriptase enzyme.

This was incubated at 42˚C for 60 minutes, and 85˚C for five minutes. The solution was then separated into 5µL aliquots and stored at -80˚C.

2.4.4 Amplification of *cam-1* and *ham-1* cDNA into dsDNA

Primers were designed to amplify *cam-1* and *ham-1* cDNA through a nested PCR. The set of primers for the first reaction were designed to amplify about 970 base pairs of the target transcript. The second, nested primers amplified a product of about 850 base pairs; each primer also included the T7 polymerase binding site so that it would be incorporated into the DNA products.

The first amplification consisted of 1x PCR Buffer with KCl (Thermo Scientific #EP0404), 14µM $MgCl_2$, 200nM each of forward and reverse primers, 1 unit of Taq polymerase, 2mM dNTPs, and molecular grade water to 20µL. This was added to a 5µL aliquot of cDNA. The thermocycler program used was 95˚C for 5 min, [95˚C for 30 seconds, 58˚C for 35 seconds, 72˚C for 1 min] – repeated 35 times, 72˚C for 10 minutes, and a 21˚C hold.

The resulting product was diluted ten-fold, then used as template DNA for the second nested reaction. This was a 50µL reaction consisting of 1x PCR Buffer with KCl, 2mM $MgCl_2$, 0.2mM dNTPs, 0.2µM of forward and reverse primers, 1 unit of Taq polymerase, and 5µL of DNA template. The thermocycler program used was 95˚C for 5 minutes, [95˚C for 30 seconds, 58˚C for 30 seconds, 72˚C for 1 minute] – repeated 35 times, 72˚C for 8 minutes, and a 21˚C hold.

2.4.5 Amplification of L4440 control dsDNA

The multiple cloning site of pPD129.36, or L4440 plasmid, was used as the control for this RNAi experiment. This plasmid has endogenous T7 polymerase recognition sites

surrounding the cloning site, so a single PCR reaction was used to obtain the dsDNA product. This reaction followed the same 50µL reaction recipe as in the preceding section, however the thermocycler program was slightly modified to the following 95˚C for 5 minutes, [95˚C for 30 seconds, 57˚C for 30 seconds, 72˚C for 45 seconds] – repeated 30 times, 72˚C for 5 minutes, and a 21˚C hold.

2.4.6 Synthesis and recovery of *cam-1, ham-1* and L4440 dsRNA

The above *cam-1, ham-1,* and L4440 double-stranded DNA products were purified using the GeneJET PCR Purification Kit (ThermoFisher #K0702), following the slightly modified protocol mentioned above. Because the samples were not concentrated enough to be used in a transcription reaction, they were then taken through a 1.5M sodium acetate and ethanol precipitation step. After adding two volumes of 100% ethanol, the samples were left to incubate overnight at -20˚C. After centrifugation and a 70% ethanol wash step, the precipitated DNA was resuspended with molecular grade ultrapure water and stored at 4˚C overnight.

The concentrated samples were used as template for a double-stranded RNA synthesis reaction using the TranscriptAid T7 High Yield Transcription Kit (ThermoFisher #K0441). Each 20µL reaction consisted of 4µL 5x TrasncriptAid buffer, 8µL of NTPs mixture, and 2µL of TranscriptAid enzyme mix. The amounts of DEPC-treated water and dsDNA template varied between genes due to differing concentrations of template. At least 600ng of template DNA was added, and the DEPC-treated water was adjusted to ensure the reaction volume remained at 20µL. The reactions were mixed by pipetting, and then briefly spun down. They were incubated at 37˚C for 2.5 hours.

A white precipitate formed at the end of the incubation. Because this could be excess dsRNA product, the tubes were spun down and the supernatant was transferred to a separate

tube. The precipitate fractions were resuspended with 6µL of DEPC-treated water, and the six tubes were then carried through the next steps in parallel. 3µL of DNAseI was added and gently mixed by pipetting, and then incubated at 37˚C for 45 minutes. 2µL of EDTA (0.5M, pH 8.0) was added to stop the reaction and DNaseI was inactivated through incubation at 65˚C for 10 minutes.

Because phenol at the correct pH was not available, a chloroform-only RNA recovery was performed. The volume of each tube was brought up to 200µL with water, and 50µL of 5M $NaCl_2$ was added. 200µL of chloroform was added and inverted multiple times to mix, then centrifuged at 12,000 xg in 4˚C for 10 minutes. Two volumes of cold 99% ethanol were added, then incubated in -20˚C overnight. The next day, samples were centrifuged at 12,000 xg in 4˚C for 10 minutes, and supernatant removed. 1mL of cold 70% ethanol was used to wash the pellet, and then centrifuged again. The pellet was air-dried and resuspended in 20µL of ultrapure water. Samples were diluted 8, 16, and 32-fold due to very high concentrations. RNA integrity was examined by running samples on a 1% agarose gel.

2.4.7 dsRNA injection

The *cam-1*, *ham-1*, and L4440 dsRNA products were each injected into the GD391*cam-1(gm122); ivIs17* and JDK114 *ham-1(kor15); ivIs17* and control GD282 *ivIs17* strains. Strains were cultured every second day and kept at 15˚C prior to injection. Each dsRNA species was injected into healthy young adult hermaphrodite gonads at a concentration of 1000ng/µL using a microinjection system. The worms were allowed to recover on NGM OP50 plates at 20˚C for six hours, and then picked over to a new plate and left overnight at 20˚C. About sixteen hours later, they were picked over once more. The progeny on the second set of plates (six hours post-

injection) were allowed to hatch over the next 48 hours, and were then screened and scored for g1P gland cell placement.

2.5 Slide/mount preparation for microscopy

L1 larvae were washed off the appropriate plate(s) with M9 buffer and into a glass well. Larvae were then transferred to a glass well slide in minimal buffer volume. Worms were anesthetized with either 500mM tetramisole hydrochloride or 50mM sodium azide. The volume of anesthetic varied due to some strains being levamisole-resistant; it approximately worked out to be a 4:1 ratio of M9 buffer to anesthetic. Once worms were rod-like and no longer moving, they were mounted on a fresh 2% agarose pad and covered with a cover slip.

2.6 Quantification and image capture

Quantification of g1P placement and image capture were completed on a Zeiss Axio Imager Z1 microscope with 100x oil immersion objective. Gland cell location was assigned a score based on the posterior-most pharyngeal muscle section it reached. If the cell body was located in two sections, it was assigned a score of 0.5. For example, a gland cell located in pm4 was given a score of 4, but if it straddled both pm4 and pm5 it was given a score of 4.5. DIC images and Z-stacks of fluorescent glands were taken of representative defects for each background/genotype using the Axiovision application. Exposure time for fluorescence was between 300 and 1000 ms. Stacks were edited in ImageJ using the stack and z-projection functions; brightness and contrast adjustments were made to enhance the projections in final images.

2.7 g1P distribution analysis

Select genotypes were selected to be compared and statistically analyzed for difference in g1P distributions. This was calculated with the use of the Mann-Whitney U test, with some manual calculations in Excel. Because sample sizes were large, the U* or z-score was used in place of the U statistic, and was calculated using the corrected standard deviation due to ties in ranks. A significance of $\alpha=0.05$ was used. Bar graphs were generated in Excel. The text for graph labels was edited in Corel Draw. Box plots were generated in Excel. The MIN, QUARTILE, MEDIAN, and MAX functions were used to calculate the minimum values, first and third quartiles, median, and maximum values of each population, respectively. The boxes are modified stacked bar graphs where the bottom bar is the first quartile, the middle bar is the first quartile subtracted from the median, and the top bar is the median subtracted from the third quartile. The bottom whisker was the value of the minimum subtracted from the first quartile. The top whisker was the value of the third quartile subtracted from the maximum. The text for boxplot labels was edited in Corel Draw.

Table 2.2 Oligonucleotides used for chromosomal linkage mapping. F=Forward, R=Reverse.

Designation	Sequence	Comments
D1007 F	AAAATATCAGGAAAGAGTTTCGG	Davis et al., 2005
D1007 R	TTTAAAGATTAAGGGTGGAGCG	Davis et al., 2005
T24B8 F	TCAAAAACTTACAATCAATCGTCG	Davis et al., 2005
T24B8 R	CCAGAAAATCTGCACAGAAGG	Davis et al., 2005
F56C9 F	AAAAATACATGTCTACACAACCCG	Davis et al., 2005
F56C9 R	TTTCTTATCACTGTGCAGTCTTACC	Davis et al., 2005
E03H12 F	AAAATGGGAAGCGTACCAAA	Davis et al., 2005
E03H12 R	TGCTTGTAGCGTTTCCAAGA	Davis et al., 2005
VC5 F	AGAAATGATCCGATGAAAAAGC	Davis et al., 2005
VC5 R	CCGATAGTGTTCATAGCATCCC	Davis et al., 2005
F11A1 F	AGCAACAAACAATGCAACTATGG	Davis et al., 2005
F11A1 R	TAAACAAGAGGGTACAAGGTATCG	Davis et al., 2005

Table 2.3 Oligonucleotides used in amplifying and Sanger sequencing *cam-1C* genomic sequence in *kor37*. F=Forward, R=Reverse.

Designation	Sequence	Comments
oJDK387	ACCCACCTTTGACTGTAGATTC	F outermost for amplifying *cam-1C*
oGD1022	ATTCCCTCCGGCACCTCTCTTT	F nested for amplifying *cam-1C*
oJDK388	TCGTGGACGGTTACCAGACG	Sanger sequencing *cam-1C*
oJDK389	TGAGCACGGCGCTTCAGAG	Sanger sequencing *cam-1C*
oJDK390	GACACGTCATCAGAATGCAC	Sanger sequencing *cam-1C*
oJDK391	AGATTCGTGCAGTTGCGACATTTG	Sanger sequencing *cam-1C*
oJDK392	GGCATGGCTGAAACTACATAACA	R nested for amplifying *cam-1C*
oGD1023	CTGAAAGACAGGTCAAATTCCTAGCTG	R outermost for amplifying *cam-1C*

Table 2.4 Additional oligonucleotides used in LGI interval mapping of *kor42.* All other LGI intervals used original primers from Davis et al., 2005. F=Forward, R=Reverse

Designation	Sequence	Comments
oJDK416	TTTCTGGGATGCCAGATTTAT	Redesign Y71G12A F
oJDK417	TACCTCATGCTGCAACGAG	Redesign Y71G12A R
oJDK418	AGCGCGAATTTGAAAACAAT	Redesign Y71G12A Fb
oJDK419	TGCTGAGCAATCGAACTTGT	Redesign Y71G12A Rb
oJDK420	TATCATCGGATAAGCCATGAG	Redesign W03D8 F
oJDK421	AAGGCGCATATACAGCAGTC	Redesign W03D8 R
oJDK422	CAAGGCAGCACAGAAATGGT	Redesign F58D5 F
oJDK423	GCACACGAGAATCCACTGTT	Redesign F58D5 R
oJDK424	GTCAACATGTCCCATATTGTCA	Redesign T06G6 F
oJDK425	CATTGATTTCTGGCTGCTGAG	Redesign T06G6 R

Table 2.5 Additional oligonucleotides used in LGX interval mapping of *kor44.* All other LGX intervals used original primers from Davis et al., 2005. F=Forward, R=Reverse

Designation	Sequence	Comments
oJDK426	ATGATTGGAACCCATTGCTAC	Redesign ZK470 F
oJDK427	GCAGTTTAGCGAAGTTGCAC	Redesign ZK470 R
oJDK428	TATTTATCCCGTTGTGAACTTC	Redesign C46F4 F
oJDK429	TCGTATCAATTCAGTGGCTG	Redesign C46F4 R
oJDK430	GAATCATCGACATGACGACG	Redesign C46F4 Fb
oJDK431	TTATTCAAGTAGTGTCGGAAAG	Redesign C46F4 Rb
oJDK432	AGTGATAATTGGTGACAAGTTG	Redesign F22E10 F
oJDK433	CGTAGTCCGACTGTACTTGA	Redesign F22E10 R
oJDK434	GACAAGTTGCGGAATATGTCG	Redesign F22E10 Fb
oJDK435	CTTGATCGTTTGGAACGGC	Redesign F22E10 Rb
oJDK436	CTTTGCAATGGATTAGAGCAC	Redesign F46G10 F
oJDK437	TGGGAGACATACATGGAACC	Redesign F46G10 R
oJDK438	CGTGGAACAGTGTCGAGAT	Redesign H13N06 F
oJDK439	AACCAGGGATAGATTCCAGTG	Redesign H13N06 R

Table 2.6 Oligonucleotides used in amplifying and Sanger sequencing *cam-1A* genomic sequence in *kor37*. F=Forward, R=Reverse.

Designation	Sequence	Comments
oJDK448	ACAACTCTTCGTGCTCCTG	F for amplification and Sanger sequencing *cam-1A* exon 1.
oJDK449	ATTGGTGGTCCTGAAAGTTG	Sanger sequencing *cam-1A* exon 1.
oJDK450	CACATACTGCATAATACTCATC	R for amplification and sequencing *cam-1A* exon 1.
oJDK451	CTGACTTAATCATTCCAGTCC	F for amplifying *cam-1A* exons 2 and 3.
oJDK452	GCTTGGAATGGTTCTGCAAG	R for Sanger sequencing *cam-1A* exons 2 and 3.
oJDK453	CTCCATTTTATGTCATGATGT	R for Sanger sequencing *cam-1A* exons 2 and 3.
oJDK454	GCTCTAGGATAATAGGTTTCG	R for amplifying *cam-1A* exons 2 and 3.

Table 2.7 Oligonucleotides used in amplifying and Sanger sequencing *cwn-2* genomic sequence in *kor32*, *kor43*, and *kor52*. F=Forward, R=Reverse.

Designation	Sequence	Comments
oJDK381	ATAGAGTCACGAACGACGT	F outermost, for sequencing *cwn-2*
oJDK382	GGAGAAGTTGTTGGCTCATAC	F nested, for amplifying *cwn-2*
oJDK383	CAATTCACAGGTCATCGATGG	Sanger sequencing *cwn-2*
oJDK384	AGCAATCAGAGTTCAGGTGG	Sanger sequencing *cwn-2*
oJDK385	CTGGGAACATTCTACAGAATGTAG	R nested, for amplifying *cwn-2*
oJDK386	GCGCCGTCTTGTACCATT	R outermost, for amplifying *cwn-2*
oJDK440	CGGAACTGTGCAATATGATAT	F outermost, for amplifying *cwn-2* promoter.
oJDK441	CTCATGCTAGTAGAAGTGCG	F nested, for amplifying *cwn-2* promoter.
oJDK442	GAACGATTTCAGTAGACATTAGG	Sanger sequencing *cwn-2*
oJDK443	TCTAGTTGGGTTCAATGTGC	Sanger sequencing *cwn-2*
oJDK444	CCAGCCAATGTATTTTGATGTC	Sanger sequencing *cwn-2*
oJDK445	CGTCTGTATCATATCAAATGTC	Sanger sequencing *cwn-2*
oJDK446	GTAATTCGCATACAGGCTTCA	R nested, for amplifying *cwn-2* promoter.
oJDK447	CCA TCG ATG ACC TGT GAA TTG	R outermost, for amplifying *cwn-2* promoter.

Table 2.8 Oligonucleotides used in amplifying *ham-1* and *cam-1* cDNA. F=Forward, R=Reverse.

Designation	Sequence	Comments
oJDK408	CAAGAATGGACGGAAGGTGT	*ham-1* cDNA F
oJDK409	GTGTGGAACTTCGGAGTGGT	*ham-1* cDNA R
oJDK410	GGGGGGTAATACGACTCACTATAGGGAAATCGGCAAATGTTCTGGA	*ham-1* cDNA F nested
oJDK411	GGGGGGTAATACGACTCACTATAGGGCGGGTCTCTGGCATATCAAC	*ham-1* cDNA R nested
oJDK412	ACAACATCTCGGACCTCCAC	*cam-1* cDNA F
oJDK413	TGCTGCTGAAGAATCGAATG	*cam-1* cDNA R
oJDK414	GGGGGGTAATACGACTCACTATAGGGACAAGCTCGTCGTTTTCCAT	*cam-1* cDNA F nested
oJDK415	GGGGGGTAATACGACTCACTATAGGGCGAATTTCCGAGAATGTTGG	*cam-1* cDNA R nested

3 Results

3.1 60 strains with gland cells defects were isolated from forward genetic screening

For this thesis, approximately 4,600 haploid genomes were screened. Combined with the previous work of Onoruza Atta and Maxwell Berg, a final total of 4,986 haploid genomes were screened, and 60 strains with aberrant gland cell migration were isolated. Brenner (1974) estimated that EMS mutagenesis has a forward mutation rate of approximately $5x10^{-4}$ per gene, which is roughly 1/2500. Therefore, to ensure a mutation in a gene involved in pharyngeal gland cell migration would be generated, it was estimated that 2,500 haploid genomes would need to be screened. Increasing the number of genomes screened by two to three times this amount would start to saturate the screen and, in this case, screening 5,000 genomes lowered the probability of missing a relevant gene. Saturation is reached when all genes that can generate the phenotype are isolated. The probability that all genes have been detected increases as multiple alleles are isolated for the same gene(s). Multiple alleles for *cwn-2* and *ham-1* were isolated, as discussed below. The 60 isolated strains are summarized in the appendix. The major phenotypic classes of the strains are homozygous viable under-migrations (25 strains), recessive lethal under-migrations (14 strains), recessive sterile under-migrations (two strains), recessive lethal mixed under- and over-migrations (14 strains), and variable phenotype (five strains). Variable phenotypes require further characterization.

3.2 Candidate Cam-1-like strains mapped to chromosomes I, II, IV, V, and X

From the 60 isolated strains, 25 were selected for further characterization. A priority assignment was made based on the following criteria: the nature of the gland cell defect, the penetrance of the phenotype and the viability of the strain. Figure 3.1 summarizes how strains were prioritized for characterization. For my thesis work, I focused on strains that had an under-

migration defect and therefore may be in the *cam-1* gene pathway. It was predicted that genes isolated with a similar phenotype would be possible down-stream regulators of CAM-1 function. Maternal effect lethality of many Wnt pathway components and the potential for redundancy in the intracellular pathways made a previous directed candidate screening approach unsuccessful. The rationale behind attempting the forward mutational screen was that an unbiased approach would be more likely to identify relevant and potentially uncharacterized pathway components. Homozygous viable under-migrations were easier to maintain. For instance, they did not need to be genetically balanced, as was the case for recessive lethal or sterile strains. Homozygous mutations made genetic mapping with physical markers and whole genome sequencing easier. In whole genome sequence analysis, identification of homozygous deleterious mutations reduced the candidate pool of possible genes. As will be discussed in following sections, the incidence of homozygous deleterious mutations was considerably lower than heterozygous mutations.

25 strains with homozygous viable under-migrations were obtained from mutagenesis. In the interest of time, only 16 strains were selected for linkage analysis. *C. elegans* has five autosomes, I to V, and a single sex X chromosome (Brenner, 1974; Nigon, 1949). Physical markers were used in linkage analysis. This methodology has been called snip-SNP since it makes use of single nucleotide polymorphisms, PCR amplification of the variable sequence, and restriction enzyme digestion to test for the presence or absence of the variable base. *C. elegans* chromosomes span a genetic distance of about 50 map units. The left arm of a chromosome is considered the “minus” end, and the right arm is the “plus” end. The majority of events will be single recombinations since, for the *C. elegans* autosomes, double recombinations do not occur frequently (Zetka and Rose, 1992). This is likely because nematode chromosomes do not have a centralized centromeric sequence. Kinetochore sites for microtubule attachment are dispersed

along the entire chromosome (Maddox et al., 2004). The autosomes have less repetitive sequence at their centers (Consortium*, 1998). Autosomes may also have sequences that suppress recombination in certain regions. The X chromosome apparently does not have these suppression sequences, and has a more uniform occurrence of recombination along the entire chromosome (McKim et al., 1988; Piano et al., 2002). The center-most physical marker was used in a linkage group analysis (Davis et al., 2005).

Mutant strains possessed the N2 background polymorphisms since GD282 was generated in this background. Hawaiian (CB4856) males were mated to the homozygous hermaphrodites to generate heterozygous F1s. Individuals were allowed to self-fertilize and the F2 progeny were separated into mutant and non-mutant groups. At least 20 F1 heterozygous hermaphrodites were used for each cross and the F2 populations were pooled and assessed as a population. For each physical marker, DraI digestion patterns of the mutant parent, Hawaiian parent, F2 mutant, and F2 non-mutant were resolved for each strain. The mutant parent digestion patterns were identical to the N2 background as expected. Since pooled F1 and F2 populations were used, the analysis of the linkage group is semi-quantitative. That is, the relative intensities of unique bands for each SNP were compared to estimate how tightly linked that area is to the causative mutation. For example, if the F2 mutant had a heterozygous snip-SNP digestion, and therefore a N2 band that was of similar intensity to the Hawaiian band, that region was displaying independent segregation and not linked to that chromosome. The F2 non-mutant served as a digestion control: worms that did not maintain the mutant phenotype most likely inherited one or two copies of the Hawaiian allele. Therefore, the F2 non-mutant digestion patterns should tend towards more of the Hawaiian pattern. If the F2 mutant tended to retain the mutant parent/N2 digestion pattern, the causative mutation was considered linked to that chromosome. Of all 16 strains that were

mapped, 12 of them had a very clear banding pattern that strongly tended to the N2 SNP on a single chromosome. Chromosomal mapping results are summarized in Table 3.1, and examples of diagnostic gels are shown in Figures 3.2-3.4. JDK99(*kor45)* showed clear linkage to LGIV. The mutant digestion pattern for LGIV is homozygous for the N2 SNP. The digestion pattern for mutant lanes at all other chromosomes are clearly heterozygous for Hawaiian and N2 SNPs (Figure 3.2). In JDK98, the *kor44* allele appeared to have linkage to chromosome X but a weak Hawaiian banding pattern was also observed (Figure 3.3). This result is consistent with *kor44* being on the X but is not tightly linked to the center of the chromosome. Further mapping of *kor44* is discussed in Section 4.3. JDK56*(kor13)* retained N2 banding patterns on chromosome III, V, and X (Figure 3.4). This could indicate multiple genes contributing to the under-migration phenotype or a complex chromosomal rearrangement. Due to the difficulty in being able to genetically map the causative mutation, this strain was not selected for further analysis. Similarly, the strains that mapped to chromosome V, JDK77*(kor23)* and JDK100*(kor46)*, were not further analyzed because the causative mutation lies on the same chromosome as the integrated *ivIs17* transgene. *ivIs17* is a large multi-copy array that was inserted into the genome using γ-irradiation. Integrated arrays suppress localized recombination making further mapping difficult (Raharjo et al., 2011). These alleles could be mated into a strain with an array with a different insertion site but this was not completed for this thesis. After mating to the Hawaiian background, JDK57, 59, and 95 displayed a reduced penetrance of phenotype. Sufficient numbers of mutant progeny could not be obtained for PCR.

3.3 Complementation crosses reveal two complementation groups on LGIV, and potential alleles of *cwn-2* and *cam-1*

Seven of the alleles with linkage group confirmation mapped to chromosome IV. To determine how many genes were represented by these alleles, complementation crosses were

performed and are summarized in Table 3.2 and Figure 3.5. Heterozygous males were generated by mating a mutant strain with N2 wild-type males. The resulting males were mated to hermaphrodites of a different strain. The progeny from this cross were screened for gland cell location. If the under-migration persisted in these hybrids, the two strains failed to complement and were placed in the same complementation group. If the hybrid progeny had normal gland cell location, the two mutations complemented and were located in different genes. To ensure that mating was successful, only plates that had 50% males in the population were considered. In summary, two independent complementation groups were designated A and B. Complementation crosses assume that alleles represent strong loss of function. During linkage group analysis the alleles behaved like recessive hypomorphs but in the interest of time this was not rigorously tested. Alleles were subsequently sequenced to confirm the nature of the lesion. JDK97(*kor43*) gave ambiguous complementation results. It failed to complement strains from both group A and group B. Within a cross, mated hermaphrodites had very different mutant progeny ratios. For example, in the JDK112(*kor15*) x JDK97(*kor43*) cross (Table 3.2), one plate had no mutant progeny, while another plate had nearly half mutant progeny. The reciprocal cross yield similar results. The other group B allele, *kor5*, when crossed to JDK97(*kor43*) had a mutant ratio over 50%. In all cases matings appeared to be successful. It is still unclear why inconsistent crossing results occur with JDK97(*kor43*). Based on sequencing results (discussed below), JDK97 was later placed in complementation group A. It was found to be a late stop codon at Q231 (Figure 3.12). It is possible that a truncated protein persisted and this was interacting with the hemizygous group B alleles in a complex genetic enhancement.

LGIV complementation group A alleles possessed a withered tail phenotype that was present in strains of *cwn-2* loss of function mutations. *kor43* and *kor49* were crossed into the

cwn-2(*ok895)* background. *ok895* is a confirmed deletion in *cwn-2* and probable null allele (The C. elegans Deletion Mutant Consortium, 2012). *ok895* was also crossed into group B representative strain *kor15*. The crosses are summarized in Table 3.3. Crossing *ok895* into another background gave variable results when crossed with *kor15* and *kor43.* Each plate gave a different proportion of mutant progeny. The averaged proportions for these crosses appeared to complement, however, this is not an accurate reflection of inconsistent individual crosses in this case. When *kor43* or *kor49* were crossed into the *ok895* background, they both failed to complement. Crossing *kor15* and *ok895* produced variable results. In one cross there was clear complementation but in the reciprocal there was a consistent failure to complement. Although there are some complex interactions presented within the complementation crosses (see Discussion), two probable allelic groups were hypothesized. Two alleles, *kor45* and *kor49,* were sent for whole genome sequencing as detailed in the following results section. In brief, mutations in *cwn-2* that generated premature stop codons were identified. Subsequent to these results the remaining three alleles, JDK76*(kor22),* JDK97*(kor43),* and JDK106*(kor52),* were sent for Sanger sequencing. These three alleles were also premature stop codons in *cwn-2* (Figure 3.12). In summary, five alleles of *cwn-2* have been identified from the genetic screen.

JDK91(*kor37*), on LGII, also possessed a low-penetrance withered tail phenotype. Because *cam-1* is located on chromosome II, a complementation cross was performed with the *cam-1(gm122)* null allele. *kor37* failed to complement *cam-1*(*gm122*) and therefore is likely a *cam-1* allele. Sanger sequencing revealed a single base pair deletion at position 9,304,751 on chromosome II (genome coordinate from Wormbase version WS230). This changes the sequence GAAATGCCACCAAAAGTTCGGGAGACGAGGT to

GAAATGCCACCAAAA-TTCGGGAGACGAGGT. This deletion occurs in the second exon of isoform A, and the first exon of isoform B. A serine residue at position 50 (relative to isoform A) is changed to an isoleucine, and a consequential frameshift then introduces a premature stop codon at amino acid 74 (Figure 3.12). Both amino acid changes occur before any defined functional domains or motifs (Figure 3.13B).

The two alleles on the V linkage group were tested for complementation. The complementation results were variable (Table 3.4). *kor23* penetrance appeared to drop to about 25% with regular maintenance of the strain. Because of this low penetrance, the *trans* progeny produced with *kor46* may have a low penetrance as well. The presence of under migration defect in progeny is 5-8%, which may indicate a failure to complement. Repeating this cross and performing the reciprocal cross may clarify this result.

3.4 Whole genome sequencing analysis

kor5, *kor15*, *kor42*, *kor44*, *kor45*, and *kor49* were selected for whole genome sequencing along with the parental GD282 and Hawaiian strains. Two alleles from each LGIV complementation group were sequenced to find common candidate genes between allelic strains. This would quickly reduce the number of candidate mutations from each strain. The GD282 parental genome was used to subtract mutations and variants that were present prior to mutagenesis. The Hawaiian mating intended to use the mapping-by-sequencing strategy (Doitsidou et al., 2010; Minevich et al., 2012). This would allow for identification of Hawaiian versus N2 SNPs throughout the entire genome and allow for large-scale SNP mapping.

Genomic DNA was submitted to Genome Quebec for library preparation and whole genome sequencing. The final sample concentrations and volumes sent for sequencing are summarized in Table 3.5. Samples were prepared for paired-end 125 basepair sequencing on the

Illumina HiSeq 2500 platform. Eight genomes were sequenced on a single lane. Between 27.1 million and 36.5 million reads were returned for each strain. FastQC reports run on raw reads returned average Phred quality scores near 36 indicating excellent sequence quality across the entire 125 base pairs of reads (Figure 3.6). Phred quality scores reflect the accuracy of bases called in each read position. The higher the Phred score, the lower the probability that a base was called incorrectly (Ewing and Green, 1998). A Phred score of at least 30 is considered standard for high-quality reads. Reads generated for this project were also consistently 125bp in length, had normal G/C content, and very low "N" or unidentified nucleotide content. Average coverage for each strain ranged from 63x-84x. This means that, on average, each nucleotide was covered or sequenced 63 to 84 times by different reads distributed along the genome. This is much higher than required for SNP discovery. For example, simultaneous mapping and sequencing required 20 to 40-fold coverage (Doitsidou et al., 2010). In a different analysis, homozygous SNP detection required about 15x coverage, while heterozygous SNPs required 33x (Bentley et al., 2008). In future experiments, about two or three times as many strains with heterozygous or homozygous mutations, respectively, could be sequenced with a single lane.

Using the CloudMap pipeline developed by the Hobert laboratory (Minevich et al., 2012), the Hawiian SNPs could be used to map the linkage of the mutation while simultaneously sequencing the genome. The populations of 50 F2 recombinants were pooled for each mutant strain. Individual plates were collected when the bacterial lawn was depleted to keep the number of worms collected from each plate relatively consistent. A random subsample of each pooled sample (~360mg of frozen pellet) was prepared for genomic extraction and whole genome sequencing. The CloudMap pipeline plotted ratios of Hawaiian SNPs over total reads at every SNP position in the genome. The resulting graphs would reveal areas along chromosome(s) that

had low occurrence of Hawaiian SNPs. These areas were likely to have retained the N2 SNP and be linked to the causative mutation. Examples of CloudMap Hawaiian mapping are included in Figures 3.7-3.9. *kor5* and *kor15* causative mutations were best estimated at 12.5 Mb on LGIV. *kor45* mapped between 8.0 and 10.5 Mb on LGIV, while the allelic strain *kor49* mapped between 10 and 12 Mb on LGIV. *kor42* appeared linked between 8 and 10 Mb on LGI. *kor44* did not link to a specific area on LGX, however the scatterplot suggests a region near 10Mb. The lack of a local decrease in Hawaiian allele could indicate modified recombination rates through a large deletion or rearrangement (Kormish and McGhee, 2005). Creating new strategies within Galaxy or manually sorting through alignments to examine these intervals was time consuming as well as unreliable due to some tools no longer being supported on the free platform. A collaboration with the UBC gene knockout consortium was established to assist in data analysis. The pipeline used for sequence variant calling did not use Hawaiian SNP mapping (Flibotte et al., 2010; Thompson et al., 2013). The sequenced Hawaiian variants were subtracted from mutagenized strains before identifying variants.

3.5 LGIV group A is *cwn-2*, and LGIV group B is *ham-1*

After aligning reads, homozygous variants were called for each strain. Variants that were present in the GD282 or Hawaiian parental strain were subtracted from the variant files. Variants that were common between non-allelic mutagenized strains were considered passenger mutations and also eliminated. Only one variant was eliminated in this step. Variant files were then filtered for changes affecting protein-coding regions including nonsense mutations, missense mutations, splice defects, and indels (insertions or deletions of basepairs). Large deletions and rearrangements were later searched for with no significant changes found in mutagenized strains. Tables of genes that fit filtering criteria for each strain are included in Tables 3.6-3.10. The only

gene common between the LGIV group A strains, *kor45* and *kor49*, was *cwn-2*. *kor45* is a guanine to adenine transition that changes the twenty-sixth amino acid from a tryptophan to a stop codon. *kor49* is a cytosine to thymine transition that changes glutamine 54 to a stop codon. The only gene common between the LGIV group B strains, *kor5* and *kor15*, was *ham-1*. *kor5* is the same as the previously reported *gm267* allele: a guanine to adenine transition that changes glycine 58 to an arginine residue. *kor15* is a new allele of *ham-1*, which is a cytosine to thymine transition changing glutamine 202 to a premature stop codon. These mutations are summarized in Figures 3.12 and 3.13.

3.6 Two candidate genes each for *kor42* LGI and *kor44* LGX

Assigning the causative gene in the strains mapping to LGI and LGX was much more difficult, as they did not have an allelic strain to compare for lesions in the same genes (candidate gene lists are included in Tables 3.8 and 3.9). Variants occurring within introns in these genomes were later examined, however only a single synonymous mutation was found within a gene. To narrow down the genetic interval and eliminate some genes that do not fall in this smaller interval, another round of snip-SNP mapping was conducted. A subsample of 30 individual F2 mutants' progeny from the whole genome sequencing preparation was used as template for interval mapping (Davis et al., 2005). The basis of recombination at the chromosomal level is depicted in Figure 3.10. The results are summarized in Figure3.11. Based on these results, *kor42* is between -6 and 5 on chromosome I, which leaves only the candidate genes *unc-73* and *srt-61* within the linked interval. *kor44* is between 2 and 11 on chromosome X, which includes *plpp-1.3* and *mig-2*. Because the F3-F4 progeny were collected for analysis, it was possible for a second single recombination event to occur. This means that within a clonal population, two separate recombinations could have occurred. With 30 clonal populations, there would be a potential of

60 recombination events to occur. This can account for the homozygous Hawaiian SNPs observed at the ends of some chromosomes.

3.7 g1P quantification of novel mutations

g1P under migration was quantified in JDK114 *ham-1(kor15)* IV, JDK119 *ham-1(kor5)* IV, JDK117 *unc-73(kor42)* I, and JDK118 *mig-2(kor44)* X. Their distributions and an image characteristic of the under migration defect are displayed in Figure 3.14. To score defects, approximately 20 hermaphrodites were picked to a plate and incubated overnight at 20˚C. The hermaphrodites were removed the next day and the eggs were allowed to hatch over the next 36-48 hours. L1-L2 larvae were collected in M9 buffer, anesthetized with tetramisole hydrochloride, and mounted on agarose pads. Scoring and Z-stack capture with a yellow fluorescence channel were done under 100x immersion oil objective. Corresponding DIC images were also taken with Z-stacks. Individuals were assigned a score based on the posterior-most muscle group the g1P cell body was located in. Penetrance of the under-migration defect and other observed phenotypes are summarized in Table 3.11.

The *ham-1* strains appear to have a bimodal defect: when the g1P cell does not make it to pm7, it is usually in pm4 or pm6, which are the anterior and posterior bulbs of the organ, respectively. The g1a cells can also appear to be disorganized and have thickened extensions. The g2 cells are usually not detectable in these backgrounds.

Multiple tail phenotypes were observed in the *ham-1* strains (Figure 3.15). The most common is a Nob-like tail, which can give a blunt appearance under a dissecting microscope. Higher magnification shows a variable rounding and misplacement of the tapering tail end (Van Auken et al., 2000). The Nob symbol stands for no back-end. *nob* mutants do not develop the posterior end of their bodies. A second slightly less common tail phenotype is *utx-1*-like. Here, a

ball-like shape forms at the very end of the tail instead of being pointy (Vandamme et al., 2012). *utx-1* is a histone demethylase that has an essential non-catalytic role in embryonic development. Loss of *utx-1* function can cause posterior defects and aberrant gonad migration and organization. A rarer occurrence in *ham-1* mutants is a forked tail spike, which appears normal except for the split cuticle at the very posterior end of the tail (Roy et al., 2000; Zhao et al., 2010). Additional phenotypes observed in the *ham-1* backgrounds include underdeveloped gonads and egg-laying defect.

The *kor42* allele displayed a low penetrance of 20% gland cell defect. In some cases, 8% (n=88), the g1p cell appeared to span multiple muscle groups. A low-penetrance egg-laying defect was apparent in some of the adults examined under a dissecting microscope. The *kor44* allele comparatively had a slightly higher penetrance and under migration severity than *kor42*. The g1P is more likely to stop in a pm6 or 6.5 position, or sometimes may have a small "finger" or point that crosses the pm6/7 border. The g1P extension can be faint or beaded, as well. The g2 cells are not detectable in this background. This strain also has a low penetrance egg-laying defect and becomes extremely bloated in the late adult stage and sometimes develops a bag of worms.

3.8 RNA interference of *ham-1* and *cam-1*

3.8.1 *ham-1* RNAi phenocopies *ham-1* strains possessing g1P under migration

Knocking down *ham-1* in GD282 *ivIs17* using RNA interference resulted in a similar g1P distribution as *kor5* and *kor15*. This suggests that strong knock-down with *ham-1* RNAi was achieved. The distributions are visualized in Figure 3.16. The *ham-1* RNAi distribution appears more like the *kor15* than the *kor5* allele. *kor15* is a truncated version of HAM-1, while *kor5* is a

non-synonymous amino acid change. It is likely that *kor15* has a stronger loss-of-function than *kor5*.

ham-1(RNAi) also produces some progeny with the Nob-like tail phenotype, further supporting that specific knock-down of *ham-1* was achieved and that *ham-1* is indeed the gene causing the g1P under migration defect in *kor5* and *kor15*.

3.8.2 *cam-1* RNAi produces an under-migration phenotype

cam-1 dsRNA produced an extremely weak phenotype when injected into the GD282 strain. The g1P distributions of *ivIs17*; *cam-1*(*RNAi*) and the control *ivIs17*; L4440 RNAi were statistically different according to the Mann-Whitney *U* test, with *cam-1* RNAi causing a slight increase in the occurrence of under migrations (Figure 3.17). No other Cam-1-like phenotypes such as a withered tail or uncoordinated movement were observed.

This weak interference was most likely caused by a suboptimal *cam-1* dsRNA product. All synthesized *cam-1* RNA products revealed a clear doublet when resolved on an agarose gel (see Figure 3.18). These double bands persisted even when each band was individually gel extracted and re-purified. This may be due to inappropriate annealing of complementary RNA strands within a sample, secondary structures, contamination with genomic DNA, or unknowingly amplifying an additional gene with the nested primers and creating an impure sample.

3.8.3 *cam-1* and *ham-1* genetic interaction

Because *cam-1* and *ham-1* mutants confer a similar g1P under-migration phenotype, it is possible that they are involved in the same pathway during embryonic development to influence gland cell position within the pharynx. To test this, either gene was knocked down with RNAi in

a background with a putative null or strong loss-of-function allele of the reciprocal gene. This created either a *cam-1(gm122); ham-1(RNAi)* or *ham-1(kor15); cam-1(RNAi)* double mutant. If the two genes are working in the same genetic pathway, the g1P distribution and under migration defect should remain the same in the double mutant background. It is important to use a null allele in the mutant background to ensure there is no residual activity of the gene. Residual activity can create a deceptive phenotype in a double mutant and lead to incorrect epistatic analysis (Huang, 2006). Alternatively, if *cam-1* and *ham-1* are working in different genetic pathways, there should be an additive effect of both mutated genes and an enhancement of the under migrated phenotype – that is, the g1P should be found in an anteriorly position even more often. The results of these RNAi experiments are summarized in Figure 3.19.

ham-1(kor15); cam-1(RNAi) distribution is not different from its control, while *cam-1(gm122); ham-1(RNAi)* is different from its control. This inconsistency may be due to the poor *cam-1* RNAi knock down. With the null *cam-1* allele and strong *ham-1* knockdown, however, there is an enhancement of the g1P under migration. This suggests that the two genes are acting in two different pathways.

Unexpectedly, the *cam-1(gm122); ham-1(RNAi)* also produced enhanced defects in other gland cells, and affected overall pharynx morphology. Samples of these defects are pictured in Figure 3.20. The g1a cells commonly had morphology defects, including lobes that expanded in the left/right direction or sometimes into the pm7 region. Many g1a extensions were also thickened, blebbed, or had curved/crooked paths through the pharynx. The overall shape of the pharynx was affected in multiple ways, including asymmetry of the left and right halves of the metacorpus/anterior bulb, a misshapen terminal bulb, or a shortened isthmus/pm5 region. While some mild asymmetry defects are observed in either single mutant alone, the defects seen in the

double mutant appear to have a higher penetrance and more severe/misaligned phenotype. Gland cells also occasionally appeared outside of the pharynx.

Although the *ham-1(kor15); cam-1(RNAi)* double mutant does not seem to affect g1P distribution, there appears to be an enhancement of improper pharyngeal-wide patterning or development. Some examples are included in Figure 3.21. The g1a morphology defects also occur in this background, sometimes making it difficult to distinguish the g1P from the g1as.

The metacorpal asymmetry defects are less penetrant and less severe than the *cam-1(gm122); ham-1(RNAi)* background. This is expected, as the anterior bulb defect occurs in *cam-1(gm122)* single mutants (Jay Kormish, personal communication). If there was very weak *cam-1* RNAi knockdown in the double mutant, it would not have the same loss-of-function strength as the putative null allele, and any phenotypes contributed by RNAi would be milder.

The misshapen terminal bulb and short isthmus phenotypes persist in the *ham-1(kor15); cam-1(RNAi)* background. In some cases, an isthmus that elongated also appeared to have weak pm5 muscle integrity. Sometimes the pm5 section buckled or was misshapen (see Figure 3.21A), and in a couple cases, the area burst when subjected to pressure from oil immersion microscopy, allowing the gland cells and other contents to be pushed outside of the organ (see Figure 3.21B).

In yet another unexpected phenotype, the g1P cell appeared to be duplicated in a low number (2-3%, n=92) of *ham-1(kor15); cam-1(RNAi)* double mutants (Figure 3.21D). This duplicated cell emitted a low level of fluorescence compared to the other gland cells, and it was difficult to determine if it possessed a nucleus. The g1p nucleus is typically large and can be identified by its relatively higher fluorescence level within the cell. Both supposed g1P cells' extensions were disjointed in the pm3 area only, making it difficult to draw conclusions about

how the cells ended up at their ultimate positions and the identity of the duplicated cell based on this single snapshot.

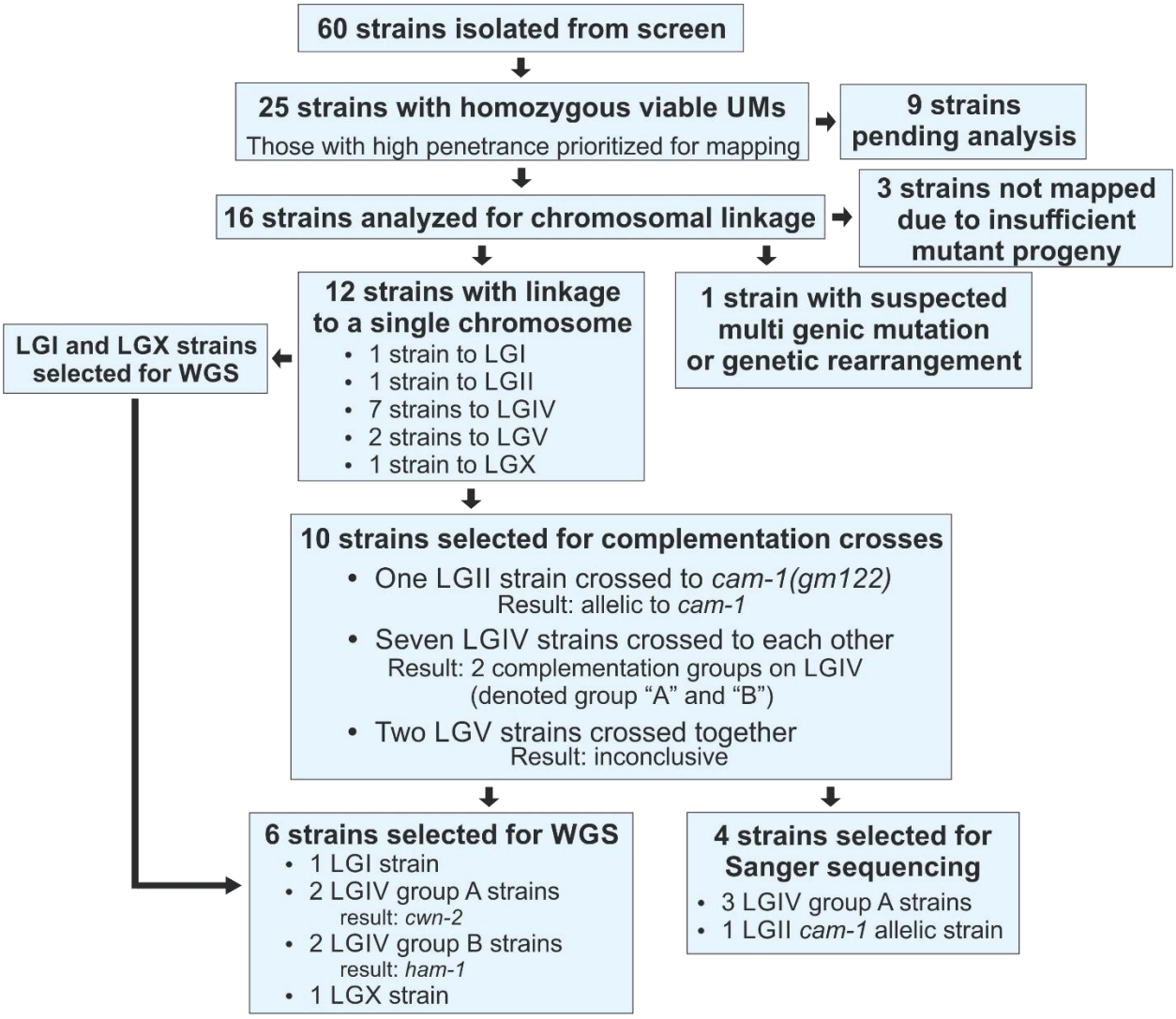

Figure 3.1 Summary of mutant strain characterization completed in this thesis. WGS=Whole genome sequencing.

Table 3.1 Summary of chromosomal linkage analysis for sixteen alleles. Twelve alleles were linked to one of the six linkage groups. Linkage of JDK56 suggests a complex chromosome rearrangement. Three strains displayed a reduction in penetrance upon crossing to the Hawaiian background and insufficient numbers of mutant progeny were available for linkage assignment (marked NA). The table summarizes the cumulative work of Onorueza Atta, Maxwell Burg, Patrycja Sroga, Stacey Line and Stephanie Tkachuk.

Strain / Allele	Linkage Group
JDK96(*kor42)*	I
JDK91(*kor37)*	II
JDK32(*kor5)*	IV
JDK58(*kor15)*	IV
JDK76(*kor22)*	IV
JDK97(*kor43)*	IV
JDK99(*kor45)*	IV
JDK103(*kor49)*	IV
JDK106(*kor52)*	IV
JDK77(*kor23)*	V
JDK100(*kor46)*	V
JDK98(*kor44)*	X
JDK56(*kor13)*	III, V, X
JDK 57(*kor14)*	NA
JDK59(*kor16)*	NA
JDK95(*kor41)*	NA

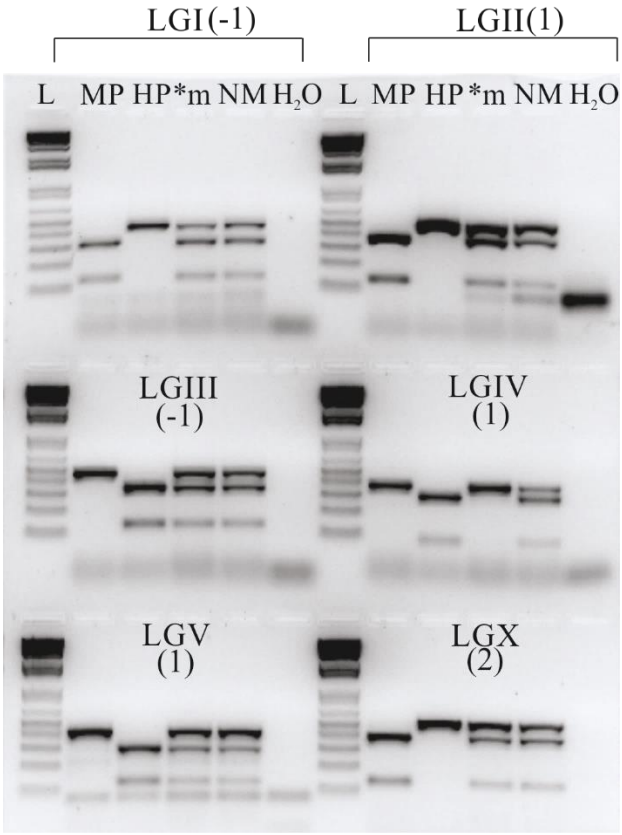

Figure 3.2 JDK99 shows linkage to chromosome IV. Example of a snip-SNP mapping gel analyzing pooled populations for linkage analysis (Davis et al., 2005) L=DNA ladder, MP=mutant parent, HP=Hawaiian parent, *m=F2 mutant progeny, NM=F2 non-mutant progeny, H20=water, negative control for PCR reactions.

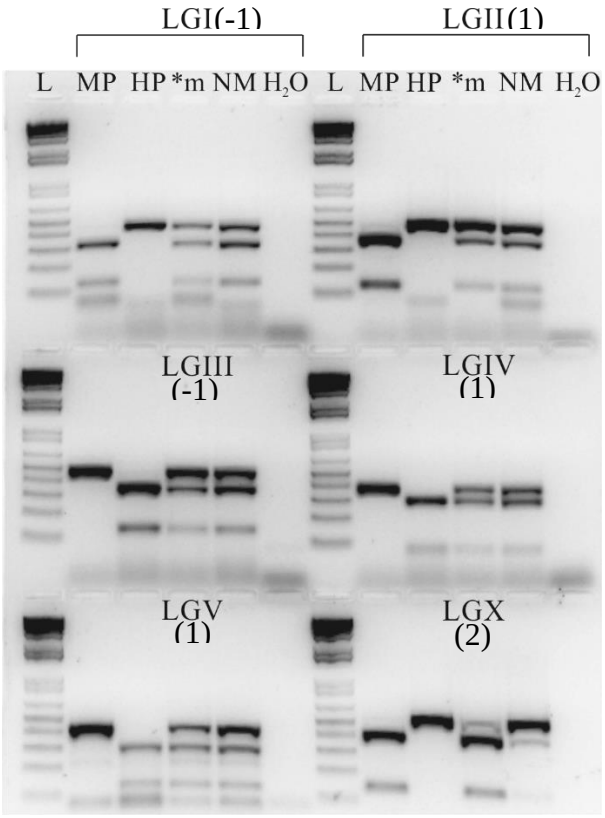

Figure 3.3 JDK98 shows linkage to the X. Example of a snip-SNP mapping gel analyzing pooled populations for linkage analysis (Davis et al., 2005) L=DNA ladder, MP=mutant parent, HP=Hawaiian parent, *m=F2 mutant progeny, NM=F2 non-mutant progeny, H20=water, negative control for PCR reactions.

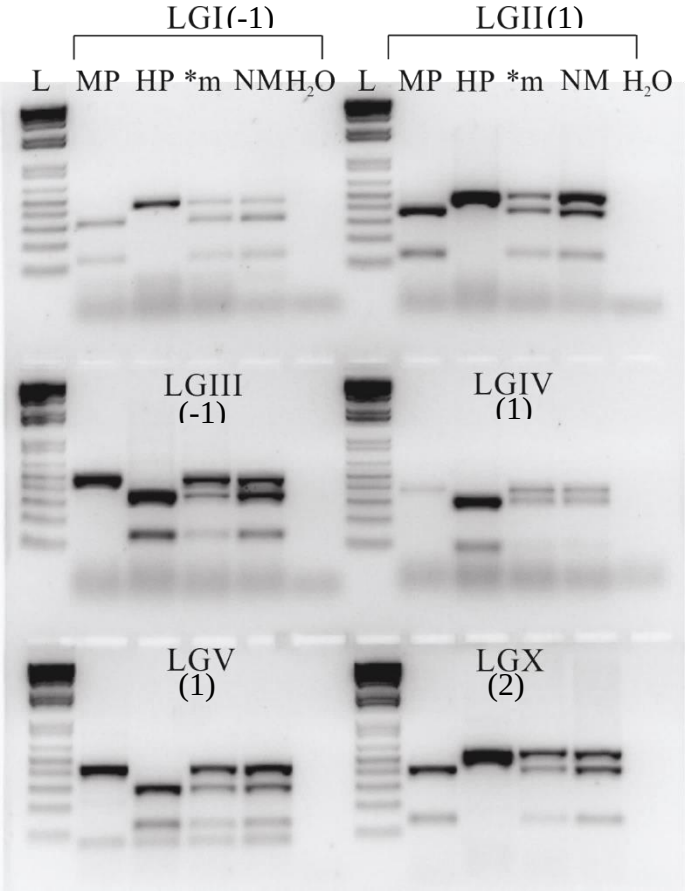

Figure 3.4 Linkage group analysis for JDK56. L=DNA ladder, MP=mutant parent, HP=Hawaiian parent, *m=F2 mutant progeny, NM=F2 non-mutant progeny, H20=water negative control for PCR reactions. Banding patterns for LGIII, V, and X display linkage to III and weak linkage to V and X. A multigenic mutation or genetic rearrangement is suspected.

Table 3.2 Results of complementation crosses between strains mapping to LG IV. Heterozygous males were mated to hermaphrodites of a different LG IV strains. Only plates that contained male progeny were scored. The number of worms with an under-migration defect and the total number of progeny on a plate were recorded. Average under-migration penetrance was calculated. Because males were heterozygous and hermaphrodites were homozygous, the maximum occurrence of homozygous recessive mutants on a plate would be 50%. Not all strains have complete penetrance, so actual proportions can be lower than 50%. If a cross resulted in 10-50% of progeny retaining the under-migration phenotype, it was considered a failure to complement. >50% was a partially dominant interaction.

Cross (male X hermaphrodite)	UM/Total Progeny						Average UM	Complement
kor22/+ X *kor5*	0/85	0/85	0/66				0%	Yes
kor22/+ X *kor43*	12/40	31/112	9/40	31/89	31/86		30%	No
kor22/+ X *kor45*	19/87	6/15	4/32	18/85			24%	No
kor22/+ X *kor52*	38/94	10/30	27/79	33/90	25/65		36%	No
kor43/+ X *kor5*	21/33						64%	No
kor43/+ X *kor22*	27/92	16/53	28/112	12/57	19/65		27%	No
kor43/+ X *kor45*	7/26	14/65	14/68	13/24			31%	Variable
kor43/+ X *kor49*	13/41	19/50	22/43	38/88			41%	Variable
kor43/+ X *kor15*	10/16	1/12	3/37	2/15	14/29		27%	Variable
kor49/+ X *kor5*	0/43	1/80					0.6%	Yes
kor49/+ X *kor22*	2/9	10/36	5/15	5/14			30%	No
kor49/+ X *kor43*	33/107	0/6	16/77	23/65			22%	No
kor49/+ X *kor45*	8/20	8/22	6/17	6/25	8/33		32%	No
kor49/+ X *kor52*	30/85	22/87	39/103	24/83	14/55	30/70	33%	No
kor49/+ X *kor15*	0/60	0/56					0%	Yes
kor52/+ X *kor45*	9/23	30/96	13/35	21/53			37%	No
kor15/+ X *kor5*	35/60	33/54	27/39	32/68	30/67		56%	No
kor15/+ X *kor22*	1/52	0/69	0/100	0/43	1/80		0.6%	Yes
kor15/+ X *kor43*	1/94	0/98	30/73				14%	Variable
kor15/+ X *kor45*	1/97	0/64	2/110	0/27	0/108	3/8	0.07%	Variable
kor15/+ X *kor49*	2/65	2/58	0/37	3/94	1/71		2%	Yes
kor15/+ X *kor52*	5/77	2/95	0/100	1/94			0.02%	Yes

A

kor	5	15	22	43	45	49	52	*ok895*
5								
15	-		+	+/-	+	+	+	+
22	+			-	-		-	
43	+/-	+/-	-		+/-	+/-		-
45								
49	+	+	-	-	-		-	-
52					-			
ok895		+/-		+/-		-		
WT	+	+	+	+	+	+	+	

B

Group A	Group B
22	5
43	15
45	
49	
52	

Figure 3.5 Summary of complementation crosses between strains mapping to LGIV. (A) Crossing punnet of LGIV strains. Y-axis is male, X-axis is hermaphrodite. + = progeny had normal gland cell placement; complementation. - = under-migrations present in progeny; failure to complement. +/- =variable complementation between individual crosses. (B) Strains organized into respective complementation groups based on crossing data in (A). *kor43* gave ambiguous crossing results and failed to complement strains in both Group A and Group B. This mutation was later placed in Group A based on sequencing results.

Table 3.3 Results of complementation crosses between strains mapping to LG IV and *cwn-2(ok895)*. Heterozygous males were mated to hermaphrodites of a different strain.

Cross (male X hermaphrodite)	Progeny with UM/Total Progeny					Average UM	Complement
ok895/+ X *kor15*	4/19	0/19	9/22	2/19		18%	Variable
ok895/+ X *kor43*	6/35	1/20	19/57			18%	Variable
ok895/+ X *kor49*	10/40					25%	No
kor15/+ X *ok895*	0/24	2/71	0/33	0/83		0.7%	Yes
kor43/+ X *ok895*	2/47	1/22	8/49	9/64	4/26	11%	No
kor49/+ X *ok895*	1/18	7/62	5/64	5/69		8%	No

Table 3.4 Results of complementation cross between strains mapping to LGV. Heterozygous males were mated to hermaphrodites.

Cross (male X hermaphrodite)	Progeny with UM/Total Progeny					Average UM
kor46/+ X *kor23*	1/39	7/51	0/37	1/28		5%
kor46/+ X *kor23*	0/25	8/38	9/49	0/41	1/39	8%

Table 3.5 Quality of genomic DNA preparations submitted to Genome Quebec for Illumina paired-end whole genome sequencing. Concentrations are averaged reads from a Quant-iT PicoGreen dsDNA assay, and the OD ratios are averaged reads from a Nanodrop 2000 spectrophotometer.

Strain	Concentration (ng/µL)	Sample volume (µL)	OD 260/280	OD 260/230
CB4856 (Hawaiian)	55	45	1.96	2.03
GD282	150	45	1.90	1.92
kor5	55	45	1.94	1.94
kor15	38	45	1.96	1.96
kor42	40	45	1.93	1.83
kor44	59	45	1.91	1.90
kor45	50	45	1.92	1.68
kor49	43	45	1.93	1.90

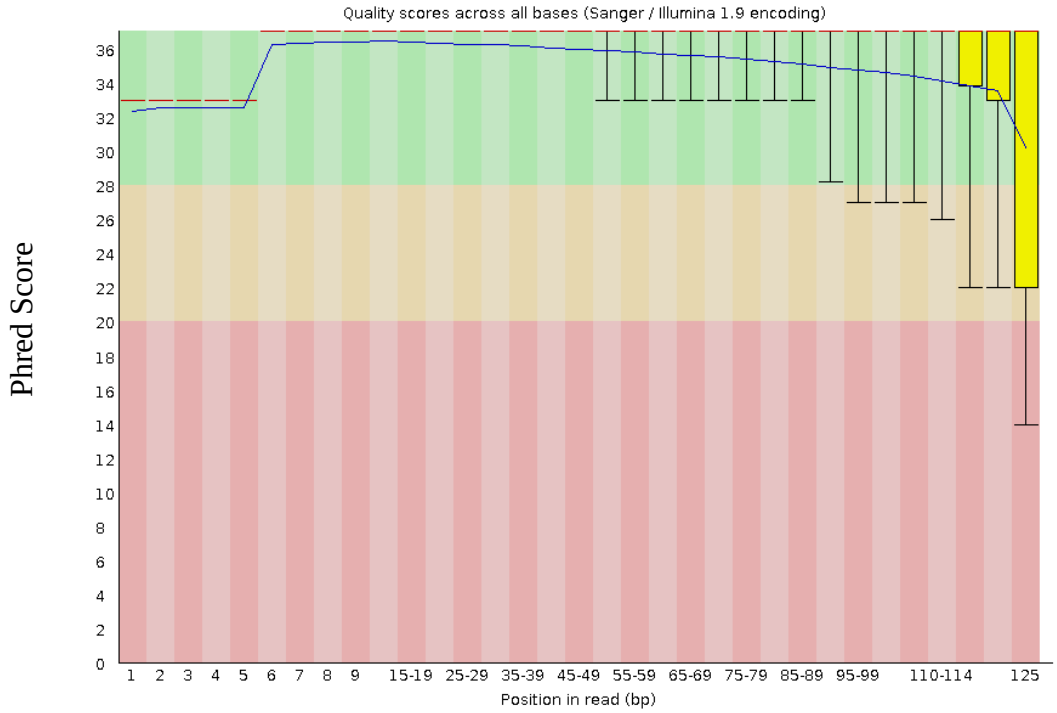

Figure 3.6 Sample output of "Per base sequence quality" from a FastQC report on raw reads. This output was generated from analysis of JDK98 forward reads. X-axis is base pair position within each 125bp read, and y-axis is Phred score associated with that position, which indicates the quality of call for each nucleotide. A higher score means higher quality. For each position, the red line represents the median score, and the blue line is the mean. Yellow boxes represent the interquartile range (25-75%), and whiskers are 10% and 90% points (maximum whiskers do not appear on this graph because the median scores are already at maximum).

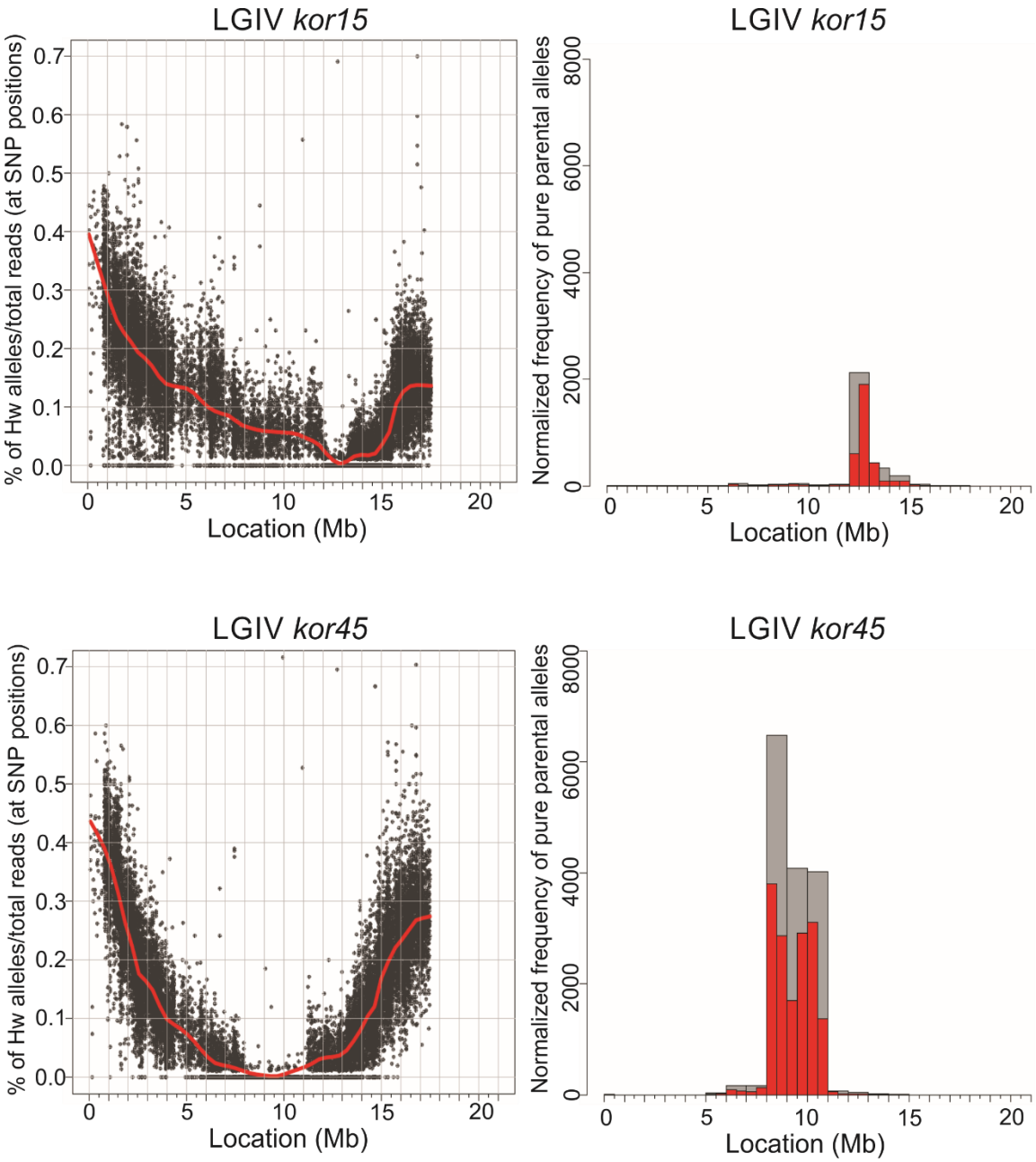

Figure 3.7 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool (Minevich et al. 2012). Only the chromosome that demonstrates linkage is displayed for each strain. (Left) Scatterplot representing the percentage of Hawaiian SNPs at each position. Data points lying exactly on the X-axis indicate regions that retain N2 SNPs, which are putatively linked to the causative mutation. Red line is a local regression (LOESS) line that weighs plotted points and provides a higher-accuracy estimate of the linked region. (Right) Frequency plot showing areas where N2 parental alleles are concentrated. These peaks correspond with dips in the LOESS line on the scatterplot, and suggest areas to begin searching for causative mutations. *kor15* and *kor45* outputs showed tight linkage, spanning areas of only about two to three megabase pairs.

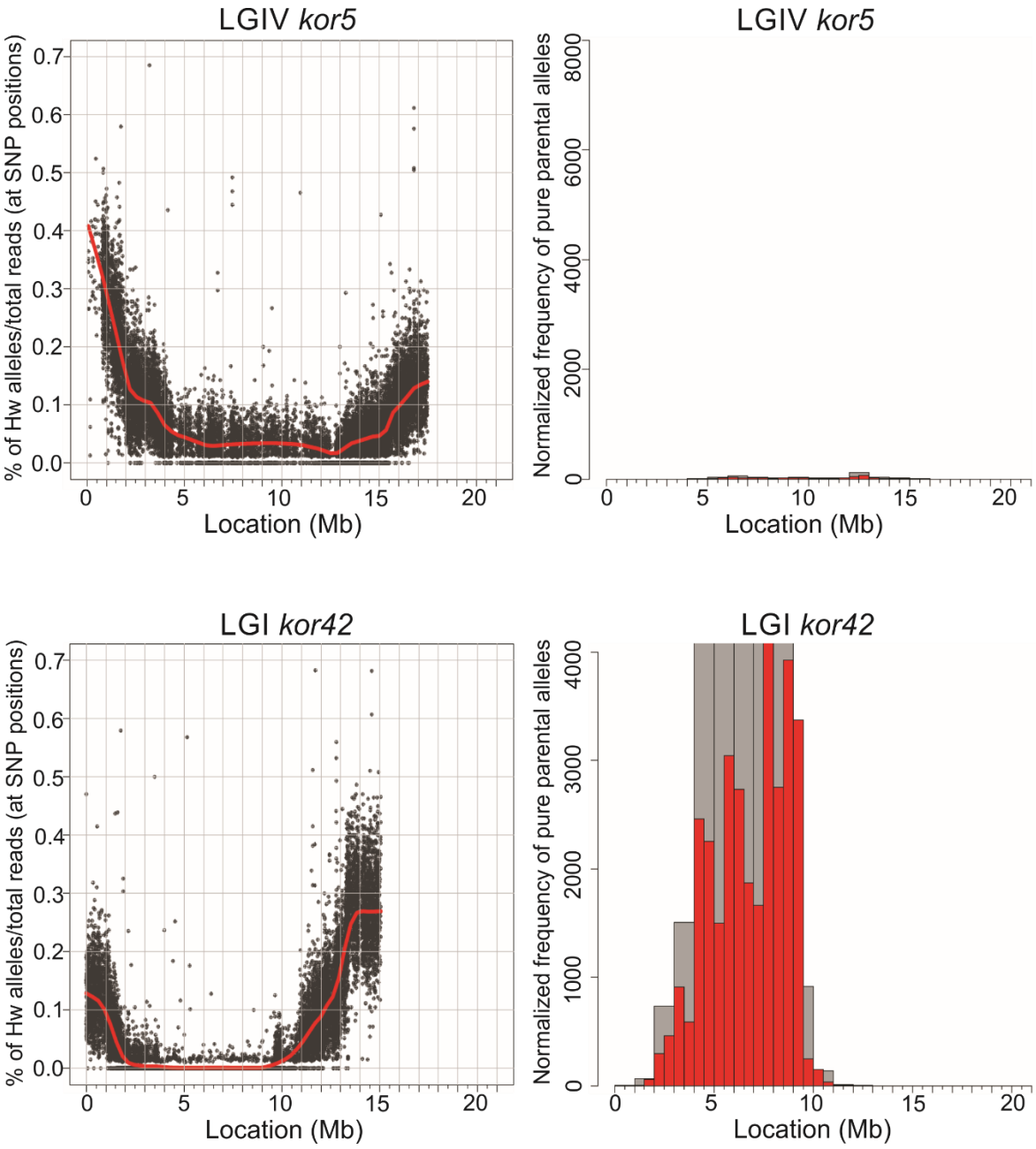

Figure 3.8 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool (Minevich et al. 2012). Only the chromosome that demonstrates linkage is shown for each strain. (Left) Scatterplot representing the percentage of Hawaiian SNPs at each position. Data points lying exactly on the X-axis indicate regions that retain N2 SNPs, which are putatively linked to the causative mutation. Red line is a local regression (LOESS) line that weighs points near the linked area and provides a higher-accuracy estimate of the linked region. (Right) Frequency plot showing areas where N2 parental alleles are concentrated. These peaks correspond with dips in the LOESS line on the scatterplot, and suggest areas to begin searching for causative mutations. *kor5* shows general linkage near the middle of LGIV, and *kor42* shows linkage to the left arm of LGI.

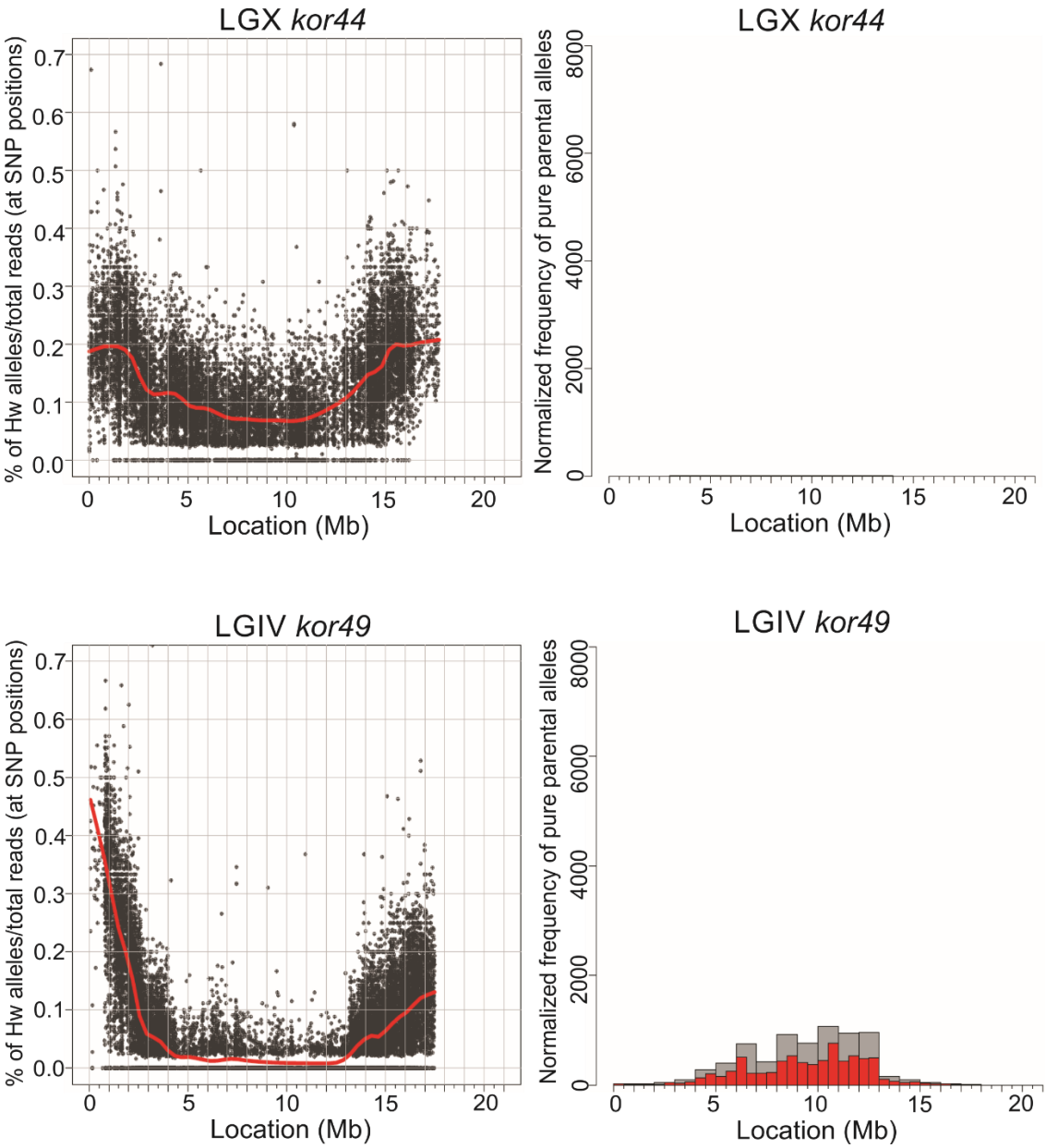

Figure 3.9 Outputs from CloudMap Hawaiian Variant Mapping with WGS data tool (Minevich et al. 2012). Only the chromosome that demonstrates linkage is displayed for each strain. (Left) Scatterplot representing the percentage of Hawaiian SNPs at each position. Data points lying exactly on the X-axis indicate regions that retain N2 SNPs, which are putatively linked to the causative mutation. Red line is a local regression (LOESS) line that weighs points near the linked area and provides a higher-accuracy estimate of the linked region. (Right) Frequency plot showing areas where N2 parental alleles are concentrated. These peaks correspond with dips in the LOESS line on the scatterplot, and suggest areas to begin searching for causative mutations. *kor44* shows weak linkage across LGX, and *kor49* shows linkage to the center of LGIV.

Table 3.6 Candidate gene list for *kor5*

Chromosome	Physical location (bp)	Sequence change	Change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
IV	4588734	C -> T	SNV	coding	F28E10.1	missense	S->L
IV	8872372	G -> A	SNV	coding	C09G9.3	missense	P->S
IV	9590050	G -> A	SNV	coding	F27C8.5	missense	A->V
IV	11640888	G -> A	SNV	coding	ttr-50	missense	P->L
IV	12542464	G -> A	SNV	coding	ham-1	missense	G->R
IV	13547437	G -> A	SNV	coding	F52G2.3	missense	P->L
IV	14333075	G -> A	SNV	coding	Y67A10A.3	missense	H->Y
V	2695946	G -> A	SNV	coding	Y45G12B.2	missense	R->H
V	17354263	G -> A	SNV	coding	F47H4.2	missense	D->N

Table 3.7 Candidate gene list for *kor15*

Chromosome	Physical location (bp)	Sequence change	Type of change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
II	3336082	G->T	SNV	coding	F14D2.19	missense	A->D
II	12756071	C->T	SNV	intron	Y46G5A.7	splicing	-
IV	3372601	C->T	SNV	coding	vha-19	missense	S->L
IV	4459672	C->T	SNV	coding	R08C7.12	missense	H->Y
IV	7394949	C->T	SNV	coding	C48A7.2	missense	S->L
IV	7465682	C->T	SNV	coding	F13B6.3	missense	G->R
IV	8918169	A->G	SNV	coding	rap-1	missense	M->T
IV	9011761	C->T	SNV	coding	C53D6.11	missense	S->F
IV	9304060	C->T	SNV	intron	F49C12.4	splicing	-
IV	12500286	C->T	SNV	coding	sru-19	missense	P->L
IV	12543618	C->T	SNV	coding	ham-1	nonsense	Q->*

Table 3.8 Candidate gene list for *kor42*

MMP Rank:Occurrence	Chromosome	Physical location (bp)	Genetic position	Sequence change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
169:48	I	593656	-18.92	G-> A	Coding	hum-7	missense	H->Y
5581:11	I	3471591	-3.22	G-> A	Coding	spg-7	missense	A->T
144:51	I	4024815	-1.88	A ->T	Coding	unc-73	missense	I->N
9901:07	I	10905843	5.83	C-> T	Coding	srt-61	nonsense	Q->*
2724:16	I	12740726	13.97	G-> A	Coding	toe-4	missense	E->K
14790:4	I	12958736	15.64	C->A	Coding	fbxa-101	missense	G->V
14790:4	I	12959117	15.64	G-> A	Coding	fbxa-101	nonsense	Q->*

Table 3.9 Candidate gene list for *kor44*

MMP Rank:Occurrence	Chromosome	Physical location (bp)	Genetic position	Sequence change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
14625:4	IV	2581079	-	C-> T	coding	nhr-242	missense	G->E
6673:10	X	4697976	-6.90	C-> T	coding	hyl-2	missense	A->T
816:28	X	7095107	-1.87	C-> T	coding	C36B7.5	missense	D->N
19577:1	X	7126792	-1.83	C-> T	coding	C36B7.8	missense	H->Y
1233:24	X	9613800	1.42	C-> T	coding	C11E4.6	missense	T->I
1647:21	X	9786941	1.55	C-> T	coding	F36G3.1	missense	E->K
9625:08	X	10696407	2.37	C-> T	coding	F13E6.5	missense	H->Y
14652:4	X	11548881	4.36	G-> A	coding	mig-2	missense	D->N

Table 3.10 Candidate gene list for *kor45*

Chromosome	Physical location (bp)	Sequence change	Type of change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
I	4026499	G-> A	SNV	coding	unc-73	missense	A->V
I	5004630	G-> A	SNV	coding	dom-3	missense	E->K
I	6367272	A-> G	SNV	coding	K04F10.7	missense	N->S
I	6408211	G-> A	SNV	coding	san-1	missense	A->T
I	6494719	G-> A	SNV	coding	uri-1	missense	A->T
I	8624102	G-> A	SNV	coding	fer-1	missense	P->S
I	9982463	G-> A	SNV	coding	sys-1	missense	P->L
I	11566764	G-> A	SNV	coding	F15D3.4	missense	A->T
IV	3697067	G-> A	SNV	coding	Y37E11AL.3	missense	R->Q
IV	5312940	G-> A	SNV	coding	msp-59	missense	P->S
IV	5551907	G-> A	SNV	coding	T12E12.1	missense	V->M
IV	7603325	G-> A	SNV	coding	tag-80	missense	A->T
IV	10067952	G-> A	SNV	coding	cwn-2	nonsense	W->*
IV	10416726	G-> A	SNV	coding	F13B12.6	missense	P->S
IV	12565327	G-> A	SNV	coding	eak-7	nonsense	Q->*
IV	13082430	G-> A	SNV	coding	C39E9.8	missense	G->D
V	16918245	G-> A	SNV	coding	fbxa-206	missense	P->L
V	17461849	G-> A	SNV	coding	F20E11.5	missense	S->F

Table 3.11 Candidate gene list for *kor49*

Chromosome	Physical location (bp)	Sequence change	Type of change	Region affected	Corresponding gene	Amino acid mutation	Change in amino acid translation
IV	3615751	C->T	SNV	coding	unc-17	missense	G->E
IV	5842520	C->T	SNV	coding	C06E7.2	missense	P->L
IV	6585243	C->T	SNV	coding	F38A5.2	missense	R->Q
IV	10068291	C->T	SNV	coding	cwn-2	Nonsense	Q->*
IV	10590699	A->C	SNV	coding	C43F9.5	missense	F->C
IV	10908682	C->T	SNV	coding	ZK596.2	missense	D->N
IV	11024377	C->T	SNV	coding	alh-3	missense	L->F
IV	15324555	C->T	SNV	coding	tbx-39	missense	T->I
IV	17301481	C->T	SNV	coding	ifta-2	missense	S->N

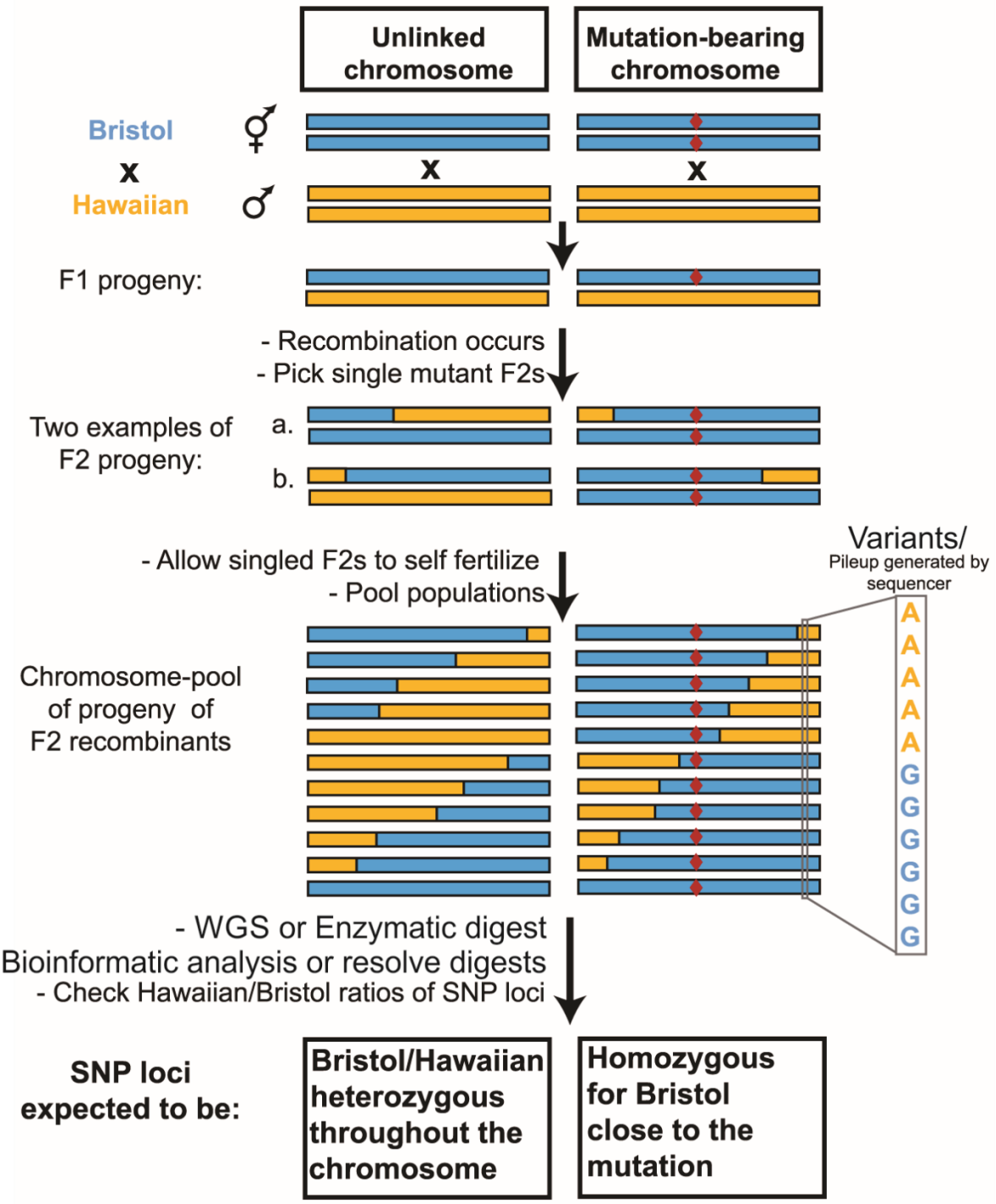

Figure 3.10 SNP mapping strategy. Crossing mutagenized strains with Hawaiian polymorphic strain introduces SNPs throughout the genome, including sites that add or remove a restriction enzyme recognition site. Recombination of these sites allows for linkage mapping of a causative mutation. Interval mapping used individual clonal populations of re-isolated recessive mutants. WGS pooled the populations together. Bristol = N2 strain, WGS = whole genome sequencing. Image modified from Doitsidou et al., 2010 and used under the terms of a Creative Commons Attribution License.

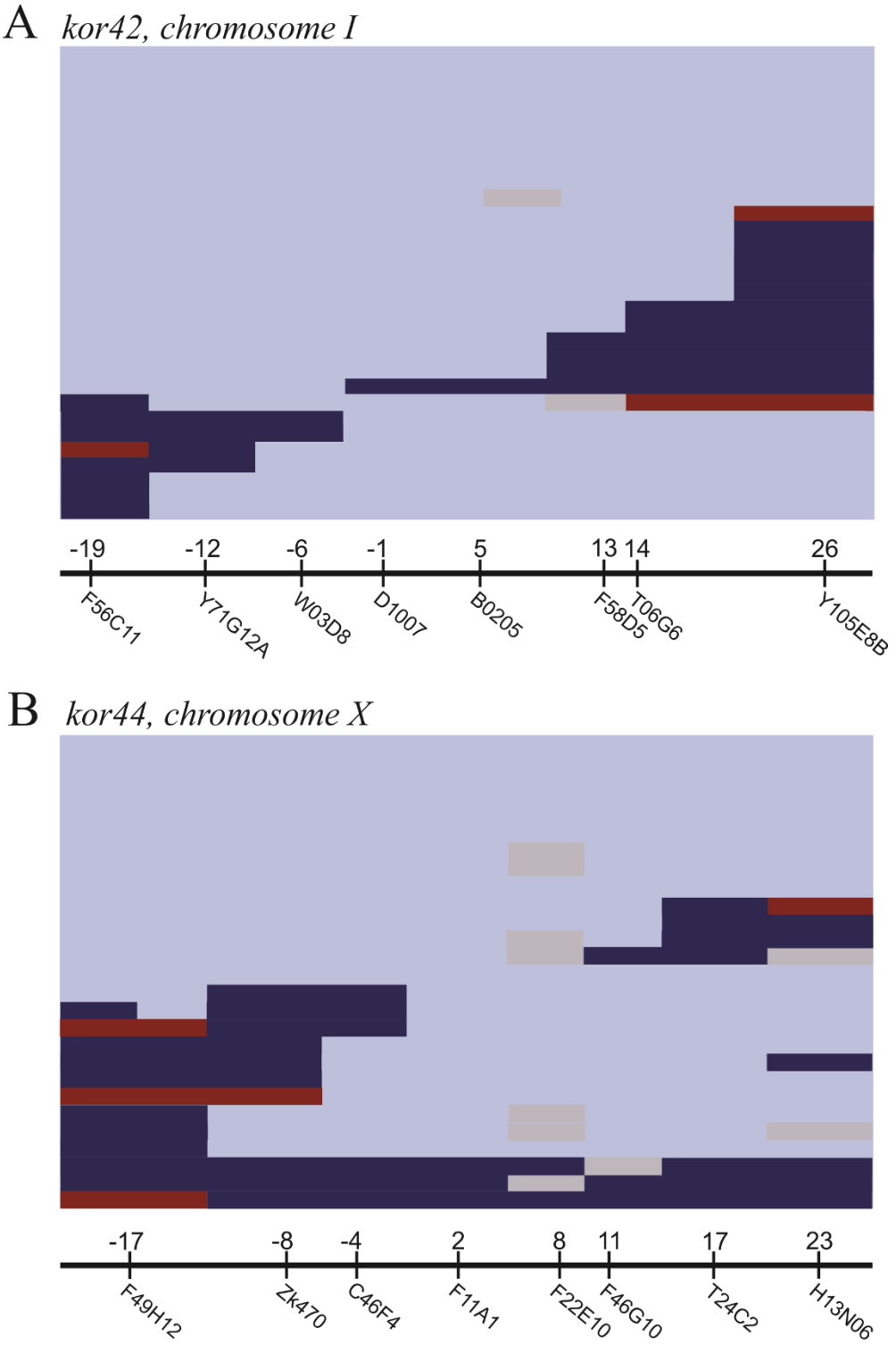

Figure 3.11 snip-SNP interval mapping results of (A) *kor42* and (B) *kor44*. Mutagenized strains were crossed with Hawaiian males. Each row or horizontal line represents a single clonal population from a mutant F2 parent. Bottom number line is a representation of the applicable chromosome, with the genetic location and clone name of each DraI snip-SNP site used (Davis et al., 2005). **Red** represents areas homozygous for the Hawaiian SNP, **light blue** areas are homozygous for N2/mutant parent SNP, **dark blue** is areas of heterozygous Hawaiian/N2 SNPs, and grey is inconclusive results due to PCR failure.

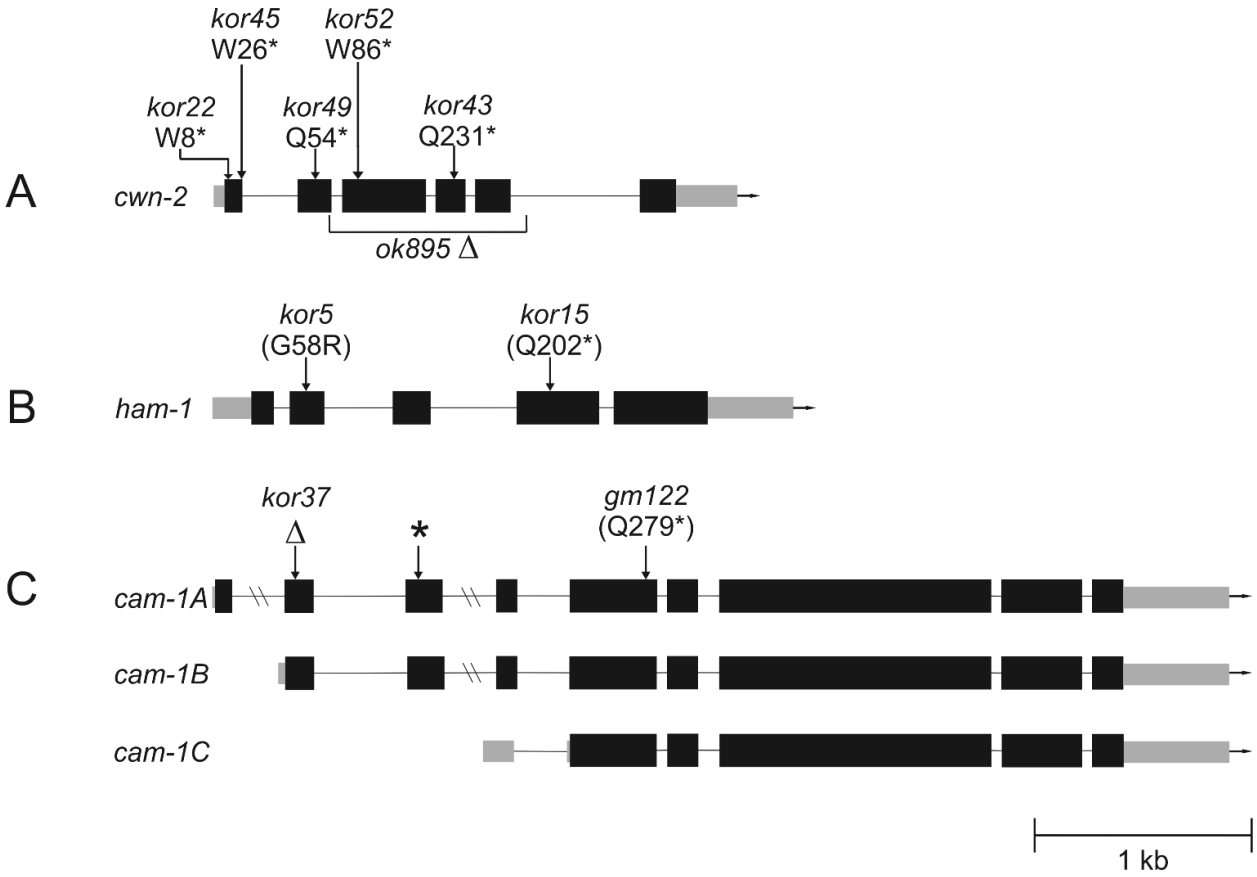

Figure 3.12 Schematic of genes indicating location and nature of mutations sequenced. Grey boxes are untranslated regions, black boxes are exons, and lines are introns. A combination of whole genome sequencing and Sanger sequencing was used to determine the nature of the causative mutations. (A) LGIV complementation group A strains have base pair transitions that introduce premature stop codons in *cwn-2*. *ok895* is a 905bp deletion that was used in complementation crosses. (B) LGIV complementation group B strains have mutations in *ham-1*. *kor5* is the same mutation as the previously reported *gm267* allele. (C) *kor37* is a single base pair deletion in *cam-1* that causes an immediate amino acid change at the deletion site, and then introduces a stop codon (asterisk) due to frameshift. This mutation has an obvious effect only in *cam-1A* and *cam-1B* isoforms. *gm122* is a premature stop codon and probable null used in the complementation cross. Hatch marks on introns indicate truncation of a very long sequence due to space limitations.

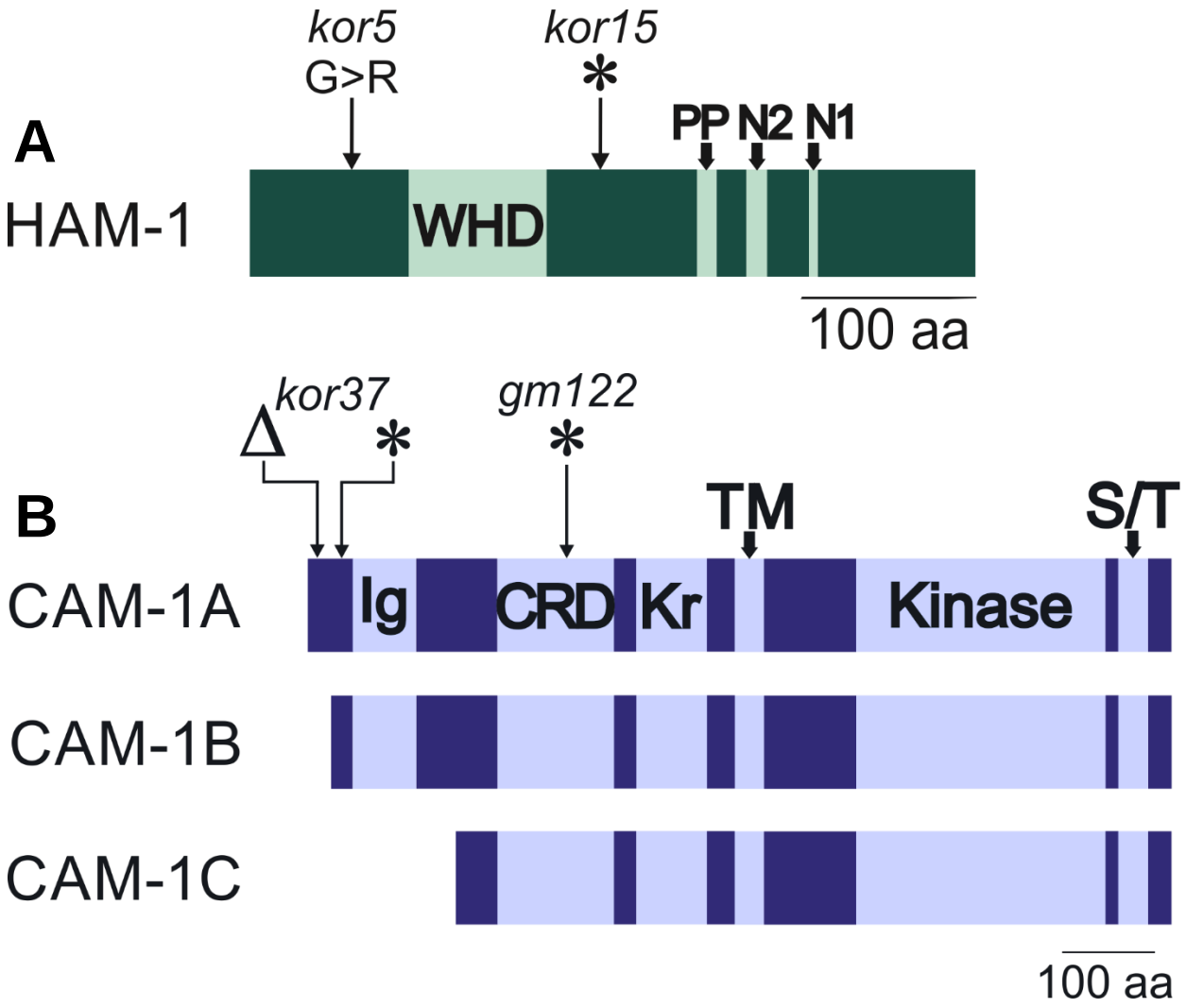

Figure 3.13 The effect of *kor5* and *kor15* on HAM-1 and *kor37* on CAM-1. (A) WHD=Winged helix domain. PP=Polyproline/SH3-like sequence. N1 and N2=Nuclear localization sequences. (B) Delta = single base pair deletion. Asterisk = early stop codon introduced as a consequence of frameshift. This mutation has an obvious effect only in CAM-1A and CAM-1B. *gm122* probable null included for reference. Ig=Immunoglobulin domain, CRD=Cysteine-rich domain, Kr=Kringle domain, TM=Transmembrane domain, Kinase=Kinase domain, S/T=Serine-Threonine-rich domain.

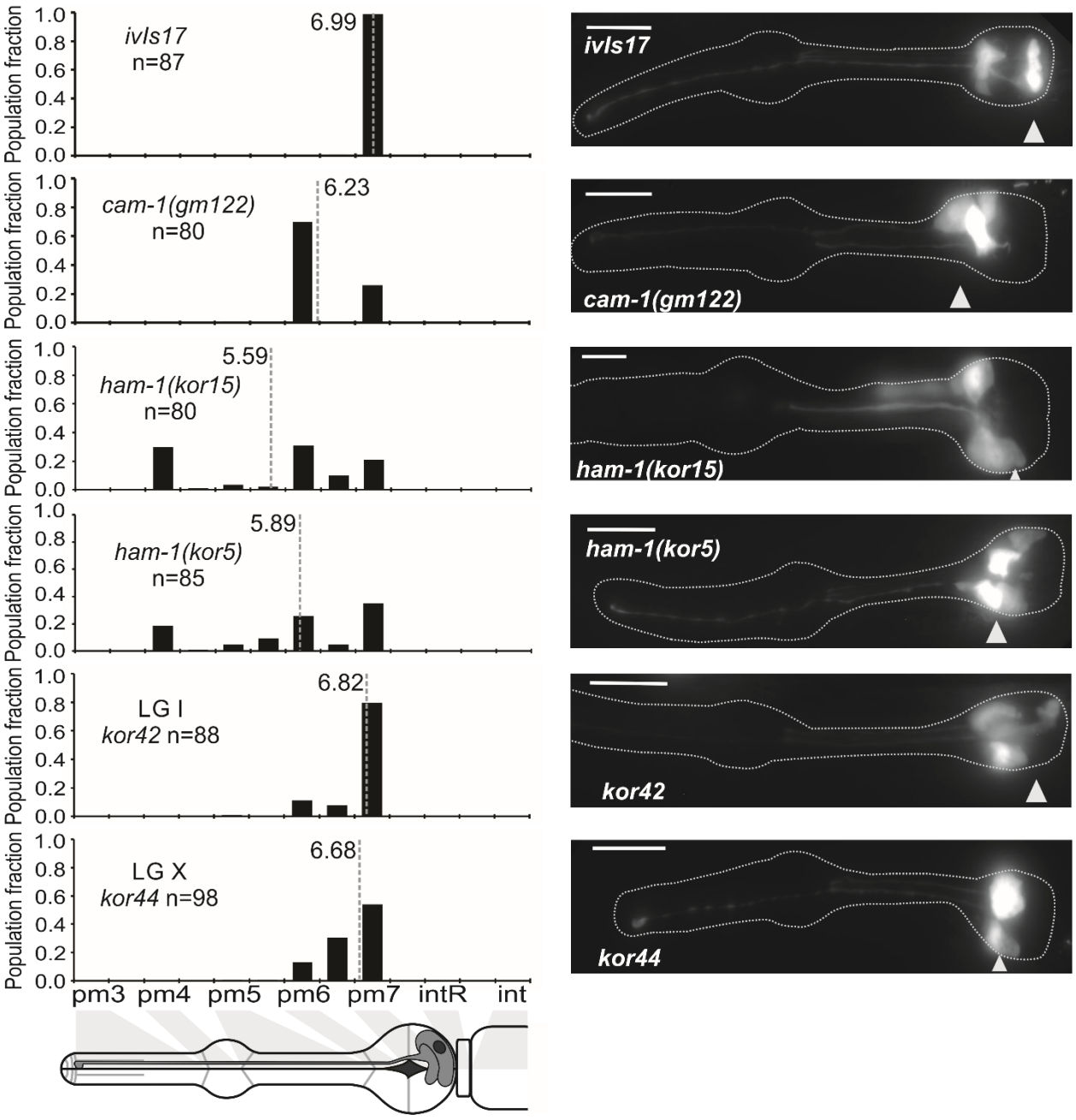

Figure 3.14 Quantitation of g1P gland cell position. (Left) Distribution of g1P cells relative to pharyngeal muscle sections in wildtype *ivIs17*, *cam-1*, *kor5*, *kor15*, and *kor42* and *kor44* genetic backgrounds. GD282 *ivIs17* [*phat-1*::YFP, *elt-2*::mTomato *rol-6(su1006)*]V was used as wild type. GD391 *cam-1(gm122)* was used to quantitate *cam-1* null mutant migration phenotype for comparison. X-axis is position of cell with regards to the pharyngeal muscle group it is closest to. Y-axis is portion of population whose g1P cell lies at that position. Grey dashed bar indicates the weighted mean for each population. (Right) Representative L1 or L2 larvae with gland cell defects. Left is anterior and right is posterior. Dotted line indicates the pharynx. Arrowhead indicates g1P position. Scale bars are 10µm.

Table 3.12 Under-migration penetrance and additional phenotypes observed in *kor5, kor15, kor42,* and *kor44*. Strains scored were outcrossed to N2 three times and maintained at 20˚C. Penetrance was calculated by dividing number of worms with under migration defect by total number of worms counted for each strain.

Strain and causative mutation	Approximate penetrance of under-migration defect	Other observed phenotypes
JDK114 *ham-1(kor15)*	0.79 (n=80)	Occasional egg-laying defect and gonad migration defect. Variable amount of dead eggs. Low penetrance of Nob-like, forked, or Utx-1-like tail defect.
JDK119 *ham-1(kor5)*	0.65 (n=85)	Low penetrance egg- laying defect. Nob-like, forked, or Utx-1-like tail defect.
JDK117 *(kor42)*	0.20 (n=88)	Low penetrance egg-laying defect/bag of worms.
JDK118 *(kor44)*	0.46 (n=98)	Low penetrance gonad migration defect. Egg-laying defect/bag of worms.

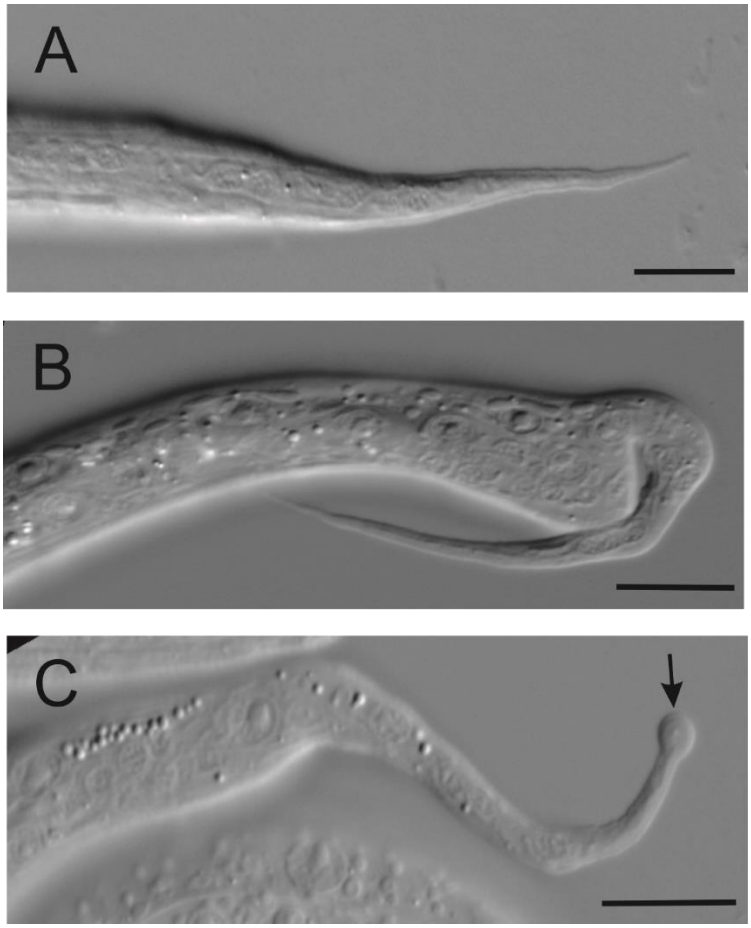

Figure 3.15 Tail defects found in strains JDK114 *ham-1(kor15); ivIs17* and JDK119 *ham-1(kor5); ivIs17*. (A) wildtype tail (B) Nob-1-like tail (C) Utx-1-like tail, with arrow pointing to cuticular ball in place of normal tail spike. Scale bars are 10µm.

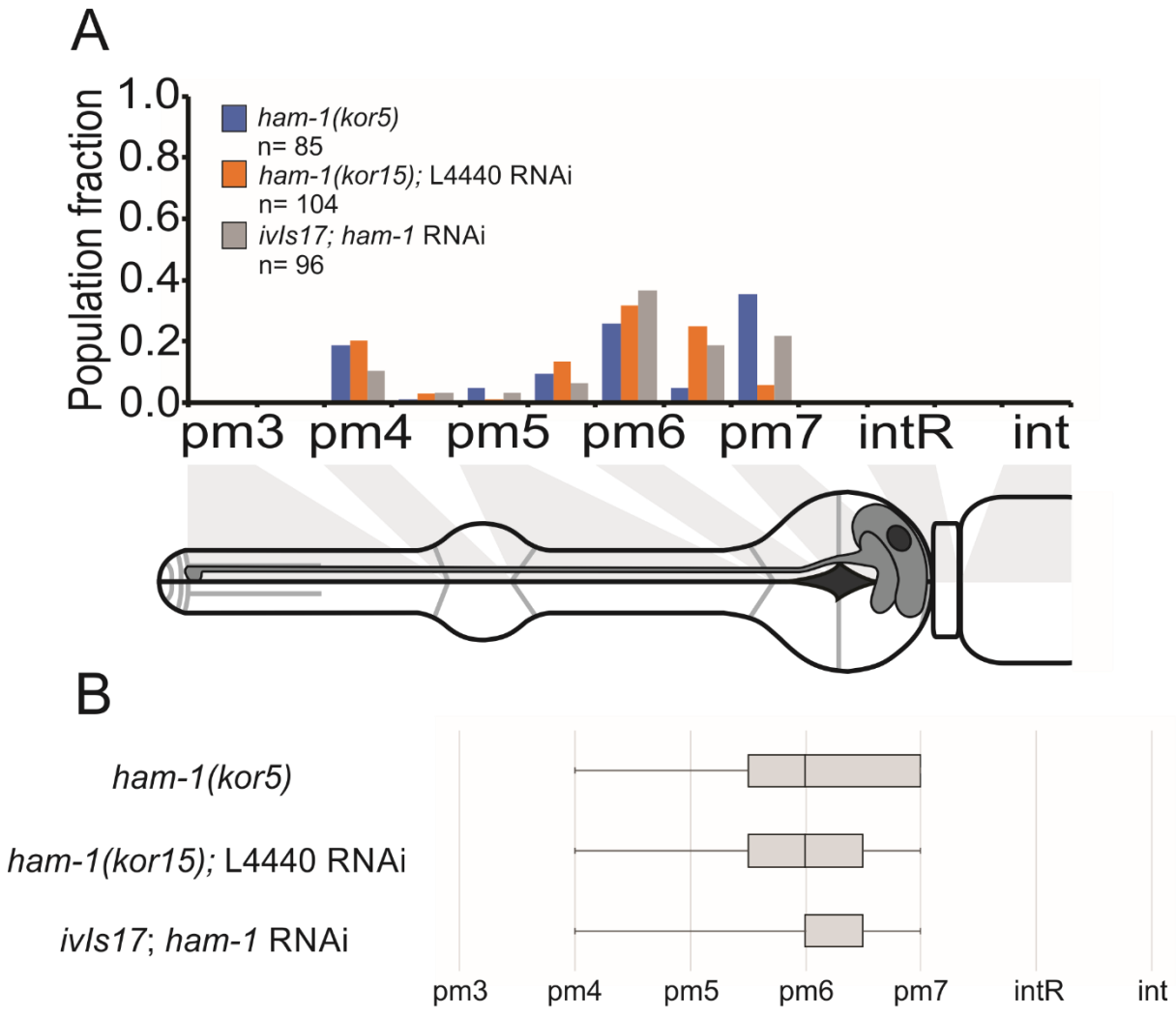

Figure 3.16 Bar graph and boxplots comparing the g1P distributions of *ham-1(kor5)*, *ham-1(kor15)*; L4440 RNAi, and ivIs17*; ham-1(RNAi)*. Horizontal axes represents relative pharyngeal muscle (pm) section the g1P was found. (B) Left-hand whiskers represent the minimum pm score, and right-hand whiskers represent the maximum pm score. Left edge of boxes, center marks, and right edge of boxes represent the Q1 lower quartile, median, and Q3 upper quartile, respectively. *ivIs17; ham-1(gm267)* maximum and Q3 are the same value. *ivIs17; ham-1*(RNAi) Q1 and median have the same score at pm6. While all three backgrounds have the same median, the overall distribution of *ham-1*(RNAi) more closely resembles that of *ham-1(kor15);* L4440 RNAi.

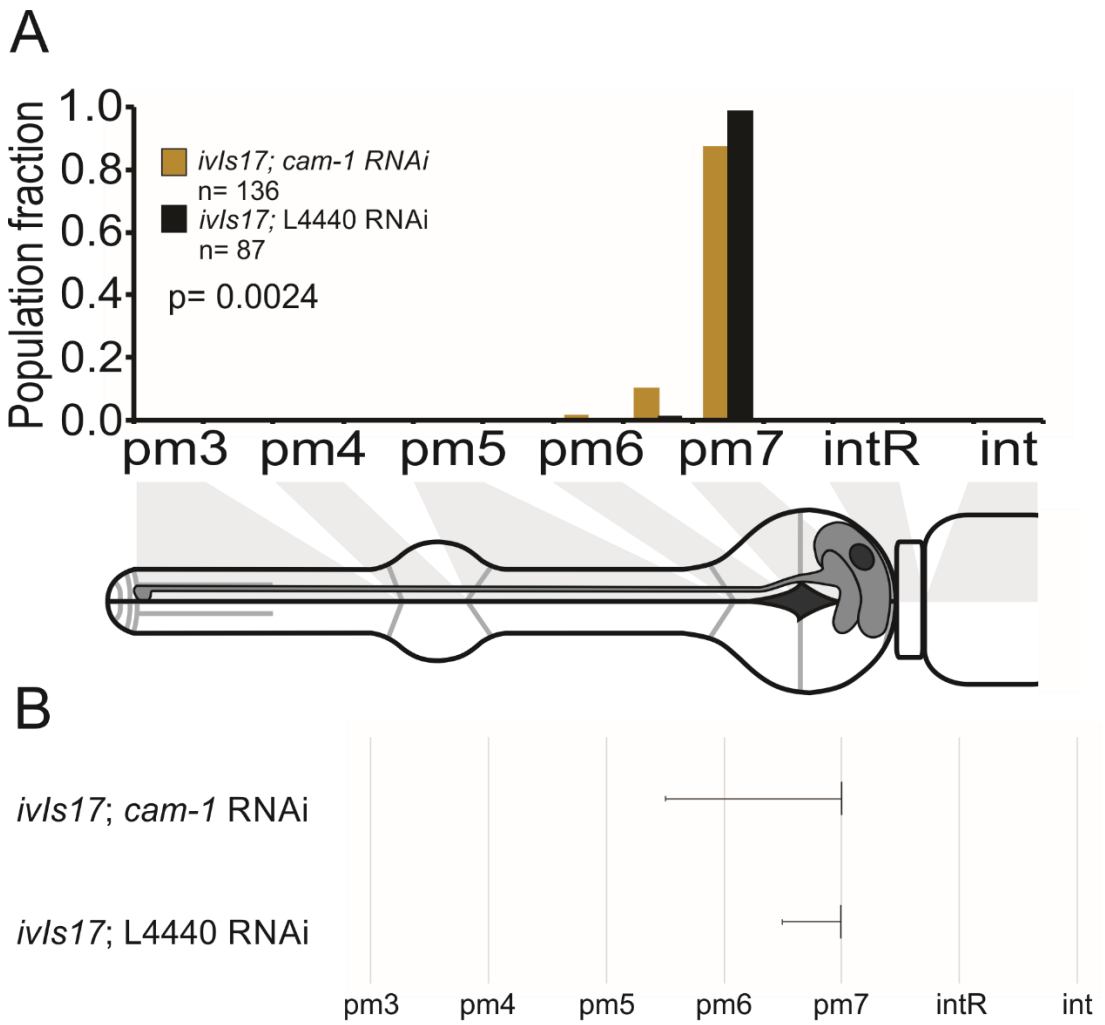

Figure 3.17 Bar graph and boxplots comparing the g1P distributions of two GD282 treatments *ivIs17*; *cam-1*(RNAi) and control *ivIs17*; L4440 RNAi. Horizontal axes represent relative pharyngeal muscle (pm) section the g1P was found in each genetic background. (B) Left-hand whiskers represent the minimum pm score, and right-hand line represents the maximum score, median, and interquartile range. According to the Mann-Whitney *U* test, these distributions are statistically different when α=0.05, with a p-value of 0.0024. This suggests that the samples scored for each treatment do not come from populations with the same distribution, and that *cam-1* RNAi confers a very weak gland cell phenotype.

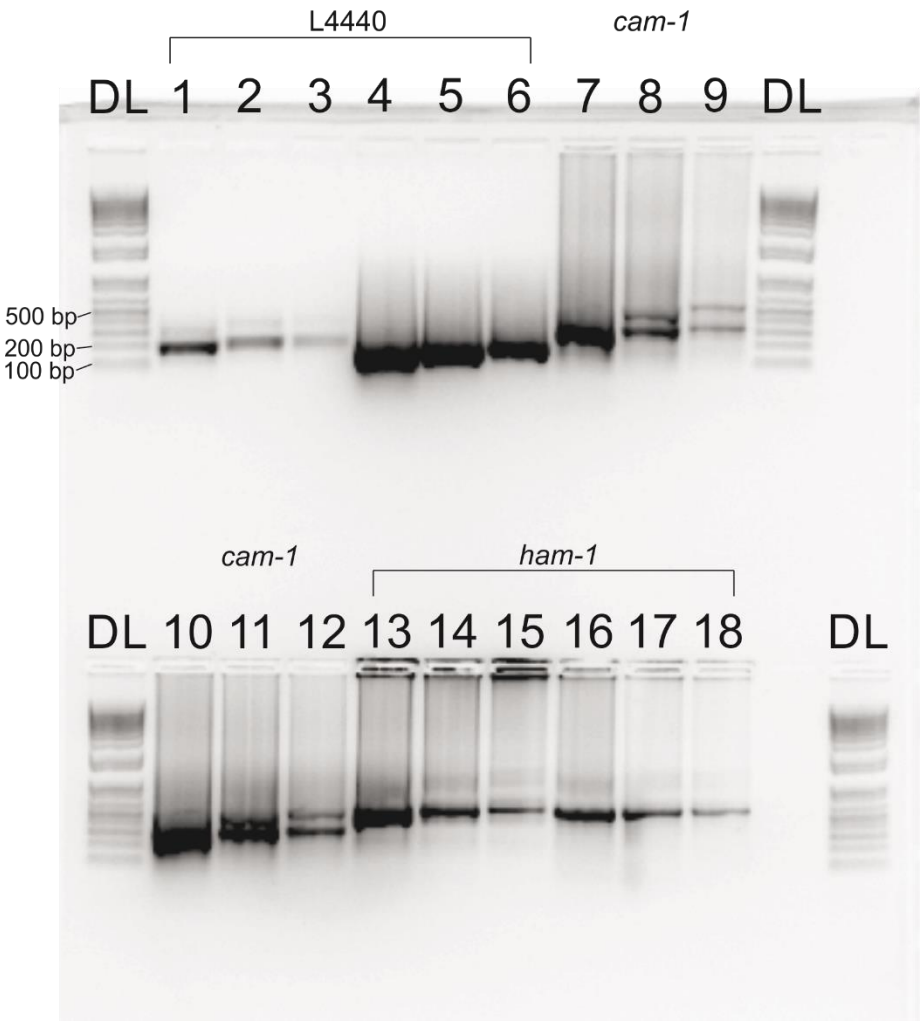

Figure 3.18 Electrophoretic gel of L4440, *cam-1*, and *ham-1* dsRNA. dsRNA was formed using the TranscriptAid T7 High Yield Transcription Kit (ThermoFisher #K0441). *cam-1 and ham-1* dsDNA templates had T7 promoters previously incorporated using primer sets oJDK414/415 and oJDK410/411, respectively. L4440 dsDNA was amplified using primers oGD30/31, which amplify endogenous T7 sites in the multiple cloning site of the L4440 plasmid. Both the dsRNA precipitate and supernatant fractions were resolved for each gene, and each fraction was diluted 8, 16, and 32-fold. L4440 expected band size is ~420 bps, *cam-1* expected band size is ~870 bps, and *ham-1* expected size is ~880 bps. (DL) DNA ladder, which is only used as an approximation of molecular weight due to different running properties than RNA. Lanes (1-3) are L4440 dsRNA precipitate fraction, (4-6) are L4440 dsRNA supernatant fraction, (7-9) are *cam-1* dsRNA precipitate, (10-12) are *cam-1* dsRNA supernatant, (13-15) are *ham-1* dsRNA precipitate, and (16-18) are *ham-1* dsRNA supernatant. Strong competing bands in *cam-1* lanes may have contributed to poor RNAi potency and weak knockdown of *cam-1* expression.

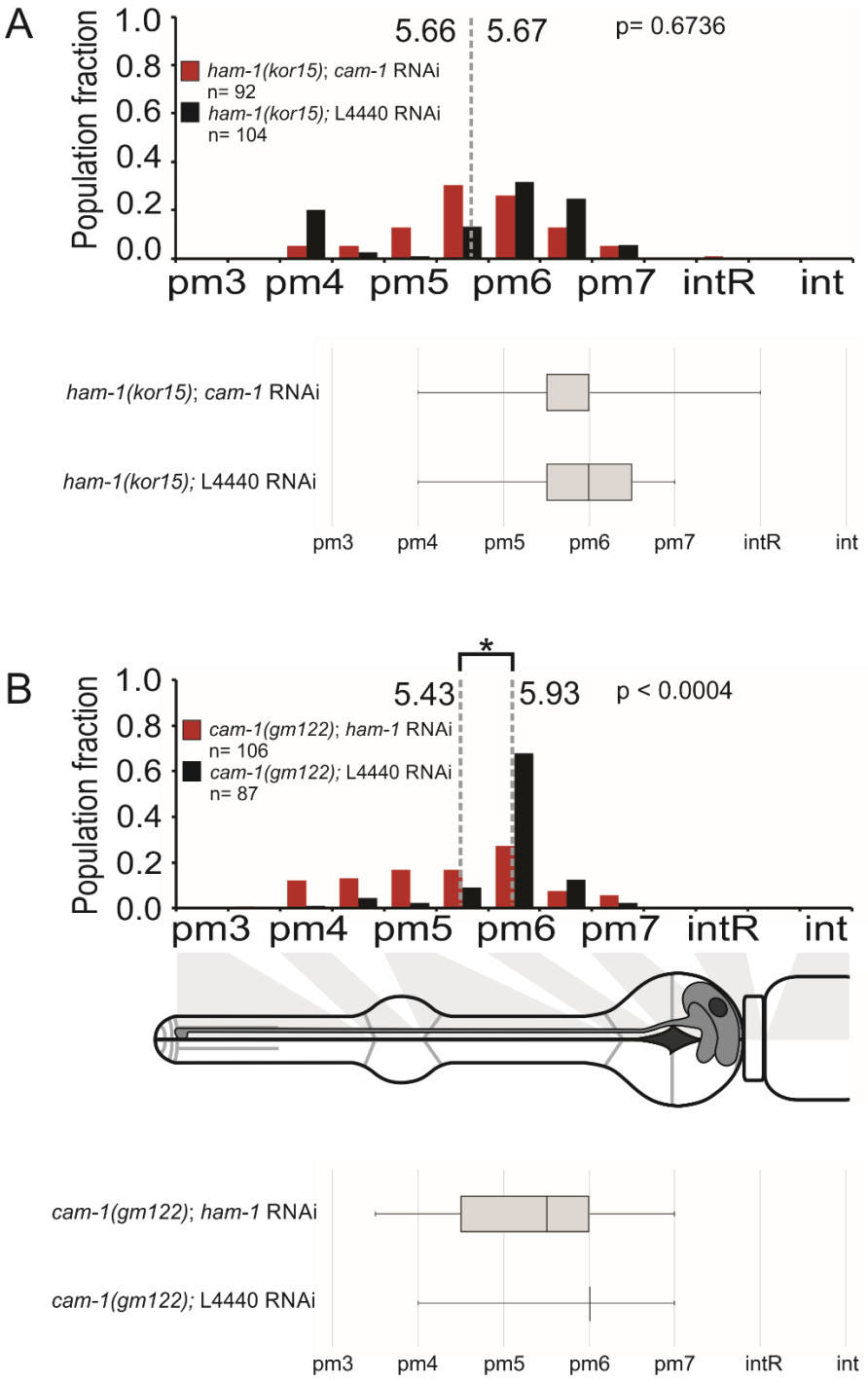

Figure 3.19 Quantitation of g1P gland cell position in (A) *ham-1(kor15);cam-1(RNAi)* double mutant and (B) *cam-1(gm122); ham-1(RNAi)* double mutant. Gland cell scoring was possible by using the *ivIs17* insertion. (Above) Distribution of g1P cells relative to pharyngeal muscle sections. Red bars are double mutant, and black bars are single mutant with control L4440 RNAi. Dashed grey bars represent weighted means for each sample. P-values were calculated using the Mann-Whitney *U* Test, suggesting that *ham-1(kor15); cam-1(RNAi)* distribution is not different from its control, while *cam-1(gm122); ham-1(RNAi)* is different from its control. This inconsistency may be due to the poor quality of *cam-1* dsRNA. (Below) Boxplots of the respective sample distributions.

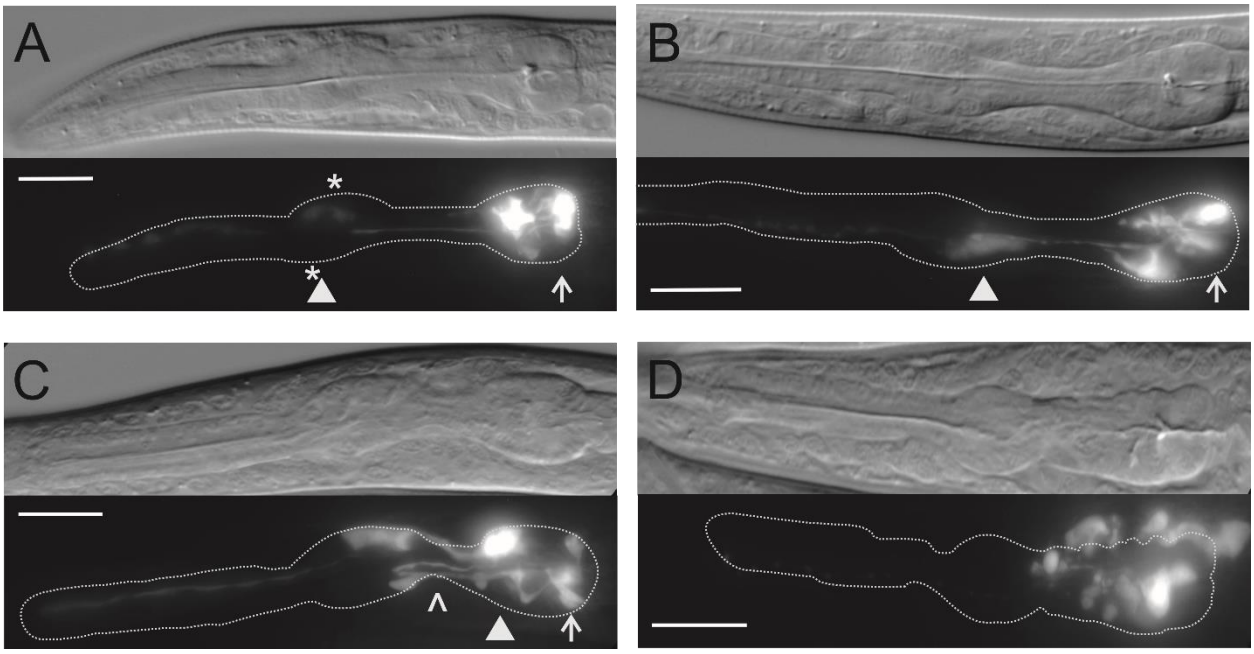

Figure 3.20 Examples of pharyngeal defects observed in *cam-1*(*gm122*); *ham-1*(*RNAi*) mutants, including DIC image (upper) and corresponding Z-stack captured with a YFP filter (lower). Gland cells were visualized with YFP fluorescence from the *ivIs17* insertion present in each strain. Arrowheads indicate position of g1P cell, and arrows indicate examples of g1a cells aberrantly located or expanded in the pm7 region. (A) and (B) Examples of asymmetric metacorpses/anterior bulbs; the asterisks in (A) designate the misaligned centers of each side of the bulb. (C) Example of an isthmus that failed to elongate (carat). (D) Example of isthmus that failed to elongate, as well as a misshapen terminal bulb and fragmented gland cells outside of the pharynx. Scale bars are 10µm.

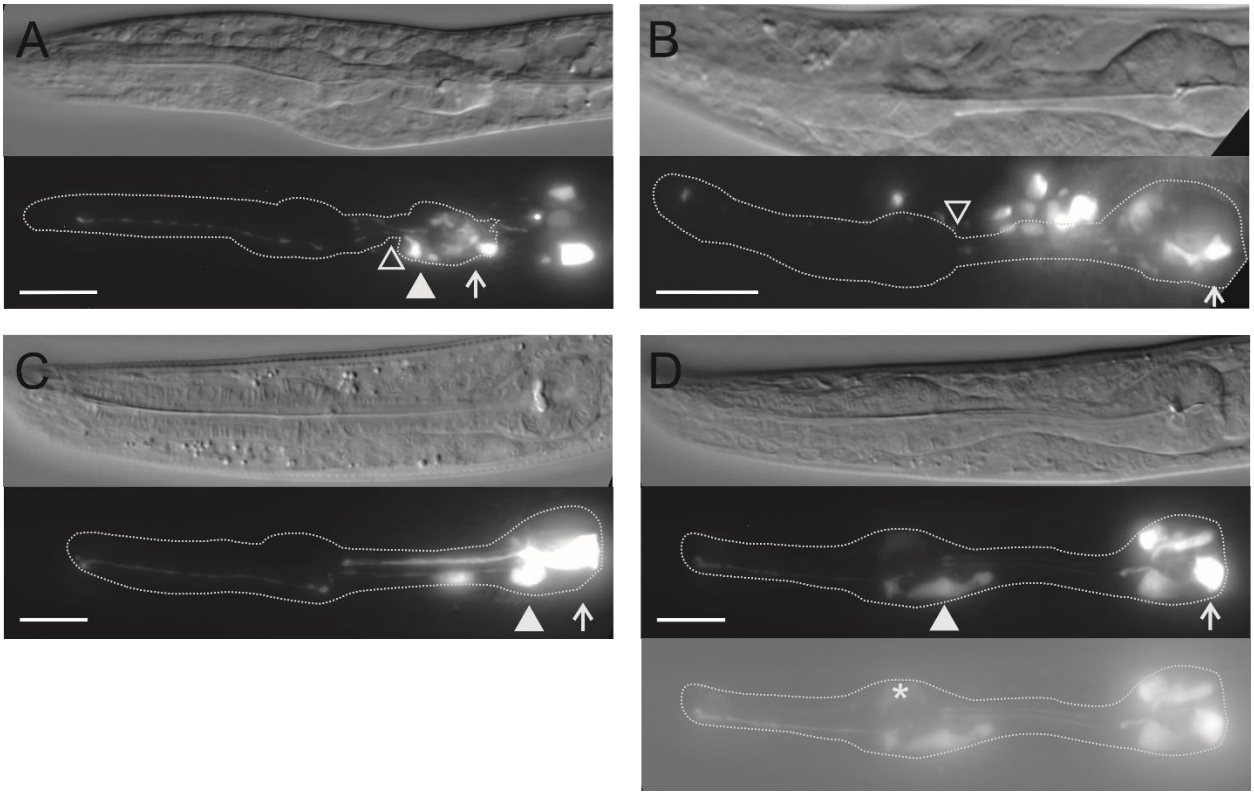

Figure 3.21 Examples of pharyngeal defects observed in *ham-1(kor15); cam-1(RNAi)* mutants, including DIC image (upper panel) and corresponding Z-stack captured with a YFP filter (lower panel). Gland cells were visualized with YFP fluorescence from the *ivIs17* insertion present in each strain. Arrowheads indicate position of g1P cell, and arrows indicate examples of g1a cells aberrantly located or expanded in the pm7 region. Scale bars are 10µm. (A) and (B) Examples of short and weakened isthmus/pm5 region. Open arrowheads point to low-integrity/underdeveloped areas of muscle. (A) Buckled half of isthmus. Gland cells are extremely fragmented, and some have apparently over migrated. The pharyngeal-intestinal valve is also undetectable. (B) Burst section of pharynx near the pm4/pm5 boundary. Some fragments of gland cell(s) and other content have been pushed outside of the pharynx. The pm3/4 muscle striations and basement membrane are also extremely difficult to distinguish. (C) Example of expanded g1a that invades the pm7 area (arrow). The g1P extension appears bent near the pm4/5 border, and has an expanded bleb within the isthmus. (D) Apparent duplication of the g1P cell. The bottom-most panel is gamma-corrected to emphasize the duplicated cell (marked with an asterisk). The extensions of both g1Ps appear to be disjointed in the pm3/4 region.

4 Discussion

4.1 Complementation results

Most complementation crosses gave consistent results between replicates. Crossing mutations into *kor22, kor45, kor49,* and *kor52* gave consistent under-migration ratios. Some complementation crosses were assigned a variable result. Inconsistent penetrance of the under-migration defect was observed between individual crosses in these series. For example, in Table 3.2, *kor15/+* X *kor43* yielded a cross with 1% mutant progeny, and another cross with 41% mutant progeny. There is a large difference between the mutant ratios of each cross, and averaging them would not be an accurate way to decide if complementation occurred.

Two crosses yielded more than the maximum of 50% mutants. These were the *kor15/+* X *kor5* and *kor43*/+ X *kor5* crosses. The *ham-1(kor5)* and *cwn-2(kor43)* alleles appear to deviate from hypomorphic allele behaviour and their *trans* progeny may have a neomorphic or gain-of-function behaviour. Since we did not know the nature of the lesion prior to sequencing the deficiency, crosses required to confirm the null or hypomorphic categorization of an allele was not rigorously tested. Although more in-depth genetic analysis is required to understand their function, the nature of each mutation may provide clues as to how the predicted gene products would function. For instance, *kor5* is the full-length HAM-1 protein with an amino acid change from glycine to arginine at the C-terminus. Arginine is positively charged under normal physiological conditions so this is a non-conservative amino acid change. This may affect protein folding. Arginine can also interact with other proteins through its guanidium group (Sokalingam et al., 2012). Post-translational modifications of arginine can also alter protein-protein interactions and signalling cascades (McBride and Silver, 2001). Without understanding the

structure of the HAM-1 protein it would be difficult to predict why this amino acid location is important. It could be that the *kor5* extra arginine allows inappropriate protein interactions to occur. This would create a neomorphic allele. It is also possible that the product would antagonize normal function. This would be an antimorphic allele. Alternatively, the substitution may prevent a negative regulation creating a more common hypermorphic allele. In these cases, only one copy of the allele would be needed to alter the function of the wild-type protein or pathway components and would behave in a dominant fashion.

It has been previously documented that *cam-1* function relies on a finely tuned balance between activation of a non-canonical Wnt pathway and inhibition of the canonical Wnt pathway (Chien et al., 2015; Green et al., 2008a). This may explain the occurrence of allele dose sensitivity generated in hemizygous strains. Furthermore, there is precedent for the requirement of a sensitized genetic background in order to observe *cam-1*-dependent defects in vulva cell precursor polarity (Forrester et al., 1999; Green et al., 2008a). Although not always consistent, this may explain the apparent lack of complementation in some allelic combination of the *cwn-2* and *ham-1* alleles. The progeny may display a phenotype caused by a complex synthetic enhancement.

The *kor43* premature stop codon creates a truncated CWN-2 product of 231 amino acids (see Figure 3.12). It is difficult to predict how this would change protein folding and post-translation modifications of the Wnt but it is possible that a truncated protein may persist. An incomplete ligand could be defective in stimulating cellular signals. If it is still able to interact with receptors, this "dead" version of the ligand could block wild-type ligands from binding. This would potentially generate a semi-dominant hypomorphic allele or antimorphic allele, respectively.

4.2 Hawaiian mapping-by-sequencing troubleshooting

If the Hawaiian mapping-by-sequencing protocol is to be used in future experiments, the procedure will need to be modified. Ensuring all strains are backcrossed to N2 wildtype can reduce the amount of background mutations in mutagenized strains. This will become especially important as strains possessing heterozygous causative mutations will be analyzed. From the whole genome sequencing experiment in this thesis, backcrossing to N2 at least two times appeared beneficial. *kor15*, which was outcrossed twice prior to sequencing, had the least number of 147 heterozygous variants. *kor5*, which was outcrossed once, had 229 heterozygous variants. *kor42*, *44*, *45*, and *49* were not outcrossed prior to sequencing. They had heterozygous calls ranging from 182 to 369 variants, with an average of 260 het variants per uncrossed strain.

Uniform representation of each recombinant within a pooled sample may also help improve the Hawaiian mapping method. Although a random subsample of pooled pellet was extracted for each sample, it is hard to guarantee that the subsample was truly random. Although care was taken to ensure that each recombinant was equally represented in the pooled population, it can be that slow growing or low fertility biased the recovery of healthier recombinants. To reduce bias, each recombinant could have gDNA isolated, quantified, and combined at equal levels. In this thesis, some bias could have been removed by using the entire pooled sample for extraction, instead of the subsample that was used. The pooled samples in this thesis were large and exceeded extraction reagent capacities. Future experiments could plate F2 mutants on small 60x15mm petri plates rather than 92x15mm plates. This would reduce the amount of material generated for extraction. Removal of allele bias unrelated to the gene being mapped would

potentially create less variability in the Hawaiian allelic ratios and, in turn, produce more reliable and more tightly linked mapping intervals.

4.3 *kor44* LGX recombination

kor44 interval mapping and the CloudMap Hawaiian mapping scatterplot show a large area of uniform recombination at the center of the chromosome (Figures 3.9 and 3.11B). This pattern suggests a large deletion or rearrangement of the X chromosome. There is no copy number variation in any strain that was sent for whole genome sequencing, suggesting that the causative mutation is not a large deletion. A complex rearrangement, such as an inversion, was more difficult to detect in whole sequencing alignments. Similar to a genetic balancer (McKim et al., 1988), this could prevent local recombination or create recombination hot spots, making it difficult to map the mutation (Kormish and McGhee, 2005).

4.4 Candidate genes on LGI and LGX

Hawaiian interval mapping narrowed down the intervals of the causative mutations for *kor42* LGI and *kor44* LGX. Combining early mapping data and the homozygous variants called for the chromosomes, *unc-73* I and *mig-2* X were possible candidates. These are a guanine nucleotide exchange factor and Rho GTPase, respectively, that have been shown to directly interact with each other (Kishore and Sundaram, 2002). The amino acid changes for each gene, however, are not entirely convincing of causing a change-of-function in the protein. *unc-73* changes an isoleucine to an asparagine, and *mig-2* changes an aspartic acid to asparagine. The other candidate genes on LGI and LGX may have more severe changes in function. *srt-61* on LGI is a serpentine receptor. Its mutation introduces a stop codon in the middle of the gene. The Cloudmap trends for *kor44* suggest a location of 7.5 Mb to 9.5 Mb on the I linkage group (Figure 3.8). This favours the *srt-61* location of 10.9 Mb over *unc-73* at 40.2 Mb. Indeed, previous work

performed by Dr. Kormish in the Gaudet lab examined an *unc-73*(*e936*) I; *ivIs12* [*phat-1*::YFP *elt-2*::mTomato *rol-6*(*su1006*)] II strain and no migration defects were observed (n=205). *unc-73* has been found to code for many distinct protein isoforms (Steven et al., 2005) but the *e936* allele, which results in a substitution removing a splice site at location 40,219,530 of the a, b, and i isoforms, is consistent with affecting the same isoforms with the substitution six exons down stream at 40,248,150 in the *kor42* allele.

The lesion detected in *mig-2* was a C to T transition at 11,548,881 resulting in a D to N amino acid change. This would result in a negative to a polar side chain. *mig-2(mu28)* is a premature stop codon at 11,548,862 in the second exon just upstream to this lesion. Previous work by Dr. Kormish did not detect gland cell migration defects for *mu28* (n=25) or when combined with the *ced-10*(*n1993*) allele (n=17). The mapping trends of the Hawaiian ratios in Figure 3.9 weakly point to a location of 10 to 10.5 Mb. This is more consistent with the location of the H to Y amino acid change for F13E6.5 at 10.7 Mb on the X. F13E6.5 has been recently named *plpp-1.3* based on its sequence homology to human phospholipid phosphatase 1. Its mutation changes a histidine to a tyrosine. Histidine is commonly involved in protein-protein interactions and the close to physiological pH pKa of its side chain allows it to be a common hydrogen donor in the enzymatic cores of proteins (Liao et al., 2013). Replacing histidine with tyrosine could be affecting protein folding and/or removing a key catalytic function of the phospholipid phosphatase.

4.5 *ham-1* is a novel factor in pharyngeal development

From the genes identified from the forward genetic screen so far, *ham-1* is a new factor involved in pharyngeal development and g1P gland cell migration. It was first discovered in a forward genetic screen for HSN motor neuron development (Desai et al., 1988). The gene

symbol and name refer to the HSN abnormal migration, where the cell bodies prematurely stop migrating along their anteriorly-destined path. *ham-1* has so far only been studied in-depth in select neuronal cell lineages, such as the PVQ/HSN/PHB, RIM, REM, and Q neuroblast lineages. It has also been implicated in pharyngeal interneuron I2 development (Frank et al., 2005). In *ham-1* mutant backgrounds, it was found that duplications of neurons occurred, while sister cells are sometimes missing. Because these cells are almost always born from asymmetrical divisions along the anterior-posterior axis, the general theory is that *ham-1* somehow contributes to the proper polarization and/or division of some asymmetrical cell divisions. HAM-1 has been suggested to position the mitotic spindle, myosin distribution, and cleavage plane of these divisions (Feng et al., 2013; Frank et al., 2005; Leung et al., 2016). When this function is abolished, any daughter cells fated to undergo apoptosis may survive, and the sizes of both sister cells change. In this regard, *ham-1* can be considered a pro-survival factor and indirectly affects cell fate (Baumeister et al., 1996; Desai et al., 1988; Guenther and Garriga, 1996; Singhvi et al., 2008). Changes of cell fate within the pharynx could affect organ polarity and provide improper environmental or intercellular cues for the g1P cell to stop migrating.

Until fairly recently, HAM-1 was considered a novel protein. It has similarity to the N-terminus of *Drosophila melanogaster* Knockout, which is involved in larval muscle development. HAM-1 shares slightly more similarity with the human and mouse protein STOX1 (Frank et al., 2005). STOX1 is part of the winged-helix protein family, and has been implicated in pre-eclampsia, Alzheimer's disease etiology, and proliferation of the epithelial cells of the inner ear (Nie et al., 2015; van Dijk et al., 2010, 2005). Through comparison with STOX1, HAM-1 was found to contain a winged helix domain, which mediates interaction with DNA and suggests transcription factor activity. Two nuclear localization signals were also found near the

C-terminus, supporting this function (Leung et al., 2016; Feng et al., 2013). This contrasts previous HAM-1 localization studies, where immunostaining only showed the protein localized at the periphery of the cell cortex (Frank et al., 2005). More recently, the nuclear localization of HAM-1 was demonstrated through visualization of a GFP fusion transgene. A subcellular fractionation experiment also detected endogenous HAM-1 in the nuclear fraction. Leung and colleagues (2016) also identified a polyproline/SH3 domain in its amino acid sequence, which suggests possible protein-protein interaction and interaction with the cortex. Cortex localization is not essential for development of all neurons, however. It is unclear if cortex localization is required for pharyngeal development, either. *ham-1(kor5)* has a non-synonymous amino acid change that does not affect any identified domain. *ham-1(kor15)* produces a truncated protein that includes the winged helix domain, but not the nuclear localization or polyproline signals.

ham-1 has been shown to interact with a handful of other genes. *unc-73* and *mig-2* have been suggested to function in the same events of HSN motor neuron development as *ham-1*, including migration and axonal outgrowth by controlling cytoskeletal rearrangement (Desai et al., 1988). *unc-73* and *mig-2* have so far appeared not required for pharyngeal development, however. In establishing cell asymmetry, *pig-1* is the only identified target of *ham-1* so far. *pig-1* is a serine-threonine kinase that is orthologous to the vertebrate Maternal Embryonic Leucine zipper Kinase (MELK). MELKs can modulate intracellular signalling and affect processes such as cell cycle and cell division (Jiang and Zhang, 2013). PIG-1 is expressed in the cytoplasm and, like *ham-1*, contributes to proper spindle and myosin positioning, regulated cell death, and therefore cell fate (Chien et al., 2013; Cordes et al., 2006; Pacquelet et al., 2015). It would be worthwhile to examine g1P migration and overall pharyngeal development in a *pig-1* background and see if it exhibits a similar phenotype, which may provide clues to piecing

together a pathway involving *ham-1*. A *pig-1* pharyngeal phenotype may also provide new candidate genes for establishment of pharyngeal polarity. For example, it has been shown that *pig-1* acts in the same genetic pathway as *par-4* (Chien et al., 2013, Pacquelet et al., 2015). *par-4* is orthologous to LKB1, a serine-threonine kinase that functions with AMP-activated kinase to polarize epithelial and neuronal cells (Nakano and Takashima, 2012).

4.6 *ham-1* probably functions in a different pathway from *cam-1*

Based on my RNAi experiments, *ham-1* and *cam-1* appear to function in different genetic pathways, at least regarding pharyngeal development. The weak *cam-1* dsRNA knock-down may have distorted the *ham-1(kor15); cam-1(RNAi)* results. It has been difficult to achieve optimal knock-down of *cam-1* using RNAi previously in the lab. Some cell types, particularly neurons, are resistant to RNAi (Asikainen et al., 2005). The gland cells display a glial-like property and may be closely related to the neuronal cell sub-type. RNAi knockdown in the gland cells is possible but not as robust as other cell types such as the intestine or pharyngeal muscle. RNAi can be made more effective in the neurons by using a sensitized genetic background. For example, an *eri-1* background can increase the amount of time dsRNA can reside in a cell before it is degraded and may improve interference potency (Kennedy et al., 2004). Expressing SID-1, a dsRNA transporter, in neuronal cells can increase their uptake of dsRNA and restore some function of RNA interference (Calixto et al., 2010; Li et al., 2015). Using another method to remove *cam-1* function may be required to properly analyze this double mutant.

Based on these results, then, during my graduate work I have been unable to identify conclusive gene candidates in the *cam-1* genetic pathway. The *kor37* allele is an allele of *cam-1* , however the resulting frame shift and early stop codon creates a truncation that removes major domains or motifs that have been identified for the receptor. The *kor37* single base pair deletion

affects the protein sequence of CAM-1A and CAM-1B, but not isoform C. The CAM-1 isoforms are reported to have different expression patterns (Koga et al., 1999; Kennerdell et al., 2009). CAM-1A appears to be expressed in some head muscles and neurons. CAM-1B is expressed in some neurons, pharyngeal muscle, body wall muscle, seam cells, and distal tip cells. Expression data for CAM-1C has not been reported. *kor37* affecting gland cell placement could mean that the expression pattern of isoform A, isoform B, or a combination of the two is required for proper pharyngeal development. CAM-1C is also lacking the immunoglobulin domain. It is possible that the immunoglobulin domain is required for CAM-1 receptor function during pharyngeal development.

There is little new functional information conferred by *kor37*, however, and the role of *cam-1* in pharyngeal development is still unknown. It is possible that *cam-1* is simply acting as a Wnt sink and sequestering Wnt ligand, and *cam-1* does not actually perform any intracellular signalling itself. Early characterization of *cam-1* showed that the kinase domain, and sometimes the entire intracellular portion of the gene, were not required for proper cell motility or outgrowth (Forrester et al., 1999, Kim and Forrester, 2003, Song et al., 2010). A recent comprehensive study showed that the *cam-1(ks52)* deletion allele may be functioning more like a hypomorphic allele in addition to its defect in kinase activity (Chien et al., 2015). Another study showed that there is reduced catalytic activity in the kinase domain of vertebrate RORs (Bainbridge et al., 2014). The enzymatic activity has been lost and the domain instead acts as a scaffold to recruit other biochemically active proteins. The Bainbridge study did find catalytic activity in CAM-1 kinase, however. Taken together, these ideas may suggest the CAM-1 kinase domain is not required for pharyngeal development. CAM-1 could alternatively cluster in a co-receptor complex with other factors, and the other factors could intracellularly activate

downstream components. For example, it has been suggested that CAM-1 complexes with *mig-13*/LRP-like and *mig-1*/Frizzled in neuronal development (Kennerdell et al, 2009) or *vang-1*/Van Gogh in vulval polarity (Green et al., 2008a).

There has been very little work investigating the serine-threonine kinase domain of CAM-1. It is difficult to say whether this domain is catalytically active or not, or whether it would have any effect on g1P or pharyngeal development. While the serine-threonine kinase is apparently not required for CAN neuron migration, it may play a role in HSN neuron migration: when a *cam-1* construct with a deleted ser/thr domain is introduced into the *cam-1(gm122)* background, the anterior movement of the HSN cell body stops prematurely, similar to a premature stop of the g1P gland cell (Kim and Forrester, 2003). Testing ser/thr deletion constructs may provide similar issues as the hypomorphic *cam-1(ks52)* kinase deletion. Neither the kinase nor ser/thr deletion constructs affect CAN migration, however. This suggests that the missing domains do indeed provide a biological function rather than the construct reducing the overall amount of CAM-1 in the system. Thus, at this point, I think the *cam-1* kinase and serine-threonine domains could be important for some aspects of embryonic development.

4.7 Strains may be temperature sensitive

Differences in g1P under migration severity were noticed when strains were subjected to both an imaging protocol and a dsRNA injection protocol. For the imaging protocol, strains were kept at a constant 20˚C prior to screening and imaging. The injection protocol incubates worms at 15˚C until they are adults, and then switches to a 20˚C temperature post-injection. After noticing some gland morphology defects between the two, such as increased incidence of extension morphology and cell body morphology defects, another incubation and screening protocol was completed. This protocol mimics the temperature shift of the injection protocol,

where worms are grown at 15˚C then transferred to individual plates and kept at 20˚C until progeny have hatched. It was expected that the g1P distributions of the temperature shift would roughly match those of the injected worms, of course keeping in mind that the trauma of injection is a variable that could still slightly affect gonadal and progeny development.

Comparing the imaging, injection, and simulation g1P distributions of GD391 *cam-1(gm122); ivIs17* and JDK114 *ham-1(kor15); ivIs17* showed that the progeny of worms at a constant 20˚C exhibit more variability in g1P position (Figure 4.1). Although the variation increases, the direction of the variation is different in the two strains: the under-migration severity decreases at 20˚C in the *cam-1* background, while the under-migration severity increases at 20˚C in the *ham-1* background. That is, the g1P was more anteriorly in *cam-1* at 20˚C compared to parents being grown in a 15˚C/20˚C mix, and the g1P was more posteriorly in *ham-1* at 20˚C compared to the 15˚C/20˚C mix. It appears that the temperature a hermaphrodite worm was raised at can contribute to differences in its progeny development. Maternal effect based on temperature of rearing can possibility influence progeny development into mid-embryogenesis (Isnenghi et al., 1983).

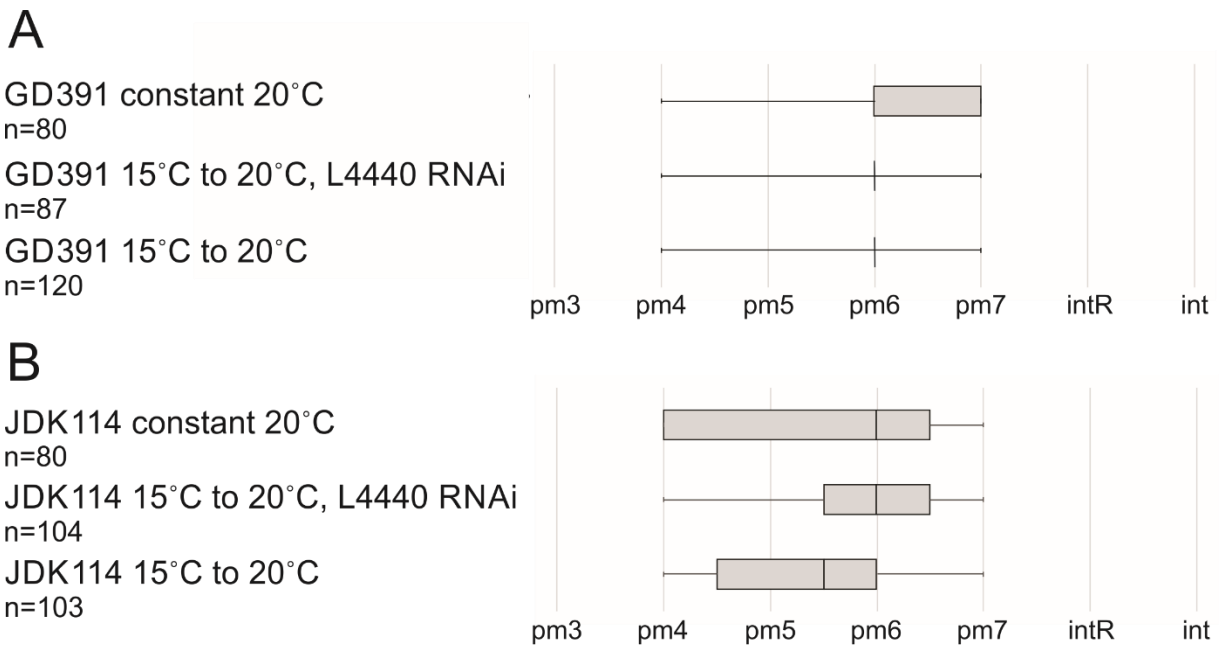

Figure 4.1 Temperature may contribute to g1P position variability. Strains were subjected to three different incubation conditions, either being maintained at a constant 20˚C; grown at 15˚C then injected with control L4440 dsRNA and then kept at 20˚C; or grown at 15˚C then switched to 20˚C without injection. This was completed for (A) GD391 *cam-1(gm122)* II*; ivIs17* V and (B) JDK114 *ham-1(kor15)* IV*; ivIs17* V

5 Conclusion and Future Directions

5.1 Summary

The g1P gland cell uses retrograde extension to migrate through the pharynx during embryonic development. The molecular genetic pathways controlling this movement are not well understood. *cam-1* is a non-canonical Wnt receptor that is required for proper g1P migration. This study used an unbiased forward screen approach to find other candidate genes that function in retrograde extension. Linkage mapping, complementation crosses, and whole genome sequencing identified three novel factors for retrograde extension: *ham-1,* a STOX1-like protein, a novel factor on LGI likely to be *srt-61*, and a novel factor on LGX likely to be *plpp-1.3*. When these genes are mutated, the g1P cell stops prematurely in an anteriorly position. *ham-1* appears to also affect organ-wide morphology. Although they confer a similar phenotype, RNAi suggests that *ham-1* functions in a different pathway than *cam-1*.

5.2 Proposed model

While *cam-1* and *ham-1* mutation both affect pharyngeal organ asymmetry and cause an under migration of the g1P gland cell, they appear to work in separate genetic pathways. CAM-1 has been shown to express in the embryo anterior and overlaps with an anterior portion of the pharyngeal primordium just before g1P begins migrating (Kormish et al., *in prep*). HAM-1 has been shown to segregate to the posterior of asymmetrically dividing cells (Frank et al., 2005; Guenther and Garriga, 1996; Leung et al., 2016), and influences the spindle positioning and cleavage planes of these divisions (Frank et al., 2005, Feng et al., 2013,). Taking these ideas together, I think that *cam-1* influences proper morphogenesis of the anterior part of the pharynx, and that *ham-1* influences the posterior part of the pharynx.

How each gene product acts at their respective end of the pharynx is currently unknown. I imagine *cam-1* is polarizing the anterior and possibly providing an initial permissive cue for the g1P cell to begin moving. The gland cell is born in an anteriorly position and most likely in contact with CAM-1 (Kormish et al., *in prep*). *ham-1* may be influencing the fate of posterior cells or polarizing the primordium at a stage prior to retrograde extension. When *ham-1* function is removed, some cells of the mid to posterior pharynx may take on improper fates or orientation. This could remove environmental permissive cues or alter the cellular scaffold the g1P cell adheres to. Either of these events could allow the g1P cell to "pause" during migration and miss the rest of its limited opportunity for movement. Alternatively, disorganized pharyngeal cells could physically block the g1P from pushing its cell body through the organ. These scenarios could be the cause of the g1P under migration in the *ham-1* background.

5.3 Future Directions

5.3.1 *srt-61* and *plpp-1.3* RNA interference

To confirm the identity of *kor42* I and *kor44* X, knockdown of candidate genes *srt-61* and *plpp-1.3* will be performed. RNAi will be performed in the same way *ham-1* and *cam-1* RNAi were (see section 2.4). RNAi phenotypes can be compared to *kor42* and *kor44* strains to confirm phenocopy.

5.3.2 Examining CAM-1 and HAM-1 distributions in the opposite genetic backgrounds

While some studies show CAM-1 and HAM-1 protein localization in a wildtype embryo, it is important to understand how these distributions would change in a mutant background. Examining CAM-1 protein distribution in a *ham-1* background and HAM-1 in a *cam-1* background, and comparing them to wild-type expression could help indicate if the two gene

products have any spatial overlap in the developing embryo. Overlap or a change in protein expression pattern could provide new evidence that the two genes share some developmental pathway components, and/or could converge on a similar pathway or process. This could be completed through either immunostaining with the appropriate antibodies, or creating a fusion of the desired protein with GFP and introducing it into the appropriate genetic background (Berkowitz et al., 2008; Mello and Fire, 1995).

5.3.3 Examining pharyngeal polarity in *cam-1* and *ham-1* backgrounds

Following the divisions and movements of all pharyngeal primordium cells would be very laborious and possibly inconclusive due to the high degree of overlap and fairly rapid changes in cell shape at the cyst stage (Rasmussen et al., 2012). An alternative way to discover which pharyngeal cells are affected by *cam-1* and *ham-1* would be to examine organ polarity once embryonic development is complete. This can be accomplished by crossing the *jcIs1* [*ajm-1*::GFP + *unc-29*(+) + *rol-6(su1006)*] transgene into *cam-1* and *ham-1* strains. AJM-1 is a coiled-coil protein that localizes to and regulates the integrity of the worm epithelial apical junctions (Köppen et al., 2001). While AJM-1 is not essential for polarity itself, tagging it with a fluorescent protein allows one to examine the pharyngeal lumen and therefore the organization of the surrounding basal cells. By comparing the expression patterns to that of wildtype, it can be determined which sections of the pharynx are not organizing properly. This can also be paired with a marker such as the nuclear-expressing version of *ceh-22*::GFP, which highlights pharyngeal muscle nuclei. This would allow for better cellular orientation and to pinpoint individual cells with morphogenetic or polarity defects. From observed results of this experiment, other specific reporter genes can then be selected to follow specific pharyngeal cells' developmental events.

5.3.4 Further testing *cam-1* kinase domain function in pharyngeal development

While previous work has shown that the *cam-1(ks52)* allele causes a g1P under migration, it is unclear whether this is due to the missing tyrosine kinase domain, or if the hypomorphic knockdown of receptor activity has removed *cam-1* function in general (Kormish et al., *in prep*). To further investigate if the kinase domain activity is required in pharyngeal development, other *cam-1* backgrounds can have the *phat-1*::YFP transgene crossed into them to observe gland cell position. For example, *cam-1(cw82)* and *cam-1(xd13)* change conserved glycine residues of the kinase domain to charged amino acids arginine and glutamic acid, respectively (Chien et al., 2015, Song et al., 2010). Western blots of these genetic backgrounds detect CAM-1 at a level much closer to wildtype expression than the *ks52* background (Chien et al., 2015).

There are currently no published alleles that affect the *cam-1* serine-threonine kinase domain. To see if this domain is required for pharyngeal development, an experiment similar to Kim and Forrester's (2003) could be repeated in the GD391 *cam-1(gm122)* II*; ivIs17* V strain to observe how rescuing constructs affect gland cell placement. The construct containing a deletion of the *cam-1* serine-threonine domain could be used, but given the debatable effect of the *ks52* deletion, it would be wise to identify key activity residues or motifs and create point mutations within the ser/thr kinase as well.

5.3.5 Identifying other genes isolated from the forward genetic screen

This study confirmed the identity of causative mutations in a subset of strains that were obtained from the forward genetic screen. There are 12 other isolated strains with homozygous viable under migrations that have not been characterized. Mapping and sequencing these strains could bring forward more novel finds like *ham-1* or finally identify a factor that works together

with *cam-1*. To better map the strains that are linked to LGV, a gland cell reporter gene located on a chromosome other than V could be crossed into the strain. For the strains that were not mapped due to unexpected phenotypic ratios or linkage to multiple chromosomes after outcrossing with the Hawaiian strain, careful outcrossing with N2 wildtype may be required to remove potentially interfering background mutations and possibly improve phenotype penetrance.

6 References

Adzhubei, A.A., Sternberg, M.J.E., Makarov, A.A., 2013. Polyproline-II Helix in Proteins: Structure and Function. J. Mol. Biol. 425, 2100–2132. https://doi.org/10.1016/j.jmb.2013.03.018

Albertson, D.G., Thomson, J.N., 1976. The pharynx of Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 275, 299–325. https://doi.org/10.1098/rstb.1976.0085

Altun, Z.F., Hall, D.H., 2009. Alimentary system, Pharynx [WWW Document]. WormAtlas. https://doi.org/10.3908/wormatlas.1.3

Altun, Z.F., Hall, D.H., 2006. Introduction to C. elegans anatomy [WWW Document]. WormAtlas. URL http://www.wormatlas.org

Asikainen, S., Vartiainen, S., Lakso, M., Nass, R., Wong, G., 2005. Selective sensitivity of Caenorhabditis elegans neurons to RNA interference. NeuroReport 16, 1995–1999. https://doi.org/10.1097/00001756-200512190-00005

Bainbridge, T., DeAlmeida, V., Izrael-Tomasevic, A., Chalouni, C., Pan, B., Goldsmith, J., Schoen, A., Quinone, G., Kelly, R., Lill, J., Sandoval, W., Costa, M., Polakis, P., Arnott, D., Rubinfeld, B., Ernst, J., 2014. Evolutionary Divergence in the Catalytic Activity of the CAM-1, ROR1 and ROR2 Kinase Domains. PLoS One 9, e102695. https://doi.org/10.1371/journal.pone.0102695

Bakker, E.R.M., Raghoebir, L., Franken, P.F., Helvensteijn, W., van Gurp, L., Meijlink, F., van der Valk, M.A., Rottier, R.J., Kuipers, E.J., van Veelen, W., Smits, R., 2012. Induced Wnt5a expression perturbs embryonic outgrowth and intestinal elongation, but is well-tolerated in adult mice. Dev. Biol. 369, 91–100. https://doi.org/10.1016/j.ydbio.2012.06.007

Bansil, R., Turner, B.S., 2006. Mucin structure, aggregation, physiological functions and biomedical applications. Curr. Opin. Colloid Interface Sci. 11, 164–170. https://doi.org/10.1016/j.cocis.2005.11.001

Baumeister, R., Liu, Y., Ruvkun, G., 1996. Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. Genes Dev. 10, 1395–1410. https://doi.org/10.1101/gad.10.11.1395

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M.J., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M.D., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara E. Catenazzi, M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D.,

Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D.V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.-D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P., Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E., Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, Andrew C., Pike, Alger C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, John, Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., VandeVondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Rogers, Jane, Mullikin, J.C., Hurles, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R., Smith, A.J., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456, 53–59. https://doi.org/10.1038/nature07517

Berkowitz, L.A., Knight, A.L., Caldwell, G.A., Caldwell, K.A., 2008. Generation of stable transgenic C. elegans using microinjection. J. Vis. Exp. JoVE. https://doi.org/10.3791/833

Blumenthal, T., Gleason, K.S., 2003. Caenorhabditis elegans operons: form and function. Nat. Rev. Genet. 4, 112–20. https://doi.org/10.1038/nrg995

Brenner, S., 1974. The Genetics of Caenorhabditis elegans. Genetics 77, 71–94.

Byerly, L., Cassada, R.C., Russell, R.L., 1976. The life cycle of the nematode Caenorhabditis elegans. Dev. Biol. 51, 23–33. https://doi.org/10.1016/0012-1606(76)90119-6

Calixto, A., Chelur, D., Topalidou, I., Chen, X., Chalfie, M., 2010. Enhanced neuronal RNAi in C. elegans using SID-1. Nat. Methods 7, 554–559. https://doi.org/10.1038/nmeth.1463

Cervantes, S., Yamaguchi, T.P., Hebrok, M., 2009. Wnt5a is essential for intestinal elongation in mice. Dev. Biol. 326, 285–94. https://doi.org/10.1016/j.ydbio.2008.11.020

Chien, S.-C., Brinkmann, E.-M., Teuliere, J., Garriga, G., 2013. Caenorhabditis elegans PIG-1/MELK acts in a conserved PAR-4/LKB1 polarity pathway to promote asymmetric neuroblast divisions. Genetics 193, 897–909. https://doi.org/10.1534/genetics.112.148106

Chien, S.J., Gurling, M., Kim, C., Craft, T., Forrester, W., Garriga, G., 2015. Autonomous and nonautonomous regulation of Wnt-mediated neuronal polarity by the C . elegans Ror kinase CAM-1. Dev. Biol. 404, 55–65. https://doi.org/10.1016/j.ydbio.2015.04.015

Consortium*, T.C. elegans S., 1998. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 282, 2012–2018. https://doi.org/10.1126/science.282.5396.2012

Cordes, S., Frank, C.A., Garriga, G., 2006. The C. elegans MELK ortholog PIG-1 regulates cell size asymmetry and daughter cell fate in asymmetric neuroblast divisions. Dev. Camb. Engl. 133, 2747–56. https://doi.org/10.1242/dev.02447

Costa, M., Raich, W., Agbunag, C., Leung, B., Hardin, J., Priess, J.R., 1998. A Putative Catenin–Cadherin System Mediates Morphogenesis of the *Caenorhabditis elegans* Embryo. J. Cell Biol. 141, 297–308. https://doi.org/10.1083/jcb.141.1.297

Crouzier, T., Boettcher, K., Geonnotti, A.R., Kavanaugh, N.L., Hirsch, J.B., Ribbeck, K., Lieleg, O., 2015. Modulating Mucin Hydration and Lubrication by Deglycosylation and Polyethylene Glycol Binding. Adv. Mater. Interfaces 2, 1500308. https://doi.org/10.1002/admi.201500308

Dann, C.E., Hsieh, J.C., Rattner, a, Sharma, D., Nathans, J., Leahy, D.J., 2001. Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. Nature 412, 86–90. https://doi.org/10.1038/35083601

Davis, M.W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S., Jorgensen, E.M., 2005. Rapid single nucleotide polymorphism mapping in C. elegans. BMC Genomics 6, 118. https://doi.org/10.1186/1471-2164-6-118

DeChiara, T.M., Bowen, D.C., Valenzuela, D.M., Simmons, M. V, Poueymirou, W.T., Thomas, S., Kinetz, E., Compton, D.L., Rojas, E., Park, J.S., Smith, C., DiStefano, P.S., Glass, D.J., Burden, S.J., Yancopoulos, G.D., 1996. The receptor tyrosine kinase MuSK is required for neuromuscular junction formation in vivo. Cell 85, 501–12.

Desai, C., Garriga, G., McIntire, S.L., Horvitz, H.R., 1988. A genetic pathway for the development of the C. elegans HSN motor neurons. Nature 336, 638–646.

Devenport, D., 2014. The cell biology of planar cell polarity. J. Cell Biol. 207, 171–9. https://doi.org/10.1083/jcb.201408039

Dixon, S.J., Roy, P. J., 2005. Muscle arm development in Caenorhabditis elegans. Development 132, 3079–3092. https://doi.org/10.1242/dev.01883

Doitsidou, M., Poole, R.J., Sarin, S., Bigelow, H., Hobert, O., 2010. C. elegans mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. PloS One 5, e15435. https://doi.org/10.1371/journal.pone.0015435

Emmons, S.W., 2005. Male development. WormBook 1–19. https://doi.org/10.1895/wormbook.1.33.1

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–94.

Feng, G., Yi, P., Yang, Y., Chai, Y., Tian, D., Zhu, Z., Liu, J., Zhou, F., Cheng, Z., Wang, X., Li, W., Ou, G., 2013. Developmental stage-dependent transcriptional regulatory pathways control neuroblast lineage progression 3847, 3838–3847. https://doi.org/10.1242/dev.098723

Fielenbach, N., Antebi, A., 2008. C. elegans dauer formation and the molecular basis of plasticity. Genes Dev. 22, 2149–65. https://doi.org/10.1101/gad.1701508

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Flibotte, S., Edgley, M.L., Chaudhry, I., Taylor, J., Neil, S.E., Rogula, A., Zapf, R., Hirst, M., Butterfield, Y., Jones, S.J., Marra, M.A., Barstead, R.J., Moerman, D.G., 2010. Whole-Genome Profiling of Mutagenesis in Caenorhabditis elegans. Genetics 185, 431–441. https://doi.org/10.1534/genetics.110.116616

Forrester, W.C., 2002. The Ror receptor tyrosine kinase family. Cell. Mol. Life Sci. 59, 83–96.

Forrester, W.C., Dell, M., Perens, E., Garriga, G., 1999. A C. elegans Ror receptor tyrosine kinase regulates cell motility and asymmetric cell division. Nature 400, 717–721.

Forrester, W.C., Kim, C., Garriga, G., 2004. The Caenorhabditis elegans Ror RTK CAM-1 Inhibits EGL-20/Wnt Signaling in Cell Migration. Genetics 1962, 1951–1962. https://doi.org/10.1534/genetics.104.031781

Francis, M.M., Evans, S.P., Jensen, M., Madsen, D.M., Mancuso, J., Norman, K.R., Maricq, A.V., 2005. The Ror Receptor Tyrosine Kinase CAM-1 Is Required for ACR-16-Mediated Synaptic Transmission at the C . elegans Neuromuscular Junction. Neuron 46, 581–594. https://doi.org/10.1016/j.neuron.2005.04.010

Frank, C.A., Hawkins, N.C., Guenther, C., Horvitz, H.R., Garriga, G., 2005. C. elegans HAM-1 positions the cleavage plane and regulates apoptosis in asymmetric neuroblast divisions. Dev. Biol. 284, 301–310. https://doi.org/10.1016/j.ydbio.2005.05.026

Friedl, P., Alexander, S., 2012. Review Cancer Invasion and the Microenvironment : Plasticity and Reciprocity. Cell 147, 992–1009. https://doi.org/10.1016/j.cell.2011.11.016

Friedman, J.R., Kaestner, K.H., 2006. The Foxa family of transcription factors in development and metabolism. Cell. Mol. Life Sci. 63, 2317–2328. https://doi.org/10.1007/s00018-006-6095-6

Gally, C., Wissler, F., Zahreddine, H., Quintin, S., Landmann, F., Labouesse, M., 2009. Myosin II regulation during C. elegans embryonic elongation: LET-502/ROCK, MRCK-1 and PAK-1, three kinases with different roles. Dev. Camb. Engl. 136, 3109–19. https://doi.org/10.1242/dev.039412

Gao, B., Song, H., Bishop, K., Elliot, G., Garrett, L., English, M.A., Andre, P., Robinson, J., Sood, R., Minami, Y., Economides, A.N., Yang, Y., 2012. Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2. Dev. Cell 20, 163–76. https://doi.org/10.1016/j.devcel.2011.01.001

Ghai, V., Gaudet, J., 2008. The CSL transcription factor LAG-1 directly represses hlh-6 expression in C . elegans. Dev. Biol. 322, 334–344. https://doi.org/10.1016/j.ydbio.2008.07.018

Ghai, V., Smit, R.B., Gaudet, J., 2012. Transcriptional regulation of HLH-6-independent and subtype-specific genes expressed in the Caenorhabditis elegans pharyngeal glands. Mech. Dev. 129, 284–297. https://doi.org/10.1016/j.mod.2012.06.005

Green, J.L., Inoue, T., Sternberg, P.W., 2008a. Opposing Wnt Pathways Orient Cell Polarity during Organogenesis. Cell 134, 646–656. https://doi.org/10.1016/j.cell.2008.06.026

Green, J.L., Inoue, T., Sternberg, P.W., 2007. The C. elegans ROR receptor tyrosine kinase, CAM-1, non-autonomously inhibits the Wnt pathway. Dev. Camb. Engl. 134, 4053–4062. https://doi.org/10.1242/dev.005363

Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008b. Ror receptor tyrosine kinases : orphans no more. Trends Cell Biol. 1, 536–544. https://doi.org/10.1016/j.tcb.2008.08.006

Guenther, C., Garriga, G., 1996. Asymmetric distribution of the C. elegans HAM-1 protein in neuroblasts enables daughter cells to adopt distinct fates. Dev. Camb. Engl. 3518, 3509–3518.

Hall, D.H., Herndon, L.A., Altun, Z.F., 2009. Introduction to C. elegans embryo anatomy [WWW Document]. WormAtlas. https://doi.org/doi:10.3908/wormatlas.1.1

Harrison, M.M., Jenkins, B. V, O'Connor-Giles, K.M., Wildonger, J., 2014. A CRISPR view of development. Genes Dev. 28, 1859–72. https://doi.org/10.1101/gad.248252.114

Haun, C., Alexander, J., Stainier, D.Y., Okkema, P.G., 1998. Rescue of Caenorhabditis elegans pharyngeal development by a vertebrate heart specification gene. Proc. Natl. Acad. Sci. U. S. A. 95, 5072–5.

Hayashi, Y., Hirotsu, T., Iwata, R., Kage-nakadai, E., Kunitomo, H., Ishihara, T., Iino, Y., Kubo, T., Early, L., 2009. A trophic role for Wnt-Ror kinase signaling during developmental

pruning in Caenorhabditis elegans. Nat. Neurosci. 12, 981–987. https://doi.org/10.1038/nn.2347
He, F., Xiong, W., Yu, X., Espinoza-Lewis, R., Liu, C., Gu, S., Nishita, M., Suzuki, K., Yamada, G., Minami, Y., Chen, Y., 2008. Wnt5a regulates directional cell migration and cell proliferation via Ror2-mediated noncanonical pathway in mammalian palate development. Dev. Camb. Engl. 135, 3871–9. https://doi.org/10.1242/dev.025767
Heiman, M.G., Shaham, S., 2009. DEX-1 and DYF-7 Establish Sensory Dendrite Length by Anchoring Dendritic Tips during Cell Migration. Cell 137, 344–355. https://doi.org/10.1016/j.cell.2009.01.057
Hodgkin, J., Doniach, T., 1997. Natural variation and copulatory plug formation in Caenorhabditis elegans. Genetics 146, 149–64.
Hodgkin, J., Horvitz, H.R., Brenner, S., 1979. Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. Genetics 91, 67–94.
Horner, M. a., Quintin, S., Domeier, M.E., Kimble, J., Labouesse, M., Mango, S.E., 1998. pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in Caenorhabditis elegans. Genes Dev. 12, 1947–1952. https://doi.org/10.1101/gad.12.13.1947
Huang, L., Sternberg, Paul, 2006. Genetic dissection of developmental pathways. WormBook. https://doi.org/10.1895/wormbook.1.88.2
Hubbard, S.R., 1999. Structural analysis of receptor tyrosine kinases. Prog. Biophys. Mol. Biol. 71, 343–358. https://doi.org/10.1016/S0079-6107(98)00047-9
Isnenghi, E., Cassada, R., Smith, K., Denich, K., Radnia, K., von Ehrenstein, G., 1983. Maternal effects and temperature-sensitive period of mutations affecting embryogenesis in Caenorhabditis elegans. Dev. Biol. 98, 465–480. https://doi.org/10.1016/0012-1606(83)90376-7
Jensen, M., Hoerndli, F.J., Brockie, P.J., Wang, R., Johnson, E., Maxfield, D., Francis, M.M., Madsen, D.M., Maricq, A. V, 2012. Wnt Signaling Regulates Acetylcholine Receptor Translocation and Synaptic Plasticity in the Adult Nervous System. Cell 149, 173–187. https://doi.org/10.1016/j.cell.2011.12.038
Jiang, P., Zhang, D., 2013. Maternal Embryonic Leucine Zipper Kinase (MELK): A Novel Regulator in Cell Cycle Control, Embryonic Development, and Cancer. Int. J. Mol. Sci. 14, 21551–21560. https://doi.org/10.3390/ijms141121551
Jones, S., 2004. An overview of the basic helix-loop-helix proteins. Genome Biol. 5, 226. https://doi.org/10.1186/gb-2004-5-6-226
Kalb, J.M., Lau, K.K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P.G., McGhee, J.D., 1998. pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the C. elegans pharynx. Dev. Camb. Engl. 125, 2171–2180.
Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., Ahringer, J., 2003. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421, 231–7. https://doi.org/10.1038/nature01278
Kay, B.K., Williamson, M.P., Sudol, M., 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J. 14, 231–241. https://doi.org/10.1096/fasebj.14.2.231
Kelly, W.G., Xu, S., Montgomery, M.K., Fire, A., 1997. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. Genetics 146, 227–38.

Kennedy, S., Wang, D., Ruvkun, G., 2004. A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature 427, 645–649. https://doi.org/10.1038/nature02302

Kennerdell, J.R., Fetter, R.D., Bargmann, C.I., 2009. Wnt-Ror signaling to SIA and SIB neurons directs anterior axon guidance and nerve ring placement in C . elegans. Development 3810, 3801–3810. https://doi.org/10.1242/dev.038109

Kestler, H.A., Kühl, M., 2008. From individual Wnt pathways towards a Wnt signalling network. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 363, 1333–47. https://doi.org/10.1098/rstb.2007.2251

Kim, C., Forrester, W.C., 2003a. Functional analysis of the domains of the C . elegans Ror receptor tyrosine kinase CAM-1. Dev. Biol. 264, 376–390. https://doi.org/10.1016/j.ydbio.2003.09.007

Kim, C., Forrester, W.C., 2003b. Functional analysis of the domains of the C . elegans Ror receptor tyrosine kinase CAM-1. Dev. Biol. 264, 376–390. https://doi.org/10.1016/j.ydbio.2003.09.007

Kishore, R.S., Sundaram, M.V., 2002. ced-10 Rac and mig-2 Function Redundantly and Act with unc-73 Trio to Control the Orientation of Vulval Cell Divisions and Migrations in Caenorhabditis elegans. Dev. Biol. 241, 339–348. https://doi.org/10.1006/DBIO.2001.0513

Klezovitch, O., Scanu, A.M., 1996. Lys and Fibrinogen Binding of Wild-Type (Trp72) and Mutant (Arg72) Human Apo(a) Kringle IV-10 Expressed in E coli and CHO Cells. Arterioscler. Thromb. Vasc. Biol. 16, 392–398. https://doi.org/10.1161/01.ATV.16.3.392

Koch, R., van Luenen, H.G., van der Horst, M., Thijssen, K.L., Plasterk, R.H., 2000. Single nucleotide polymorphisms in wild isolates of Caenorhabditis elegans. Genome Res. 10, 1690–6.

Koga, M., Take-uchi, M., Tameishi, T., Ohshima, Y., 1999. Control of DAF-7 TGF-(alpha) expression and neuronal process development by a receptor tyrosine kinase KIN-8 in Caenorhabditis elegans. Dev. Camb. Engl. 126, 5387–98.

Komiya, Y., Habas, R., 2008. Wnt signal transduction pathways. Organogenesis 4, 68–75. https://doi.org/10.4161/org.4.2.5851

Komuro, H., Yacubova, E., 2003. Recent advances in cerebellar granule cell migration. Cell. Mol. Life Sci. 60, 1084–1098. https://doi.org/10.1007/s00018-003-2248-z

Koppen, M., Simske, J.S., Sims, P.A., Firestein, B.L., Hall, D.H., Radice, A.D., Rongo, C., Hardin, J.D., 2001. Cooperative regulation of AJM-1 controls junctional integrity in Caenorhabditis elegans epithelia. Nat. Cell Biol. 3, 983–91. https://doi.org/10.1038/ncb1101-983

Kormish, J.D., McGhee, J.D., 2005. The C. elegans lethal gut-obstructed gob-1 gene is trehalose-6-phosphate phosphatase. Dev. Biol. 287, 35–47. https://doi.org/10.1016/J.YDBIO.2005.08.027

Kormish, J., Raharjo, W., Kim, S., Rohs, P., Srayko, M., Gaudet, J. *in prep.* cam-1/ROR and egl-15/FGFR receptor tyrosine kinases regulate retrograde extension and migration of the gland cells during pharynx morphogenesis in C. elegans.

Kühl, M., Sheldahl, L.C., Park, M., Miller, J.R., Moon, R.T., 2000. The Wnt/Ca2+ pathway: a new vertebrate Wnt signaling pathway takes shape. Trends Genet. TIG 16, 279–83.

Kurosaka, S., Kashina, A., 2008. Cell Biology of Embryonic Migration. Birth Defects Res. Part C 84, 102–122. https://doi.org/10.1002/bdrc.20125

Kutscher, L.M., Shaham, S., 2014. Forward and reverse mutagenesis in C. elegans. WormBook. https://doi.org/doi/10.1895/wormbook.1.167.1
Labouesse, M., 2006. Epithelial Junctions and Attachments [WWW Document]. Wormbook. https://doi.org/doi/10.1895/wormbook.1.56.1
Lai, C.-H., Chou, C.-Y., Ch'ang, L.-Y., Liu, C.-S., Lin, W., 2000. Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics. Genome Res. 10, 703–713.
Leung, A., Hua, K., Ramachandran, P., Hingwing, K., Wu, M., Luan, P., Hawkins, N., 2016. C . elegans HAM-1 functions in the nucleus to regulate asymmetric neuroblast division. Dev. Biol. 410, 56–69. https://doi.org/10.1016/j.ydbio.2015.12.011
Levitan, D., Doyle, T.G., Brousseau, D., Lee, M.K., Thinakaran, G., Slunt, H.H., Sisodia, S.S., Greenwald, I., 1996. Assessment of normal and mutant human presenilin function in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 93, 14940–14944. https://doi.org/10.1073/pnas.93.25.14940
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. 25, 2078–9. https://doi.org/10.1093/bioinformatics/btp352
Li, L., Yuan, H., Xie, W., Mao, J., Caruso, A.M., McMahon, A., Sussman, D.J., Wu, D., 1999. Dishevelled proteins lead to two signaling pathways. Regulation of LEF-1 and c-Jun N-terminal kinase in mammalian cells. J. Biol. Chem. 274, 129–34.
Li, W., Koutmou, K.S., Leahy, D.J., Li, M., 2015. Systemic RNA Interference Deficiency-1 (SID-1) Extracellular Domain Selectively Binds Long Double-stranded RNA and Is Required for RNA Transport by SID-1. J. Biol. Chem. 290, 18904–18913. https://doi.org/10.1074/jbc.M115.658864
Liu, Y., Bodine, P.V.., Billiard, J., 2007. Ror2, a novel modulator of osteogenesis. J Musculoskelet Neuronal Interact 7, 323–324.
Liu, Yan, Ross, J.F., Bodine, P.V.N., Billiard, J., 2007. Homodimerization of Ror2 tyrosine kinase receptor induces 14-3-3(beta) phosphorylation and promotes osteoblast differentiation and bone formation. Mol. Endocrinol. Baltim. Md 21, 3050–61. https://doi.org/10.1210/me.2007-0323
Lowery, L.A., Vactor, D. Van, 2009. The trip of the tip : understanding the growth cone machinery 10. https://doi.org/10.1038/nrm2679
Maddox, P.S., Oegema, K., Desai, A., Cheeseman, I.M., 2004. "Holo"er than thou: Chromosome segregation and kinetochore function in C. elegans. Chromosome Res. 12, 641–653. https://doi.org/10.1023/B:CHRO.0000036588.42225.2f
Mak, M., Spill, F., Kamm, R.D., Zaman, M.H., 2016. Single-Cell Migration in Complex Microenvironments: Mechanics and Signaling Dynamics. J. Biomech. Eng. 138, 021004. https://doi.org/10.1115/1.4032188
Mango, S.E., Lambie, E.J., Kimble, J., 1994. The pha-4 gene is required to generate the pharyngeal primordium of Caenorhabditis elegans. Dev. Camb. Engl. 120, 3019–3031.
Mangone, M., Manoharan, A.P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S.D., Mis, E., Zegar, C., Gutwein, M.R., Khivansara, V., Attie, O., Chen, K., Salehi-Ashtiani, K., Vidal, M., Harkins, T.T., Bouffard, P., Suzuki, Y., Sugano, S., Kohara, Y., Rajewsky,

N., Piano, F., Gunsalus, K.C., Kim, J.K., 2010. The Landscape of C. elegans 3'UTRs. Science 329, 432–435. https://doi.org/10.1126/science.1191244
Matsuda, T., Nomi, M., Ikeya, M., Kani, S., Oishi, I., Terashima, T., Takada, S., Minami, Y., 2001. Expression of the receptor tyrosine kinase genes, Ror1 and Ror2, during mouse development. Mech. Dev. 105, 153–156.
McBride, A.E., Silver, P.A., 2001. State of the Arg: Protein Methylation at Arginine Comes of Age. Cell 106, 5–8. https://doi.org/10.1016/S0092-8674(01)00423-8
McCombie, W.R., Adams, M.D., Kelley, J.M., FitzGerald, M.G., Utterback, T.R., Khan, M., Dubnick, M., Kerlavage, A.R., Craig Venter, J., Fields, C., 1992. Caenorhabditis elegans expressed sequence tags identify gene families and potential disease gene homologues. Nat. Genet. 1, 124–131. https://doi.org/10.1038/ng0592-124
McKay, S.E., Hislop, J., Scott, D., Bulloch, A.G., Kaczmarek, L.K., Carew, T.J., Sossin, W.S., 2001. Aplysia ror forms clusters on the surface of identified neuroendocrine cells. Mol. Cell. Neurosci. 17, 821–41. https://doi.org/10.1006/mcne.2001.0977
McKim, K.S., Howell, A.M., Rose, A.M., 1988. The effects of translocations on recombination frequency in Caenorhabditis elegans. Genetics 120, 987–1001.
Mello, C.C., Fire, A., 1995. DNA Transformation. Methods Cell Biol. 48, 451–482.
Mello, C.C., Kramer, J.M., Stinchcomb, D., Ambros, V., 1991. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10, 3959–70.
Minami, Y., Oishi, I., Endo, M., Nishita, M., 2010. Ror-family receptor tyrosine kinases in noncanonical Wnt signaling: their implications in developmental morphogenesis and human diseases. Dev. Dyn. Off. Publ. Am. Assoc. Anat. 239, 1–15. https://doi.org/10.1002/dvdy.21991
Minevich, G., Park, D.S., Blankenberg, D., Poole, R.J., Hobert, O., 2012. CloudMap: a cloud-based pipeline for analysis of mutant genome sequences. Genetics 192, 1249–69. https://doi.org/10.1534/genetics.112.144204
Modzelewska, K., Lauritzen, A., Hasenoeder, S., Brown, L., Georgiou, J., 2013a. Neurons Refine the Caenorhabditis elegans Body Plan by Directing Axial Patterning by Wnts 11, 1–22. https://doi.org/10.1371/journal.pbio.1001465
Mojica, F.J.M., Diez-Villasenor, C., Garcia-Martinez, J., Soria, E., 2005. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J. Mol. Evol. 60, 174–182. https://doi.org/10.1007/s00239-004-0046-3
Nakano, A., Takashima, S., 2012. LKB1 and AMP-activated protein kinase: regulators of cell polarity. Genes Cells Devoted Mol. Cell. Mech. 17, 737–47. https://doi.org/10.1111/j.1365-2443.2012.01629.x
Nie, X., Zhang, K., Wang, L., Ou, G., Zhu, H., Gao, W.-Q., 2015. Transcription factor STOX1 regulates proliferation of inner ear epithelial cells *via* the AKT pathway. Cell Prolif. 48, 209–220. https://doi.org/10.1111/cpr.12174
Nigon, V., 1949. Les modalités de la reproduction et le déterminisme du sexe chez quelques Nematodes libres. Ann. Sci. Nat. Zool. Biol. Anim. 11, 1–132.
Nomi, M., Oishi, I., Kani, S., Suzuki, H., Matsuda, T., Yoda, A., Kitamura, M., Itoh, K., Takeuchi, S., Takeda, K., Akira, S., Ikeya, M., Takada, S., Minami, Y., 2001. Loss of mRor1 enhances the heart and skeletal abnormalities in mRor2-deficient mice: redundant and pleiotropic functions of mRor1 and mRor2 receptor tyrosine kinases. Mol. Cell. Biol. 21, 8329–35. https://doi.org/10.1128/MCB.21.24.8329-8335.2001

Oishi, I., Suzuki, H., Onishi, N., Takada, R., Kani, S., Ohkawara, B., Koshida, I., Suzuki, K., Yamada, G., Schwabe, G.C., Mundlos, S., Shibuya, H., Takada, S., Minami, Y., 2003. The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway. Genes Cells 8, 645–654. https://doi.org/10.1046/j.1365-2443.2003.00662.x

Okkema, P.G., Fire, A., 1994. The Caenorhabditis elegans NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. Development 2186, 2175–2186.

Okkema, P.G., Ha, E., Haun, C., Chen, W., Fire, A., 1997. The Caenorhabditis elegans NK-2 homeobox gene ceh-22 activates pharyngeal muscle gene expresion in combination with pha-1 and is required for normal pharyngeal development. Development 124, 3965–3973.

Pacquelet, A., Uhart, P., Tassan, J.-P., Michaux, G., 2015. PAR-4 and anillin regulate myosin to coordinate spindle and furrow position during asymmetric division. J. Cell Biol. 210, 1085–99. https://doi.org/10.1083/jcb.201503006

Pandey, P., Bhardwaj, A., Babu, K., 2017. Regulation of WNT Signaling at the Neuromuscular Junction by the Immunoglobulin Superfamily Protein RIG-3 in Caenorhabditis elegans. Genetics 206, 1521–1534. https://doi.org/10.1534/genetics.116.195297

Peggs, K.S., Allison, J.P., 2005. Co-stimulatory pathways in lymphocyte regulation : the immunoglobulin superfamily 809–824. https://doi.org/10.1111/j.1365-2141.2005.05627.x

Piano, F., Schetter, A.J., Morton, D.G., Gunsalus, K.C., Reinke, V., Kim, S.K., Kemphues, K.J., 2002. Gene clustering based on RNAi phenotypes of ovary-enriched genes in C. elegans. Curr. Biol. CB 12, 1959–64. https://doi.org/10.1016/S0960-9822(02)01301-5

Piekny, A.J., Johnson, J.-L.F., Cham, G.D., Mains, P.E., 2003. The Caenorhabditis elegans nonmuscle myosin genes nmy-1 and nmy-2 function as redundant components of the let-502/Rho-binding kinase and mel-11/myosin phosphatase pathway during embryonic morphogenesis. Development 130, 5695–5704. https://doi.org/10.1242/dev.00807

Podbilewicz, B., Gruenbaum, Y., 2006. Live Imaging of Caenorhabditis elegans: preparation of samples. CSH Protoc. 2006, pdb.prot4601. https://doi.org/10.1101/pdb.prot4601

Portereiko, M.F., Mango, S.E., 2001. Early morphogenesis of the Caenorhabditis elegans pharynx. Dev. Biol. 233, 482–94. https://doi.org/10.1006/dbio.2001.0235

Pozzi, N., Chen, Z., Pelc, L.A., Shropshire, D.B., Di Cera, E., 2014. The linker connecting the two kringles plays a key role in prothrombin activation. Proc. Natl. Acad. Sci. U. S. A. 111, 7630–5. https://doi.org/10.1073/pnas.1403779111

Priess, J.R., Hirsh, D.I., 1986. Caenorhabditis elegans morphogenesis: The role of the cytoskeleton in elongation of the embryo. Dev. Biol. 117, 156–173. https://doi.org/10.1016/0012-1606(86)90358-1

Qu, W., Ren, C., Li, Y., Shi, J., Zhang, J., Wang, X., Hang, X., Lu, Y., Zhao, D., Zhang, C., 2011. Reliability analysis of the Ahringer Caenorhabditis elegans RNAi feeding library: a guide for genome-wide screens. BMC Genomics 12, 170. https://doi.org/10.1186/1471-2164-12-170

Raharjo, I., Gaudet, J., 2007. Gland-specific expression of C . elegans hlh-6 requires the combinatorial action of three distinct promoter elements 302, 295–308. https://doi.org/10.1016/j.ydbio.2006.09.036

Raharjo, W.H., Ghai, V., Dineen, A., Bastiani, M., Gaudet, J., 2011. Cell Architecture : Surrounding Muscle Cells Shape Gland Cell Morphology. Genetics 189, 885–897. https://doi.org/10.1534/genetics.111.132449

Ramakrishnan, K., Ray, P., Okkema, P.G., 2014. CEH-28 activates dbl-1 expression and TGF- β signaling in the C . elegans M4 neuron. Dev. Biol. 390, 149–159. https://doi.org/10.1016/j.ydbio.2014.03.015
Rasmussen, J.P., English, K., Tenlen, J.R., Priess, J.R., 2008. Notch Signaling and Morphogenesis of Single-Cell Tubes in the C. elegans Digestive Tract. Dev. Cell 14, 559–569. https://doi.org/10.1016/j.devcel.2008.01.019
Rasmussen, J.P., Feldman, J.L., Reddy, S.S., Priess, J.R., 2013. Cell interactions and patterned intercalations shape and link epithelial tubes in C. elegans. PLoS Genet. 9, e1003772. https://doi.org/10.1371/journal.pgen.1003772
Rasmussen, J.P., Reddy, S.S., Priess, J.R., 2012. Laminin is required to orient epithelial polarity in the C . elegans pharynx. Development 139, 2050–2060. https://doi.org/10.1242/dev.078360
Roy, P.J., Zheng, H., Warren, C.E., Culotti, J.G., 2000. mab-20 encodes Semaphorin-2a and is required to prevent ectopic cell contacts during epidermal morphogenesis in Caenorhabditis elegans. Dev. Camb. Engl. 127, 755–767.
Santella, A., Du, Z., Nowotschin, S., Hadjantonakis, A.-K., Bao, Z., 2010. A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D. BMC Bioinformatics 11, 580. https://doi.org/10.1186/1471-2105-11-580
Schambony, A., Wedlich, D., 2007. Wnt-5A/Ror2 regulate expression of XPAPC through an alternative noncanonical signaling pathway. Dev. Cell 12, 779–92. https://doi.org/10.1016/j.devcel.2007.02.016
Shaye, D.D., Greenwald, I., 2011. OrthoList: A Compendium of C. elegans Genes with Human Orthologs. PLoS ONE 6, e20085. https://doi.org/10.1371/journal.pone.0020085
Singhal, A., Shaham, S., 2017. Infrared laser-induced gene expression for tracking development and function of single C. elegans embryonic neurons. Nat. Commun. 8, 14100. https://doi.org/10.1038/ncomms14100
Singhvi, A., Frank, C.A., Garriga, G., 2008. The T-Box Gene tbx-2, the Homeobox Gene egl-5 and the Asymmetric Cell Division Gene ham-1 Specify Neural Fate in the HSN/PHB Lineage. Genetics 179, 887–898. https://doi.org/10.1534/genetics.108.088948
Smit, R.B., Schnabel, R., Gaudet, J., 2008. The HLH-6 transcription factor regulates C. elegans pharyngeal gland development and function. PLoS Genet. 4. https://doi.org/10.1371/journal.pgen.1000222
Sokalingam, S., Raghunathan, G., Soundrarajan, N., Lee, S.-G., 2012. A Study on the Effect of Surface Lysine to Arginine Mutagenesis on Protein Stability and Structure Using Green Fluorescent Protein. PLoS ONE 7, e40410. https://doi.org/10.1371/journal.pone.0040410
Song, S., Zhang, B., Sun, H., Li, X., Xiang, Y., Liu, Z., Huang, X., 2010. A Wnt-Frz / Ror-Dsh Pathway Regulates Neurite Outgrowth in Caenorhabditis elegans. PLoS Genet. 6. https://doi.org/10.1371/journal.pgen.1001056
Sossin, W.S., 2006. Tracing the Evolution and Function of the Trk Superfamily of Receptor Tyrosine Kinases. Brain. Behav. Evol. 68, 145–156. https://doi.org/10.1159/000094084
Stern, M.J., Marengere, L.E., Daly, R.J., Lowenstein, E.J., Kokel, M., Batzer, A., Olivier, P., Pawson, T., Schlessinger, J., 1993. The human GRB2 and Drosophila Drk genes can functionally replace the Caenorhabditis elegans cell signaling gene sem-5. Mol. Biol. Cell 4, 1175–1188.
Steven, R., Zhang, L., Culotti, J.G., Pawson, T., 2005. The UNC-73/Trio RhoGEF-2 domain is required in separate isoforms for the regulation of pharynx pumping and normal

neurotransmission in C. elegans. Genes Dev. 19, 2016–2029. https://doi.org/10.1101/gad.1319905
Struhl, K., Segal, E., 2013. Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. 20, 267–273. https://doi.org/10.1038/nsmb.2506
Sugi, T., 2016. Genome Editing in C. elegans and Other Nematode Species. Int. J. Mol. Sci. 17, 295. https://doi.org/10.3390/ijms17030295
Sulston, J.E., 1977. Post-embryonic Cell Lineages of the Nematode , Caenorhabditis elegans. Dev. Biol. 156, 110–156.
Sulston, J.E., Schierenberg, E., White, J.G., Thomson, J.N., 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol. 100, 64–119. https://doi.org/10.1016/0012-1606(83)90201-4
Taylor, J., Abramova, N., Charlton, J., Adler, P.N., 1998. Van Gogh: a new Drosophila tissue polarity gene. Genetics 150, 199–210.
The C. elegans Deletion Mutant Consortium, 2012. Large-Scale Screening for Targeted Knockouts in the Caenorhabditis elegans Genome. G3 2, 1415–1425. https://doi.org/10.1534/g3.112.003830
Thompson, O., Edgley, M., Strasbourger, P., Flibotte, S., Ewing, B., Adair, R., Au, V., Chaudhry, I., Fernando, L., Hutter, H., Kieffer, A., Lau, J., Lee, N., Miller, A., Raymant, G., Shen, B., Shendure, J., Taylor, J., Turner, E.H., Hillier, L.W., Moerman, D.G., Waterston, R.H., 2013. The million mutation project: A new approach to genetics in Caenorhabditis elegans. Genome Res. 23, 1749–1762. https://doi.org/10.1101/gr.157651.113
Timmons, L., Court, D.L., Fire, A., 2001. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene 263, 103–12.
Tzur, Y.B., Friedland, A.E., Nadarajan, S., Church, G.M., Calarco, J.A., Colaiácovo, M.P., 2013. Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. Genetics 195, 1181–1185. https://doi.org/10.1534/genetics.113.156075
Van Auken, K., Weaver, D.C., Edgar, L.G., Wood, W.B., 2000. Caenorhabditis elegans embryonic axial patterning requires two recently discovered posterior-group Hox genes. Proc. Natl. Acad. Sci. U. S. A. 97, 4499–503. https://doi.org/10.1073/PNAS.97.9.4499
van Dijk, M., Mulders, J., Poutsma, A., Könst, A.A.M., Lachmeijer, A.M.A., Dekker, G.A., Blankenstein, M.A., Oudejans, C.B.M., 2005. Maternal segregation of the Dutch preeclampsia locus at 10q22 with a new member of the winged helix gene family. Nat. Genet. 37, 514–519. https://doi.org/10.1038/ng1541
van Dijk, M., van Bezu, J., van Abel, D., Dunk, C., Blankenstein, M.A., Oudejans, C.B.M., Lye, S.J., 2010. The STOX1 genotype associated with pre-eclampsia leads to a reduction of trophoblast invasion by α-T-catenin upregulation. Hum. Mol. Genet. 19, 2658–2667. https://doi.org/10.1093/hmg/ddq152
Vandamme, J., Lettier, G., Sidoli, S., Di Schiavi, E., Nørregaard Jensen, O., Salcini, A.E., 2012. The C. elegans H3K27 Demethylase UTX-1 Is Essential for Normal Development, Independent of Its Enzymatic Activity. PLoS Genet. 8, e1002647. https://doi.org/10.1371/journal.pgen.1002647
Veeman, M.T., Axelrod, J.D., Moon, R.T., 2003. A Second Canon : Functions and Mechanisms of -Catenin-Independent Wnt Signaling 5, 367–377.

Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J., Sulston, J., 1992. A survey of expressed genes in Caenorhabditis elegans. Nat. Genet. 1, 114–123. https://doi.org/10.1038/ng0592-114

Wicks, S., Yeh, R., Gish, W., Waterson, R., Plasterk, R., 2001. Rapid gene mapping i n C aenorhabditis elegans using a high density polymorphism map. Nat. Genet. 28, 160–164.

Williamson, M.P., 1994. The structure and function of proline-rich regions in proteins. Biochem. J. 297, 249–260. https://doi.org/10.1042/bj2970249

Wilson, C., Goberdhan, D.C.I., Steller, H., 1993. Dror , a potential neurotrophic receptor gene , encodes a Drosophila homolog of the vertebrate Ror family of Trk-related receptor tyrosine kinases. Proc. Natl. Acad. Sci. U. S. A. 90, 7109–7113.

Witte, F., Chan, D., Economides, A.N., Mundlos, S., Stricker, S., 2010. Receptor tyrosine kinase-like orphan receptor 2 (ROR2) and Indian hedgehog regulate digit outgrowth mediated by the phalanx-forming region. Proc. Natl. Acad. Sci. U. S. A. 107, 14211–6. https://doi.org/10.1073/pnas.1009314107

Wolff, T., Rubin, G.M., 1998. Strabismus, a novel gene that regulates tissue polarity and cell fate decisions in Drosophila. Dev. Camb. Engl. 125, 1149–59.

Xu, Y.K., Nusse, R., 1998. The Frizzled CRD domain is conserved in diverse proteins including several receptor tyrosine kinases. Curr. Biol. 8, R405–R406. https://doi.org/10.1016/S0960-9822(98)70262-3

Yamada, M., Udagawa, J., Matsumoto, A., Hashimoto, R., Hatta, T., Nishita, M., Minami, Y., Otani, H., 2010. Ror2 is required for midgut elongation during mouse development. Dev. Dyn. Off. Publ. Am. Assoc. Anat. 239, 941–53. https://doi.org/10.1002/dvdy.22212

Yamamoto, H., Yoo, S.K., Nishita, M., Kikuchi, A., Minami, Y., 2007. Wnt5a modulates glycogen synthase kinase 3 to induce phosphorylation of receptor tyrosine kinase Ror2. Genes Cells 12, 1215–1223. https://doi.org/10.1111/j.1365-2443.2007.01128.x

Yamamoto, Y., Takeshita, H., Sawa, H., 2011. Multiple Wnts Redundantly Control Polarity Orientation in Caenorhabditis elegans Epithelial Stem Cells. PLoS Genet. 7, e1002308. https://doi.org/10.1371/journal.pgen.1002308

Yu, J., Chen, L., Cui, B., Ii, G.F.W., Shen, Z., Wu, R., Zhang, L., Zhang, S., Briggs, S.P., Kipps, T.J., 2016. Wnt5a induces ROR1 / ROR2 heterooligomerization to enhance leukemia chemotaxis and proliferation 126, 585–598. https://doi.org/10.1172/JCI83535DS1

Zetka, M.-C., Rose, A.M., 1992. The Meiotic Behavior of an Inversion in Caenorhabditis elegans. Genetics 131, 321–332.

Zhang, W., Coldefy, A.-S., Hubbard, S.R., Burden, S.J., 2011. Agrin binds to the N-terminal region of Lrp4 protein and stimulates association between Lrp4 and the first immunoglobulin-like domain in muscle-specific kinase (MuSK). J. Biol. Chem. 286, 40624–30. https://doi.org/10.1074/jbc.M111.279307

Zhao, Z., Boyle, T.J., Liu, Z., Murray, J.I., Wood, W.B., Waterston, R.H., 2010. A Negative Regulatory Loop between MicroRNA and Hox Gene Controls Posterior Identities in Caenorhabditis elegans. PLoS Genet. 6, e1001089. https://doi.org/10.1371/journal.pgen.1001089

7 Appendix

The following tables are a summary of mutants isolated from a forward genetic screen for pharyngeal gland cell defects. Penetrance is estimated unless "n" is given. Hw: Hawaiian, UM: under-migration, OM: over-migration, X: outcross to N2, LG: linkage group, het: heterozygous, NM: non-mutant, WGS: whole genome sequencing

Table A1 Recessive lethal under-migration strains

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK17	*kor1*	-	-	-	-	Recessive lethal. Severe UM. Penetrance <5%.
JDK26	*kor6*	-	-	-	-	L4 recessive lethal. UM.
JDK27	*kor54*	-	-	-	-	L1 recessive lethal. g1P UM. Penetrance ~27%.
JDK28	*kor55*	-	-	-	-	Recessive lethal. UM. Penetrance 40%
JDK 45	*kor7*	-	-	-	-	Recessive lethal. UM. Penetrance high.
JDK 71	*kor17*	-	-	-	-	UM. Recessive lethal. Clr phenotype.
JDK 72	*kor18*	-	-	-	-	UM. Recessive lethal.
JDK 74	*kor20*	-	-	-	-	UM. Recessive lethal. Some males present.
JDK 84	*kor30*	-	-	-	-	UM @ pm4. L1 recessive lethal.
JDK 85	*kor31*	-	-	-	-	UM. Recessive lethal. Fragmented extensions, some gland cells outside pharynx.
JDK 88	*kor34*	-	-	-	-	UM. L1 recessive lethal. Pun. Pinched extensions. Penetrance <10%.
JDK 90	*kor36*	-	-	-	-	UM. Recessive lethal. Penetrance <5%
JDK 93	*kor39*	-	-	-	-	UM. Recessive lethal. Protruding vulva. Penetrance 30-50%.
JDK 94	*kor40*	-	-	-	-	UM. Recessive lethal. Penetrance 10%.

Table A2 Recessive lethal strains with mixed under-migration and over-migration phenotype.

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK20	*kor2*	-	-	-	-	Recessive lethal. g1P UM, some g1A OM. Penetrance 5%.
JDK21	*kor3*	-	-	-	-	Recessive lethal. UM and g1P OM. Penetrance 5-10%.
JDK29	*kor56*	-	-	-	-	Recessive lethal. UM and g1p OM. Penetrance 48% penetrance.
JDK30	*kor57*	-	-	-	-	L1 recessive lethal. UM and g1P OM. Penetrance 50%
JDK46	*kor10*	-	-	-	-	Recessive lethal. Penetrance high. UM. May be allelic to JDK47 and 48.
JDK 47	*kor58*	-	-	-	-	UM. May be allelic to JDK46.
JDK 48	*kor59*	-	-	-	-	UM. May be allelic to JDK46.
JDK 81	*kor27*	-	-	-	-	UM, some OM. L1 recessive lethal. Severe Dpy.
JDK 82	*kor28*	-	-	-	-	UM and OM, L1 recessive lethal. Penetrance ~5%. Dead eggs.
JDK 83	*kor29*	-	May be allelic to JDK 82			UM and OM L1 recessive lethal. From same P0 as JDK82.
JDK 86	*kor32*	-	-	-	-	OM, some UM. Recessive lethal. Some gland cells outside pharynx.
JDK 92	*kor38*	-	-	-	-	UM. Occasional lethal L1 extreme OM. Penetrance 5-15%.
JDK 104	*kor50*	-	-	-	-	UM and OM. L1 recessive lethal. Penetrance 15-20%. Gland cell fragmentation.
JDK 105	*kor51*	-	-	-	-	UM and OM. L1 recessive lethal. Slow-growing.

Table A3 Strains with recessive sterile under-migration phenotype.

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK 40	*kor8*	-	-	-	-	UM. Recessive sterile. Penetrance 70% Large body size.
JDK 80	*kor26*	-	-	-	-	UM. Recessive sterile. Penetrance 1-5%.

Table A4 Strains with homozygous viable under-migration phenotype

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK32	*kor5*	LG IV	Allelic to JDK58 LG IV Group B	WGS Dec 2016, *ham-1(gm267/kor5)* G>A SNP, Gly58Arg	JDK33 1x JDK50 2x JDK119 3x	UM. Homozygous viable. Penetrance 65% (n=85)
JDK53	*kor11*	-	-	-	-	UM. Homozygous viable. Penetrance 25-30%.
JDK 56	*kor13*	LG III, V, X	-	-	-	UM. Homozygous viable. Slow-growing. Penetrance decreased with outcrossing. Complex chromosome rearrangement or multigenic basis of phenotype suspected.
JDK 57	*kor14*	Not completed*	-	-	-	UM. Homozygous viable. Variable YFP expression. *Penetrance decreased with outcrossing.
JDK 58	*kor15*	LG IV	Allelic to JDK 32 (LG 4 Group B)	WGS Dec 2016, *ham-1(kor15)* C>T SNP, Gln202STOP	JDK110 1x JDK112 2x JDK114 3x	UM. Homozygous viable. Penetrance 79% (n=80).
JDK 59	*kor16*	Not completed*	-	-	-	UM. Homozygous viable. Penetrance 15-30%. *Hw outcrossing decreased penetrance, sick progeny
JDK 73	*kor19*	-	-	-	-	UM. Homozygous viable. Penetrance ~30%. Males present. Clr phenotype. Slow-growing.
JDK 75	*kor21*	-	-	-	-	UM. Homozygous viable. Penetrance 1-5%.
JDK 76	*kor22*	LG IV	Allelic to JDK 99, 103, 106 (LG 4 Group A)	Sanger sequencing, Jul-Aug 2017, *cwn-2* G>A SNP, W8STOP	-	UM. Homozygous viable. Penetrance 60-70%.
JDK 77	*kor23*	LG V	Allelic to JDK 100 15-18% mutants in trans progeny	-	-	UM. Homozygous viable. Penetrance 40-50%. Appears allelic to JDK100.
JDK 78	*kor24*	-	-	-	-	UM. Homozygous viable. Penetrance 5-10%.

Table A5 Strains with homozygous viable under-migration phenotype, continued.

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK 79	*kor25*	-	-	-	-	UM. Homozygous viable. Penetrance 5%.
JDK 87	*kor33*	-	-	-	-	UM. Homozygous viable. Pentrance 10-15%.
JDK 89	*kor35*	-	-	-	-	UM. Homozygous viable. Penetrance 50-70%.
JDK 91	*kor37*	LG II	Fails to complement *cam-1(gm122)*	Sanger Jul-Aug 2017, *cam-1* single bp deletion.	-	UM. Homozygous viable. Wit. Single bp deletion occurs in exon 2 of isoform A, causes a frameshift and early stop codon in proceeding exon.
JDK 95	*kor41*	Not completed*	-	-	-	UM. Homozygous viable. Penetrance 40%.*Severe penetrance reduction with Hw outcrossing.
JDK 96	*kor42*	LG I	-	WGS Dec 2016, in progress	JDK117 3x	UM. Homozygous viable. Penetrance 20% (n=88). Egl and bag of worms.
JDK 97	*kor43*	LG IV	Ambiguous results for both LG 4 Group A and B	Sanger Jul-Aug 2017, *cwn-2* C>T SNP, Q236STOP	-	UM. Homozygous viable. Penetrance 70%.
JDK 98	*kor44*	LG X	-	WGS Dec 2016, in progress	JDK118 3x	UM. Homozygous viable. Penetrance 46% (n=98).
JDK 99	*kor45*	LG IV	LG 4 "Group A," see JDK 76	WGS Dec 2016, *cwn-2* G>A SNP, W26STOP	-	UM. Homozygous viable. Penetrance 40%. Constitutive dauer.
JDK 100	*kor46*	LG V	Allelic to JDK77?	-	-	UM. Homozygous viable. Mild Unc.
JDK 101	*kor47*	-	-	-	-	UM. Homozygous viable. Penetrance 10-15%
JDK 102	*kor48*	-	-	-	-	UM. Homozygous viable. Penetrance 15-20%.
JDK 103	*kor49*	LG IV	LG 4 Group A, see JDK 76	WGS Dec 2016, *cwn-2* C>T SNP, Gln54STOP	-	UM. Homozygous viable., Penetrance 80%.
JDK 106	*kor52*	LG IV	LG 4 Group A, see JDK 76	Sanger Jul-Aug 2017, *cwn-2* G>A SNP, W89STOP	-	UM. Homozygous viable. Penetrance 70%.

Table A6 Strains with variable phenotype

Strain	Allele	Linkage Mapping	Complementation	Sequence	Outcross	Description
JDK25	*kor4*	-	-	-	-	Uncharacterized.
JDK31	*kor9*	-	-	-	-	Dominant sterile. UM. Penetrance 60%
JDK 49	*kor60*	-	-	-	-	UM. Uncharacterized.
JDK 54	*kor12*	-	-	-	-	Ectopic fluorescence of nucleus in posterior bulb.
JDK 107	*kor53*	-	-	-	-	UM and OM. Variable phenotype. Homozygous viable.