UNIVERSITY OF CALIFORNIA

Santa Barbara

From guts to glory: Investigating modulators of the *C. elegans* endoderm gene regulatory network

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy in Molecular, Cellular and Developmental Biology

By Geneva Louise Alok

Committee in charge:

Professor Joel Rothman, Chair

Professor Denise Montell

Professor William Smith

Professor Thomas Turner

December 2020

The dissertation of Geneva Louise Alok is approved.

_____
Denise Montell

_____
Thomas Turner

_____
William Smith

_____
Joel Rothman, Committee Chair

December 2020

DEDICATION

This dissertation is dedicated to *all* of my parents- Donna Bradford, David Miller, Glenn Bradford, Beena Singh, and Alok Singh. You have all continuously supported me on this journey, and your words of encouragement have meant the world to me. To my siblings, Caleb and Elle- both of you have motivated me to keep going, even when my goals felt impossible. Thank you for believing in me.

To all of the Rothman lab members that sat with me through marathon lab meetings, I am inspired by and grateful for your ability to give meaningful feedback, even three hours into a presentation. I truly found my Santa Barbara family in the lab, and I couldn't have survived this process without you guys. It's been an honor to work with all of you, and you've set a high bar for all of my future co-workers.

Finally, this dissertation is dedicated to my husband, Apurva. There aren't words to describe how appreciative I am for the love and support you've bestowed upon me during this leg of our journey. From driving me into lab every day to always helping me when I want to throw my computer at the wall, you're simply the best.

ACKNOWLEDGEMENTS

First, I'd like to thank my advisor, Joel Rothman, for countless hours of curious and creative scientific discussions. Your support has pushed me to be a better scientist, and I'm proud of the work we've accomplished together. I'd also like to thank Dr. Pradeep Joshi, who was always ready and willing to talk through experiments and ideas with me at a moment's notice. Additionally, the Rothman lab couldn't run without the dutiful work of Cricket Wood. Thank you for taking care of us.

Thank you to my committee, Denise Montell, William Smith, and Thomas Turner, for guiding me through this process and elevating my research.

I'd also like to acknowledge the different departments within the university that helped me complete this dissertation. My home department, MCDB, and the NRI provided tremendous support, both financially and mentally. Thank you to Jen Smith with the BNL for providing crucial help with my sequencing experiments. I'm grateful to the CNSI, the O'Malley lab, and the Daugherty lab for allowing access to the FACS machine, without which two of the following chapters wouldn't exist. Also thank you to the Center for Stem Cell Biology and Engineering for granting access to the qPCR machine.

Several graduate students from outside the Rothman lab provided vital support when I was attempting to use new equipment or explore new methods, so I would like to acknowledge Joel Bozekowski, Yunning Chen, Justin Yoo, and Mohamed Faynus for their willingness to assist me as I stumbled through new processes.

Lastly, I'd like to acknowledge the worm community. It was my pleasure to work on a model organism with such a dedicated following. The community is always striving to share knowledge and resources, and I am eternally grateful for all of the scientists who have contributed to WormBase, which has been my top visited website for the better part of my PhD.

CIRICULUM VITAE—Geneva Louise Alok

Education

Bachelor of Arts in Biological Sciences from Columbia University in the City of New York—
Granted May 22, 2013
G.P.A.—3.39/4.00

Molecular, Cellular and Developmental Biology PhD program at University of CA, Santa Barbara—
July 2014 - present
G.P.A—3.97/4.00

Grants/honors and awards

Amgen Scholar, summer 2011
Carbon Clarke Fellowship, summer 2014
Storke Award, May 2015
Devlin Fellowship, December 2015
MCDB Graduate Program Fellowship, summer 2019
Amgen Fellowship, March 2020

Teaching Experience

University of California, Santa Barbara, TAships:
Intro Biology Lab 1AL- fall 2013, instructor: Douglas Bush
Eukaryotic Genetics 101B- winter 2016, instructor: Tau Mu Yi; winter 2017, 2018 & 2019, instructor: Julie Simpson
Genetics Lab 101L- fall 2018, fall 2019 (head TA), fall 2020, instructor: Pradeep Joshi
Prokaryotic Genetics 101A- winter 2020, instructor: Steve Pool; summer 2020 (head TA), instructor: Doug Thrower

Research experience

University of California, Santa Barbara July 2014-Present
Role: graduate student in Molecular, Cellular and Developmental Biology PhD program
1st rotation (July 2014-Sept. 2014):
Principle Investigator: Dr. Denise Montell
Research Area: Understanding the process of anastasis in mammalian cell culture
Bio II Building, 4th floor, University of California, Santa Barbara
2nd rotation (Oct. 2014- Dec. 2014):
Principle Investigator: Dr. Tony De Tomaso
Research Area: Vascular fusion and rejection in *B. schlosseri*

Life Sciences Building, 1st floor, University of California, Santa Barbara
PhD work (Jan. 2015- present):
Principle Investigator: Dr. Joel Rothman
Research Area: Endoderm development in *C. elegans*
Bio II Building, 2nd floor, University of California, Santa Barbara

The Rockefeller University June 2013- May 2014
Role: Research assistant
Principle Investigator: Dr. Howard Hang
Research Area: The role of commensal bacteria in innate immunity
Collaborative Research Center, The Rockefeller University, New York City, NY, USA

Columbia University Depart. Of Biological Sciences Sept 2009- May 2013
Roles: Assistant lab technician, undergraduate researcher
Principle Investigator: Dr. Martin Chalfie
Research area: Mechanosensation transduction in the touch neurons of *C. elegans*
Sherman Fairchild Center, Columbia University, New York City, NY, USA

Publications

- **Alok, G.**, Ewe, C.K., Cleuren, Y.N.T., Joshi, P., and Rothman, J.H. (2020). BRAP-2 negatively regulates Wnt-mediated endoderm specification in *C. elegans*. *(in preparation)*
- **Alok, G.**, Ackley, C., Shen, Y., Wood, C., Mengarelli, I., Ewe, C.K., Flowers, S., Joshi, P., and Rothman, J.H. (2020). Largescale mutagenesis, sorting, and bulk sequencing of *C. elegans* embryos reveals *act-4* as a novel modulator of the endoderm gene regulatory network. *(in preparation)*
- Ewe, C.K., **Alok, G.**, and Rothman, J.H. (2020). Stressful Development: integrating endoderm development, stress and longevity. Developmental Biology.
- Ewe, C.K., Cleuren, Y.N.T., Flowers, S.E., **Alok, G.**, Snell, R.G., and Rothman, J.H. (2020). Natural cryptic variation in epigenetic modulation of an embryonic gene regulatory network. Proc. Natl. Acad. Sci.
- Ewe, C.K., Torres Cleuren, Y.N., **Alok, G.**, and Rothman, J.H. (2019). ICD-1/BTF3 antagonizes SKN-1-mediated endoderm specification in Caenorhabditis elegans. MicroPublication Biol.

ABSTRACT

From guts to glory: Investigating modulators of the *C. elegans* endoderm gene regulatory network

-Geneva Louise Alok-

Proper specification of cell fates and organ identity is critical to the creation of a fully functional organism during animal development. Uncovering the dynamic events underlying deployment of gene regulatory networks (GRNs) that dictate cell fate and organ identity is critical to understanding the mechanisms that ensure developmental robustness and the relationship between the genesis and function of organs. In *C. elegans*, the entire intestine arises from a single cell, the E (endoderm) blastomere, which is specified by the seven-cell stage of embryogenesis. As with nearly all other cells throughout development of the animal, the E cell undergoes a stereotyped pattern of cell divisions, giving rise to 20 intestinal cells in the fully developed embryo before hatching into a first-stage larva. While many of the key, phylogenetically conserved regulatory factors comprising the endoderm GRN, including a cascade of GATA-type transcription factors, have been identified, the function and identity of modulators that regulate faithful execution of the GRN during endoderm development have not been comprehensively revealed. To access the broad set of genetic modulators underlying the endoderm GRN, I have developed a high-throughput pipeline based on bulk sorting of *C. elegans* embryos and large-scale genomic sequencing and applied it to identifying genomic regions responsible for natural variation in its execution. This method revealed quantitative trait loci (QTLs) that overlap with those identified by classical strategies, thus validating the approach, and also identified additional QTLs underlying variation in execution of endoderm development. I found that one candidate gene, encoding

BRAP-2 (BRCA1-associated protein-2), regulates endoderm specification by modulating expression of a core regulator, the END-1 GATA factor, likely by tuning the activity of key maternal regulatory inputs into the GRN. I further applied this method to the bulk analysis of embryos following mutagenesis to assess genetic regions that, when mutated, alter gut development. This latter strategy revealed that components in actin networks play a previously unknown role in ensuring the proper number of endoderm cells. This study contributes to our understanding of how regulatory inputs fine-tune an embryonic GRN, resulting in the robust production of an entire germ layer, the endoderm, and a crucial organ, the gut.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendix 1

Appendix 2

Appendix 3

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 1

Introduction to the endoderm GRN in *C. elegans*

1.1 Primary transcriptional cascade

Development is driven by progressive activation of transcriptional programs in a gene regulatory network (GRN). Robust cellular specification and differentiation depend on strict spatial and temporal regulation of gene expression (Levine & Davidson, 2005; Little et al., 2013). Suboptimal transcriptional inputs may lead to developmental failure and aberrant cellular behavior. Evolutionarily conserved GRNs exist throughout the animal kingdom to maintain core function of basic organs. Diversity in body plan and specialized functions arise from genetic modulators acting on the core network (reviewed in Peter & Davidson, 2011). In addition, modulators help to fine-tune gene expression to ensure reproducibility despite environmental or genetic perturbation. It is of great interest to completely understand every genetic and cellular interaction that exists within a single GRN so that we may fully connect genotype to phenotype.

The development of the *C. elegans* endoderm provides a highly tractable model to study developmental robustness in a complex GRN. The entire germ layer, which makes up the twenty cells of the intestine, arises from the E blastomere (Sulston et al., 1983). The E blastomere is the posterior cell born out of an asymmetrical division of the mesoendodermal precursor EMS. In E, maternal SKN-1/Nrf transcription factor activates a cascade of GATA-like transcription factors, starting with *med-1/2* and culminating with *elt-2*, whose expression is maintained throughout the life of the worm to support transcription of genes necessary for proper gut function (Blackwell et al., 1994; Broitman-Maduro et al., 2005; Fukushige et al.,

1998; Maduro et al., 2001, 2005; reviewed in Maduro & Rothman, 2002). MED-1/2 activates *end-3*, END-3 activates *end-1*, END-1 activates *elt*-7, and ELT-7 activates *elt*-2 (Fig. 1) (Broitman-Maduro et al., 2005; Fukushige et al., 1998; Maduro et al., 2001, 2007; McGhee et al., 2007; Sommermann et al., 2010; Zhu et al., 1997). Duplicated genes *elt-7* and *elt-2* maintain their expression through autoregulation, as well as their ability to upregulate each other (Sommermann et al., 2010).

Ensuring robustness of this process, the primary endoderm transcriptional cascade uses duplicated genes in a redundant fashion and employs a feed-forward loop motif, where upstream factors activate the next gene in the cascade as well as the preceding gene (Fig. 1) (reviewed in Maduro, 2017). For example, *end-1/3* are the products of a gene duplication event, and, in the core pathway, END-3 upregulates *end-1* transcription, but it can also activate *elt-7* transcription, traditionally thought of as the target of END-1 (Sommermann et al., 2010; Zhu et al., 1997). ELT-7 turns on *elt-2*, but it can also activate ELT-2 target genes, and END-1 can substitute for ELT-7 in upregulating *elt-2* (reviewed in Maduro, 2017; Sommermann et al., 2010; Wiesenfahrt et al., 2016). This system ensures fidelity of the pathway even if one gene is compromised, although of the terminal differentiation factors ELT-2 and ELT-7, ELT-2 appears more necessary as *elt-2* mutants sometimes lack gut granules in the most anterior and posterior intestinal rings, as well as the middle ring, and gut integrity diminishes after hatching, whereas *elt-7* mutants display no notable mutant gut phenotype (T. Fukushige et al., 1998; Sommermann et al., 2010). This cascade begins very early on in the embryo, as soon as the E cell is born at the 8-cell stage and relies on the presence of SKN-1 in EMS at the 4-cell stage. Surprisingly, it has been shown that even in the absence of maternal SKN-1, ~30% of embryos can still produce endoderm (Bowerman et

al., 1992), indicating the existence of alternative activators. It has been shown that SKN-1-independent gut development is primarily dependent on signals from the EMS-adjacent cell, P2 (Bei et al., 2002; Lin et al., 1995, 1998; Meneghini et al., 1999; Rocheleau et al., 1997, 1999; Shetty et al., 2005; Shin et al., 1999; Sumiyoshi et al., 2011; Thorpe et al., 1997).

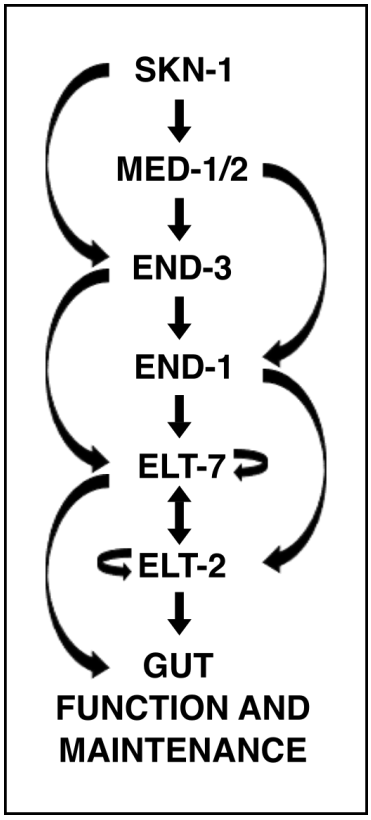

Figure 1 Primary endoderm gene cascade

Diagram showing the order of the primary endoderm transcriptional cascade, ending in the transcription of 1000's of genes required for normal gut function and structural maintenance. Arrows indicate a feed-forward loop, where transcription factors can activate genes immediately below them as well as the target in the next tier. Autoregulation of ELT-7 and ELT-7 are shown in addition to their ability to activate each other.

1.2 Signaling from P2 (see also Fig. 2)

Convergence on POP-1

Without intervention, *end-1* and *end-3* are repressed in EMS-daughter cells by TCF/LEF orthologue POP-1. In the anterior EMS-daughter cell, MS, *end-1/3* remains repressed by POP-1, and MED-1/2 alternatively activates MS-lineage specifier *tbx-35*

(Broitman-Maduro, 2006). However, POP-1 repression is relieved in the E cell by Wnt /MAPK and Src signaling from the adjacent P2 cell. Wnt/MAPK and Src signaling converge on and phosphorylate POP-1, turning it from a repressor of *end-1/3* gene expression to an activator of the *end* genes (Bei et al., 2002; Lin et al., 1995, 1998; Meneghini et al., 1999; Rocheleau et al., 1997, 1999; Shetty et al., 2005; Shin et al., 1999; Sumiyoshi et al., 2011; Thorpe et al., 1997). This coincides with low nuclear concentrations of POP-1 in its activating state, and high nuclear concentrations in its repressive state (Maduro et al., 2005; Shetty et al., 2005). These expression levels are inversely correlated with nuclear levels of Wnt-regulated SYS-1, a β-catenin that acts as a transcriptional co-activator with POP-1 (Huang et al., 2007; Liu et al., 2008; Phillips et al., 2007). As such, the E cell contains low nuclear POP-1 and high nuclear SYS-1, when compared to its anterior sister cell, MS. In fact, this reciprocal POP-1/SYS-1 patterning is maintained throughout all anterior-posterior (A-P) divisions, showing a crucial role for POP-1 and SYS-1 in A-P cell fate patterning (Huang et al., 2007; Phillips et al., 2007). In terms of endoderm development, Wnt/MAPK and Src signaling from P2 to E are responsible for modifying POP-1 to allow progression of the endoderm GRN.

Wnt signaling

Canonical Wnt signaling is initiated when a Wnt ligand is recognized by a Frizzled receptor, which activates Dishevelled in the receiving cell, in turn inhibiting GSK3β such that β-catenin can activate TCF/LEF. Activated TCF/LEF is free to induce transcription of its targets (reviewed in Wodarz & Nusse, 1998). In endoderm specification in *C. elegans*, the maternal Wnt ligand MOM-2 is released from P2 and binds to the MOM-5/Frizzled receptor at the posterior end of EMS, the end which is to become E (Rocheleau et al., 1997; Thorpe et

al., 1997). MOM-5 activates SGG-1/GSK3β which then activates WRM-1/β-catenin (Maduro et al., 2001). When activated, WRM-1 phosphorylates POP-1/TCF, which serves dual functions. Phosphorylated POP-1 no longer represses *end-1/3* and is predominantly exported from the nucleus (Lin et al., 1995, 1998; Maduro et al., 2005; Rocheleau et al., 1999; Shetty et al., 2005). The remaining nuclear POP-1 additionally positively contributes to *end-1/3* expression with its co-factor SYS-1/β-catenin (Huang et al., 2007; Maduro et al., 2005; Phillips et al., 2007; Shetty et al., 2005).

MAPK signaling

Another maternal factor, MAPKKK-related protein MOM-4/Tak1, is necessary in EMS for proper E lineage differentiation. MOM-4 phosphorylates NEMO-like kinase (NLK) LIT-1, which functions with WRM-1 to phosphorylate POP-1 (Meneghini et al., 1999; Rocheleau et al., 1999; Shin et al., 1999; Thorpe et al., 1997). Evidence suggests that MOM-2/Wnt from P2 is may actually be the activating signal for MOM-4 in EMS (Smit et al., 2004).

Src signaling

Src signaling has been shown to work in parallel with Wnt/MAPK, causing phosphorylation and nuclear localization of HMP-2/β-catenin which functions redundantly with WRM-1 in endoderm induction (Bei et al., 2002; Sumiyoshi et al., 2011). MES-1, a transmembrane tyrosine kinase localized at the junction between P2 and EMS, signals to tyrosine kinase SRC-1, which phosphorylates HMP-2 (Bei et al., 2002; Berkowitz & Strome, 2000; Sumiyoshi et al., 2011). HMP-2 is mostly recognized for its contribution to cell adhesion processes (Costa et al., 1998; Korswagen et al., 2000). APC homolog APR-1 and

Fer like kinase FRK-1 work to repress HMP-2, and either *apr-1* or *frk-1* knockdowns result in hyperproliferation in the endoderm (Putzke & Rothman, 2010; Rocheleau et al., 1997).

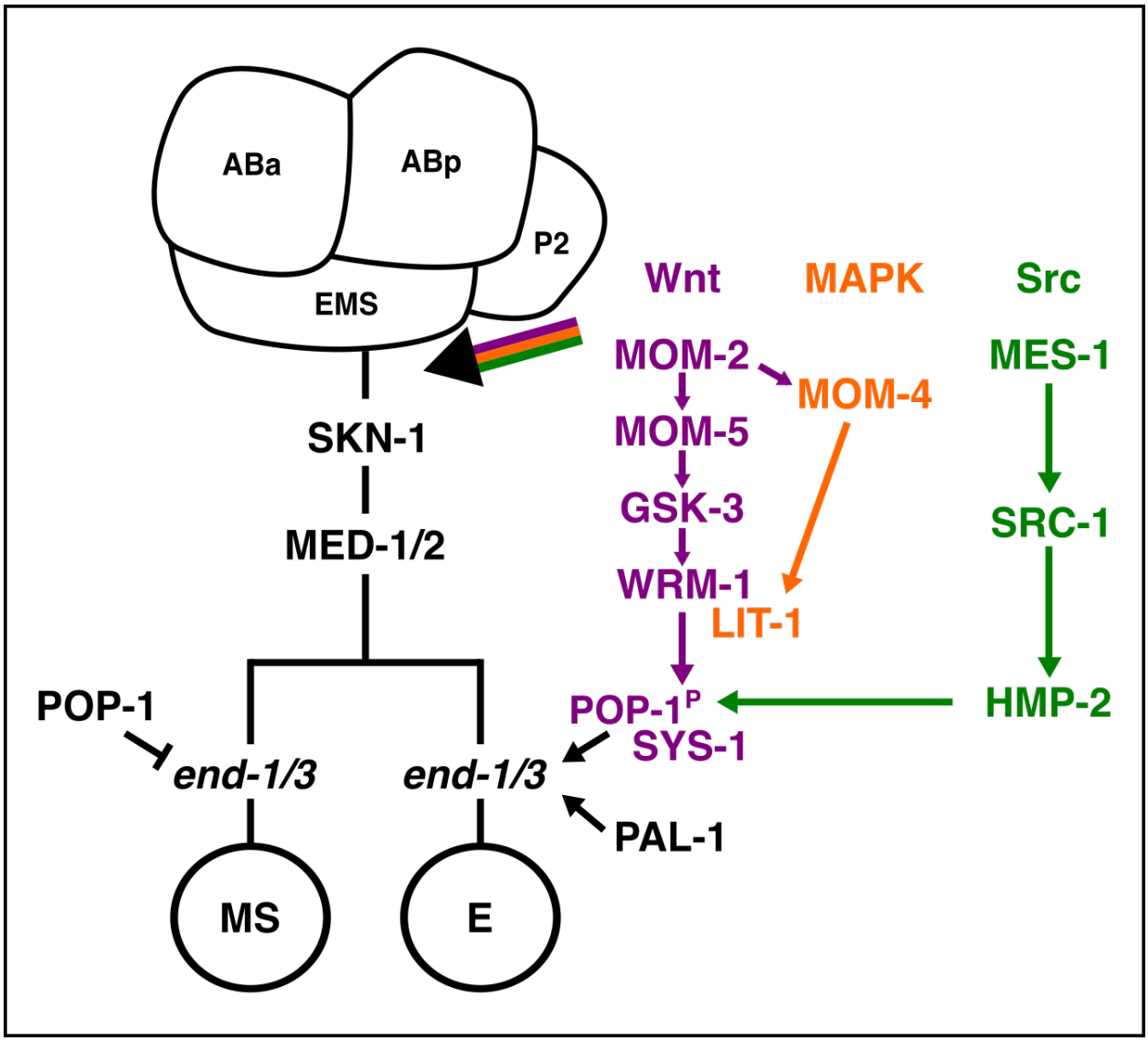

Figure 2 Endoderm GRN is reliant on signaling from P2

In the 4-cell embryo, Wnt/MAPK, and Src signaling from P2 to EMS converge on POP-1, turning it from a repressor of *end-1/3* into an activator. The loss of POP-1 repression defines the E lineage, and the transcriptional cascade beginning with SKN-1 leads to the activation of *end-1/3*.

1.3 Known modulators of the endoderm GRN

PAL-1

The Caudal homeoprotein PAL-1 is the major specifier of mesectodermal cell fate in the C and D lineages, producing muscle (Hunter & Kenyon, 1996; Lei et al., 2009). PAL-1 is a maternal factor, and it is first activated in C, the granddaughter of P1, which is the posterior cell of the 2-cell embryo. P1 divides to create EMS and P2, and P2 divides into C and P3. P3 continues to divide into D and P4 (J. E. Sulston et al., 1983). In order to properly specify SKN-1 activity in EMS and PAL-1 activity in C and D, evidence suggests that PIE-1 protein limits SKN-1 and PAL-1 activity in the germline precursor cells (P1, P2, P3, and P4) (Hunter & Kenyon, 1996; Seydoux et al., 1996). In the EMS lineage, ruled by SKN-1, the prevailing view is that SKN-1 inhibits PAL-1 function, as antibody staining confirms the presence of both in EMS (Bruce Bowerman et al., 1993; Hunter & Kenyon, 1996). PAL-1 can function in the C and D lineages in the absence of PIE-1, which is selectively sorted into the germline precursors (Mello et al., 1996; Seydoux et al., 1996), and it is thought that SKN-1 has decayed in C and D lineages, so it can no longer inhibit PAL-1, allowing mesectodermal differentiation to proceed (Hunter & Kenyon, 1996). Additionally, SGG-1/GSK3β can repress MED-1/2 in the C lineage (Maduro et al., 2001). In this way, SKN-1 activity is limited to the EMS lineage, and PAL-1 is allowed to function in C and D.

In the absence of SKN-1, EMS produces PAL-1 dependent muscle cells (Hunter & Kenyon, 1996). Surprisingly, PAL-1 also provides a minor positive input into the endoderm GRN (Fig. 2). Knockdown of SKN-1 and POP-1 together, the two dominant activators of the of endoderm GRN, still shows 5-32% of embryos (depending on use of mutant strains or

RNAi) that can make gut cells. However, RNAi knockdown of *skn-1*, *pop-1*, and *pal-1* in combination completely represses endoderm specification (Maduro et al., 2005).

ICD-1

Additionally, a homologue of human BTF3, ICD-1, antagonizes SKN-1 input in *mom-2* (Wnt) or *mom-4* (MAPKKK) mutants (Ewe et al., 2019). RNAi knockdown of ICD-1 in either of these mutants increases the number of embryos that can make endoderm, and we hypothesize that ICD-1 inhibits SKN-1 action by binding to and repressing *end-1/3* promoters. ICD-1 is added to a list of factors that seemingly work to fine-tune the activating inputs in the endoderm GRN.

Genetic screen for genes that repress endoderm hyperproliferation

Previous work from the Rothman lab identified 58 genes from a genome-wide RNAi screen that cause excess endoderm in embryos (Mengarelli, 2006). This list includes genes with a wide-breadth of known and predicted functions, including cyclin orthologue *cyl-1*, NEDD8 ubiquitin-like modifier orthologue *ned*-8, tensin orthologue *tns*-1, SF1 (splicing factor 1) orthologue *sfa*-1, and many others. Within this list, we see a high GO enrichment (>100 fold) for genes involved in cell proliferation, centrosome regulation, and p53 signaling. It is clear from this screen alone that we do not fully understand all of the regulatory factors at play within the endoderm GRN.

QTL analysis

More dissertation work from the Rothman lab has solidified that we do not have a complete understanding of all the modulators acting on the endoderm GRN. While studying the genomic regions controlling variability in endoderm determination among different isotypes of *C. elegans*, analysis pointed to several large regions across the genome (Torres

Cleuren, 2016). Out of these candidate regions, work narrowed in on Pur alpha like Protein PLP-1 and TORC2 signaling component RICT-1, as knockdown of either was previously shown to rescue endoderm in the absence of *skn-1* (Ruf et al., 2013; Witze et al., 2009). Newly reported, however, was that their knockdown also suppresses endoderm in the absence of *mom*-2 (Torres Cleuren et al., 2019), although the specific mechanism through which they are acting is unknown. This reciprocal relationship between factors effecting SKN-1 and MOM-2 activating inputs is also seen with MIG-5, a Dishevelled homologue known to act downstream of Wnt signaling in the AB lineage (Walston et al., 2004). MIG-5 has the opposite effect as PLP-1 and RICT-1; RNAi knockdown of *mig-5* decreases endoderm in the absence of *skn-1* and increases endoderm in the absence of *mom-2* (Torres Cleuren et al., 2019). This reciprocal relationship between SKN-1 and Wnt signaling among different isotypes is evidence of rapid evolution within the endoderm GRN and points to plasticity within the network. While two genes on chromosomes II (*rict-1* and *mig-5*) and one gene on chromosome IV (*plp-1*) were identified from work on the wild isolates, QTL analysis points to other regions on chromosomes I, V, and X that contribute to the plasticity of the GRN. We do not yet have a complete understanding of all genes deployed within the network, nor do we fully comprehend how all of these genes are interacting together.

1.4 Aims of present study

As there is still much to uncover about the discreet modulators of the endoderm GRN, the present study aims to increase our understanding of this complex network. In Chapter 2 of this study, we identified another novel regulator of the GRN, BRAP-2 (BRCA-1 Associated Protein), and we sought to investigate its mode of action using classic genetics

methods. In addition, we explored two new screening methods to expand our search for genes that contribute to the network. Both of these new methods take advantage of the ability to collect embryos in bulk with a FACS machine and the capacity to rapidly perform whole-genome sequencing (WGS). The first new method, presented in Chapter 3, further investigates the genomic regions regulating the variability seen within the GRN among two different isolates and aims to identify modulators which are the substrates of evolutionary forces, resulting in the plasticity of GRN. The second method, further explained in Chapter 4, explores genes controlling increased endoderm differentiation, which helps to define the pathways contributing to the fine-tuning of activating inputs, ensuring robustness of the GRN and faithfully producing L1 larvae with 20 gut cells at the end of embryonic development. Validation of both new methods expands the possibilities in the field to rapidly interrogate a developmental GRN.

Chapter 2

Using recombination to find cryptic loci

2.1 Introduction

Most of the work done in the model organism *C. elegans* used descendant lines from a strain found in Bristol, England, in the 1950's (Nicholas et al., 1959; Nigon & Félix, 2017). We refer to this strain as the N2 reference strain, but since the turn of the century, more researchers are using *C. elegans* strains from around the globe, called wild isolates (compared to the domesticated N2 strain), to study the genetic variation within the species (Barriere, 2005; Barrière & Félix, 2005; Cook et al., 2017; Denver et al., 2003; Haber et al., 2005; Hodgkin & Doniach, 1997; Knight et al., 2001; Schulenburg & Müller, 2004; Sivasundar & Hey, 2003). These wild isolates represent over 400 unique, inbred haplotypes, and researchers have used this variation to study the basis of diverse phenotypes within *C. elegans*. Genome wide association studies (GWAS) using many strains in one survey have been used to pinpoint quantitative trait loci (QTLs) that are responsible for this species diversity (Andersen et al., 2012; Cook et al., 2016; Evans et al., 2017). Furthermore, recombinant inbred lines (RILs) created by mating two different isolates has been beneficial community resources for identifyingQTLs underlying various traits (Elvin et al., 2011; Greene et al., 2016; Shook et al., 1996; Shook & Johnson, 1999). These wild isolates present a unique opportunity to study how a gene regulatory network can vary among members of the same species.

Previous work from the lab has shown that the SKN-1 requirement for endoderm varies among naturally occurring wild isolates (Torres Cleuren et al., 2019). Out of nearly

one hundred strains studied, it appears that some isolates have an absolute requirement for SKN-1 and can never produce embryos with endoderm in the absence of *skn-1*, while in other strains ~60% of embryos develop endoderm under the same conditions. We see a range of phenotypes between these two extremes in the other strains tested. Earlier work from the lab used several methods to identify QTLs responsible for the variation in preponderance of gut in absence of *skn-1* across the different wild isolates. GWAS across all the isolate strains identified a large QTL (> 13Mb) on chromosome IV (Torres Cleuren, 2016; Torres Cleuren et al., 2019). A type of machine learning using ElasticNet regression models found smaller, yet still expansive, QTLs on all chromosomes except chromosome III, ranging from ~2Mbp to 6Mbp (Torres Cleuren, 2016). Additionally, RILs were generated between the laboratory reference strain (N2), which exhibits a partial dependence on *skn-1* producing ~30% of embryos with endoderm in a *skn-1* knockdown, and wild isolate strain MY16, which exhibits a greater dependence on *skn-1* producing just ~2.5% of embryos with endoderm under the same conditions. Single and multi-QTL analysis and bulk-segregant analysis (BSA) was performed using phenotypic and sequencing data derived from the RIL set, and more large QTLs were identified. A table of these different methods and the QTLs they produced is available in Yamila Torres-Cleuren's doctoral thesis (Torres Cleuren, 2016), although the ElasticNet analysis was later revised (unpublished data from Melissa Alcorn).

We aimed to test a method to further narrow in on those candidate regions. eXtreme-QTL (X-QTL) analysis has been described in yeast and *C. elegans* (*ce*X-QTL), where mating, phenotyping, and genotyping is carried out in bulk, allowing for the analysis of thousands of progeny (Burga et al., 2019; Ehrenreich et al., 2010). Here we describe a

similar method, although we used a slightly different genetic approach to ensure bulk mating among RILs. After 16 generations of continuous bulk mating among RILs, we collected embryos that can still make endoderm under *skn-1 (RNAi)* using a fluorescently activated cell sorter (FACS) machine, collecting embryos that expressed a gut-specific GFP marker. It has been previously shown that embryos can be successful sorted and collected using a FACS machine (Stoeckius et al., 2009). We then sequenced the collected population in bulk, along with an unsorted control population. Using a G-statistic as laid out in the *ce*X-QTL method (Burga et al., 2019), we were able to identify SNPs whose representation was most significantly altered between a bulk of unsorted embryos and a bulk of gut+ embryos (gut+ refers to their ability to make endoderm). We have compiled a list of candidate genes that surround those SNPs, and we are pursuing an RNAi screen. In support of this method, the majority of top SNPs identified by this method also corresponded with the QTL regions identified previously (Torres Cleuren, 2016). Validation of this method shows it potential strength in quickly identifying genomic regions that underlie variation in the endoderm GRN and may be applied across a multitude of developmental GRNs.

2.2 Methods and Materials:

Strains: JR3970 *spe-27(it132)* IV; *rrIs1(elt-2p::GFP)* X, JR3915 MY16; *wEX1738(elt-2::GFP + myo-2::mcherry)*, JR4101-JR4180—GPF+ RILs, JR3936— *dpy-13(e184) skn-1(zu67) IV / nT1[qIs51] (IV;V)*

Generating spe-27(it132) homozygous RILs: MY16 hermaphrodites with an *elt-2p::GFP* extrachromosomal array were mated with *ts-spe-27*; *elt-2p::GFP* males with an otherwise N2

background. The *ts-spe-27* allele has an E to K mutation at AA727 (G→A at nucleotide 5865726 IV). SPE-27 is necessary for hermaphroditic sperm production but not for male sperm production (Minniti et al., 1996). As such, at the restrictive temperature, only males can produce sperm turning the line into an obligate mating strain. GFP+ F2 lines were selected and tested for self-sterility at 25°C and fertility at 15°C, signifying a homozygous *ts-spe-27*. 2-4 larvae from each line were placed at 25°C for two weeks, and lines that succeeded in starving out the plate were discarded. On plates that did not starve out, the number of progeny was counted. Additionally, individual L4 larvae from each strain thought to be *ts-spe-27* homozygous and from both parental strains were grown at both 15°C and 25°C for one week, and the number of progeny produced was counted (2-13 replicates for each strain). 28 out of 80 total RILs were also sequenced confirmed, using PCR primers fwd TTTGCATATTCATGTTACTCGACA and rv TCAATTCGAACCATTATTCTTGA to amplify *spe-27* and sequencing through UC Berkeley Sequencing Facility using sanger sequencing. We did not see any false positives through the sequencing confirmation. Populations were maintained at 15°C, and male stocks were generated using a brief heat shock treatment in which 4-5 L4 hermaphrodites were placed at 37°C for 40 minutes. To test for mating efficiency at 25°C, one hermaphrodite from a subset of RILs was plated with several males from the same line, and progeny were counted after several days.

Bulk matings: About 30 males (L4 and adult) and 10 hermaphrodites (young adult and adult) from each line were placed together such that all lines were on one 150mm plate. Approximately 50,000 F1 were collected in liquid and redistributed to multiple 150mm plates to allow the bulk of worms to continue growing and mating. From the F2-F16,

approximately 250,000 worms were collected and replated at each generation. The fraction of males in the population was assessed at each generation to make sure it was still a mating population as successful matings result in half male progeny (Altun & Hall, 2006). GFP expression was also confirmed in the F16 population using an Olympus SZX12 stereo microscope.

Selection for embryos with skn-1-independent endoderm: L4s in the 16th generation were fed *skn-1 (RNAi)* for 24 hours before being transferred to another *skn-1 (RNAi)* plate for egg laying. Adults were washed off after several hours, and remaining embryos were allowed to develop for 12 hours. Arrested embryos were scraped from the plate and collected in M9 buffer. After several washes, embryos were resuspended in PBST and strained through a 40uM cell strainer and then kept on ice until sorting. Embryos were sorted for GFP expression using a BD FACSAria II system using a 100μM nozzle at low sheath fluid pressure (20 psi).

Whole genome sequencing: GFP+ sorted embryos, an unsorted control, and both parental strains were prepared for whole genome sequencing. Libraries were made using NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina and NEBNext® Multiplex Oligos for Illumina, and samples were sequenced at the UCSB Biological Nanostructures Lab (BNL) using 75bp Paired End reads on the Illumina NextSeq 500.

Generating a SNP map: FASTQ sequencing files were uploaded onto the Galaxy public server at usegalaxy.org (Afgan et al., 2018), and processed first with FASTQ joiner

(Blankenberg et al., 2010) and then Trimmomatic (Bolger et al., 2014) (average base quality of 25 required across 4 bases) before aligning to the WS235 version of the *C. elegans* reference genome from wormbase.org with Bowtie for Illumina (Langmead et al., 2009; *WormBase : Nematode Information Resource*). After aligning, the resulting BAM files were processed with Samtools Sort, Samtools Merge, RmDup, and finally Generate Pileup, all through the Galaxy workspace (H. Li et al., 2008, 2009; H. Li, 2011a, 2011b). The output is a 10-column file with read and consensus information for each base, and this information was imported into R and processed with a custom script to find SNPs between the parent strains. All SNPs used for mapping were filtered for coverage greater than 3X and less than 100X and base quality greater than 20 in each parental sample. SNPs found between parental strains JR3970 and JR3915 were used for mapping.

Bulk analysis using ceX-QTL: Alignment and quality control within Galaxy was the same for the unsorted and gut+ sequencing files as for the parental files used in map-making, described above. The resulting BAM files were processed with Naïve Variant Caller and Variant Annotator in Galaxy (Blankenberg et al., 2014; Dickins et al., 2014), and output files were imported into R and processed with a custom script. Only SNPs from the map with greater than 10X and less than 300X coverage were considered. Wildtype and MY16 parental allele counts from each sample were used to generate a G-statistic and a G-statistic' after a smoothing function for each SNP, as described in previous yeast and *C. elegans ce*X-QTL studies (Burga et al., 2019; Magwene et al., 2011). SNPs whose reads were less than 1% or greater than 99% wildtype in the unsorted control were removed, along with SNPs with equal proportions of wildtype reads in both the unsorted and gut+ samples. Unfiltered

samples referenced in Table 3, Table 4, and Figure 10 retained SNPs whose WT allele frequencies were equal in both the unsorted and gut+ populations for downstream analysis.

Picking candidates with VEP: A new variant map was made using the Naïve Variant Caller on usegalaxy.org program on the JR3915 MY16 parental BAM file against the WS235 reference genome from wormbase.org (Afgan et al., 2018; Blankenberg et al., 2014; *WormBase : Nematode Information Resource*). Any variant 5kb upstream or downstream of a top SNP was processed through Ensemble Variant Effect Predictor (VEP), and those with a moderate or high impact were considered as candidates (McLaren et al., 2016).

Comparison with SNPs from previous analysis: We looked for N2 vs. MY16 SNPs identified in a previous analysis (Torres Cleuren, 2016) that were closest to the top SNPs identified in the present analysis. RILs from the previous analysis were individually scored for the percentage of embryos producing gut granules under *skn-1 (RNAi)*, as well as individually sequenced. We analyzed whether lines carrying differential alleles (N2 or MY16) of the previously identified SNPs showed significant difference in their *skn-1 (RNAi)* phenotypes.

RNAi screen for candidate genes: RNAi clones for candidate genes were obtained from either the Vidal (Rual et al., 2004) or Ahringer (Kamath et al., 2003) libraries and transformed into HT115 *E. coli*. Bacteria containing the empty L4440 RNAi feeding vector was used as a negative control. The RNAi clones were grown overnight at 37°C in 3mL LB with 50µg/ml ampicillin and 50µg/ml tetracycline. The cultures were diluted 1:10 and grown

for 4 hours at 37°C in LB with antibiotic selection. 1mM IPTG was added to each culture before seeding 60µL onto NGM plates supplemented with 1mM IPTG and 25 µg/ml carbenicillin. JR3936 L4 dumpy worms harboring a *skn-1* point mutation were fed RNAi for 1 day before being transferred to a new RNAi plate where they were allowed to lay eggs for several hours and subsequently removed. Arrested embryos from the second RNAi plate were scored for the presence of gut granules 6-24 hours after the removal of parents. Any embryos scored more than 12 hours after the parents were removed were kept at 15°C until scoring.

2.3 Results:

Generation and quality control testing of *spe-27*$^{-/-}$ RILs

Determining precise QTLs is a function of the extent and complexity of recombination. To achieve high rates of recombination without setting up numerous tedious crosses, we devised a strategy to mate RILs in bulk for several generations. This strategy relies on generating *spe-27*$^{-/-}$ recombinant lines, creating obligate mating strains that will foster recombination without supervision. 80 *spe-27*$^{-/-}$ RILs from a cross between *spe-27*$^{-/-}$ hermaphrodites and MY16 males were selected by picking GFP+ F2 lines which can be maintained at 15°C but not at 25°C. The GFP transgene is a transcriptional reporter for the gut-specific GATA factor, ELT-2. While lines vary in brood size, they all phenocopy the *spe-27*$^{-/-}$ parent strain at 25°C (1.15 ±0.9 progeny for N2 derived parent) (Fig. 1A), and sterility can be rescued through mating (Fig. 1C). Surviving progeny from the RILs at 25°C could be selected for and could possibly take over the mating population, diminishing the extent of recombination through multiple generations. To test this, we plated 2-4 L1s and

kept them at 25°C for one week. The MY16 parent completely starves out the plate under these conditions, producing over 20,000 progeny, while all of the RILs produce vastly fewer progeny (8.89 ± 13.94 progeny, mean and standard deviation across all RILs), with 60 progeny found in the most fertile line (Fig. 1B). Furthermore, the RILs did not show any obvious growth rate differences. Taken together, these results suggest that takeover by any one RIL is unlikely over the course of this experiment.

As expected, the N2 derived *spe-27*$^{-/-}$ parent strain showed ~30% of embryos with endoderm under *skn-1 (RNAi)*, while the MY16 parent showed ~2%, both scored using gut granules. Gut granules are autofluorescent, lysosomal-like structures in intestinal cells of *C. elegans* (Clokey & Jacobson, 1986) and serve as a proxy to identify endoderm cells. It has been previously shown that gut granules correspond reliably with *elt-2p*::GFP expression (Torres Cleuren et al., 2019), and we are therefore confident in using this method to identify cells that have undergone endoderm differentiation. A selection of RILs tested show variation in their ability to form endoderm under *skn-1 (RNAi)*, ranging from ~21% to ~71% (Fig. 1D). This range in *skn-1 (RNAi)* phenotypes signals that these RILs harbor different genetic loci controlling the endoderm GRN and validates their use in this experiment.

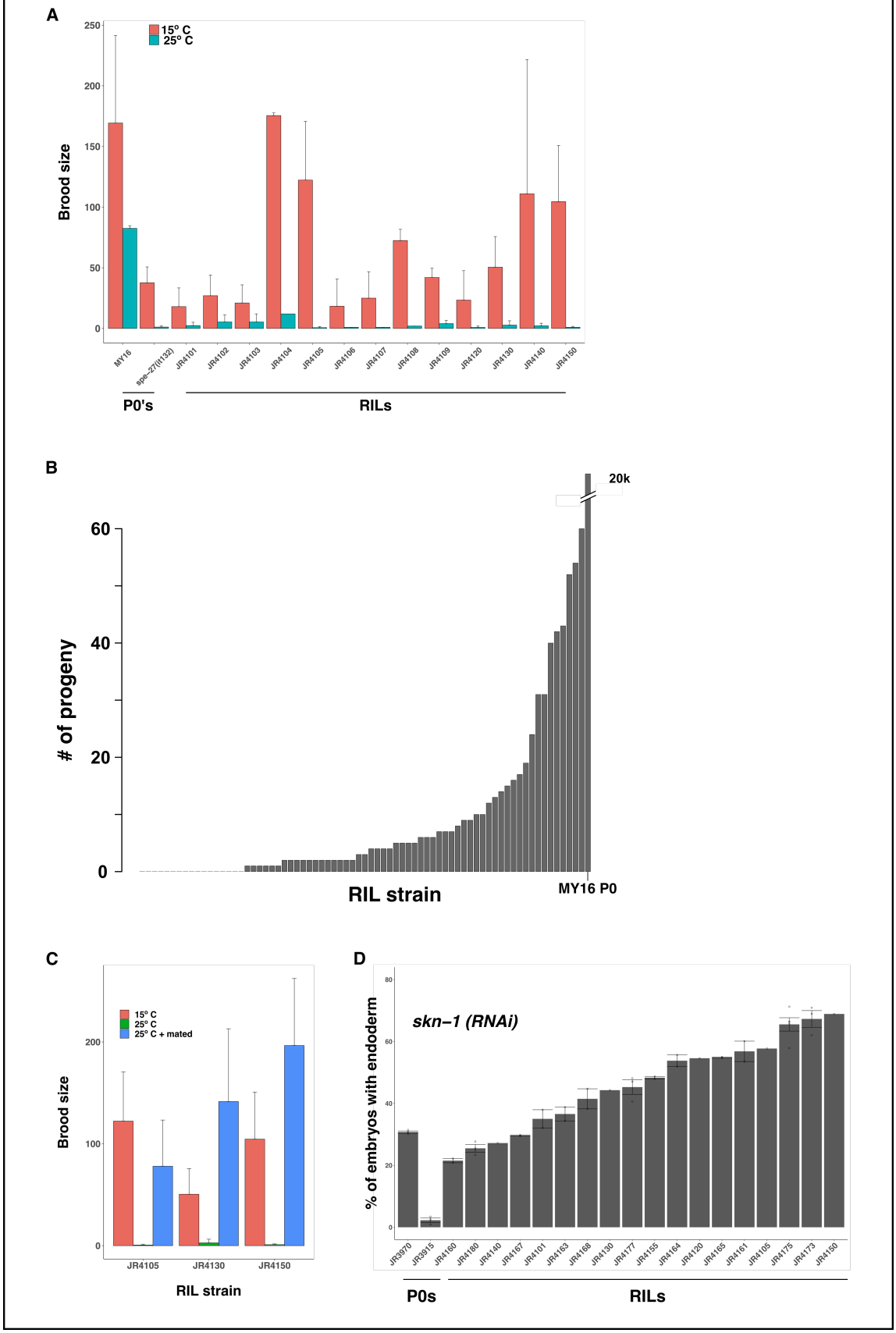
A
15° C
25° C
Brood size
250
200
150
100
50
0
MY16
spe-27(ii132)
JR4101
JR4102
JR4103
JR4104
JR4105
JR4106
JR4107
JR4108
JR4109
JR4120
JR4130
JR4140
JR4150
P0's
RILs
B
20k
60
40
20
0
# of progeny
RIL strain
MY16 P0
C
15° C
25° C
25° C + mated
Brood size
200
100
0
JR4105
JR4130
JR4150
RIL strain
D
skn–1 (RNAi)
% of embryos with endoderm
80
60
40
20
0
JR3970
JR3915
JR4160
JR4180
JR4140
JR4167
JR4101
JR4153
JR4168
JR4130
JR4177
JR4155
JR4164
JR4120
JR4165
JR4161
JR4105
JR4175
JR4173
JR4150
P0s
RILs

Figure 1 Quality control testing of RILs

A) Quantification of brood size of individual worms from parental strains and a subsection of RILs at both 25°C and 15°C. B) Bar graph showing the number of progeny found on individual plates after 2-4 L1s were placed at 25°C for 1 week. C) Quantification of brood size for 3 RILs at 15°C, 25°C, and 25°C with males from the same strain. D) Bar graph showing the percent of embryos showing gut granules after *skn-1 (RNAi)* in parental strains plus a subset of RILs.

Generation and sorting of recombinant embryos

Efficient QTL mapping is dependent on extensive recombination between the N2 and MY16 derived chromosomes. In *C. elegans* on average there is a single crossover per chromosome per gamete (Barnes et al., 1995). To increase the number and distribution of recombination events along the length of the chromosome males and hermaphrodites from all RIL strains were plated together and allowed to mate for 16 generations. Self-fertilized hermaphrodites only produce ~0.1% male progeny, but mated hermaphrodites produce 50% males (Altun & Hall, 2006). The male population was maintained around 50% throughout the experiment, signifying that mating has been robust at each generation (Fig. 2A). Additionally, the *elt-2p::*GFP transgene, introduced from the N2 hermaphrodite parent as a genomic insert and from the MY16 male parent as an extrachromosomal array, showed high expression in the F16 bulk population (Fig. 2B). The maintained expression of *elt-2p*::GFP is essential in order to sort embryos for endoderm differentiation based on a fluorescent proxy.

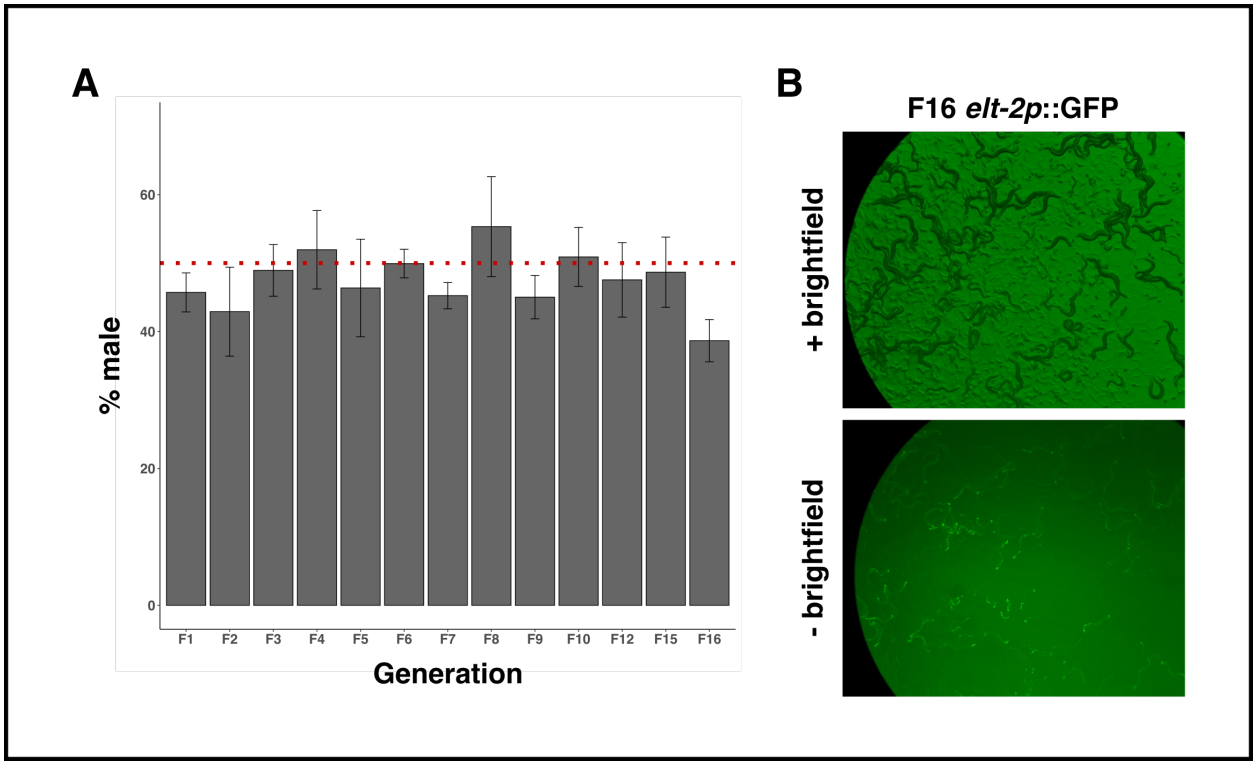

Figure 2 Male and GFP maintenance during bulk mating

A) Bar graph showing the percent of males found at several generations throughout the bulk mating. The red dotted line shows 50%, or the expected percent of males with successful matings. B) Adult worms from the F16 population visualized with a GFP filter, both with the brightfield light and without.

Trial runs on the FACS machine showed that it was possible to distinguish between populations of embryos based on their *elt-2p::*GFP expression following *skn-1 (RNAi)* (Fig. 3A). F16 parents were fed *skn-1 (RNAi)*, and the resulting arrested embryos were sorted for *elt-2p::*GFP expression with a FACS machine, alongside controls (Fig. 3B). Over the course of 5 different days of sorting, over 64,000 embryos were collected according to the FACS-analysis software. This represents ~3% of the embryo population (Fig. 3C). This was confirmed by visual inspection of an unsorted population which shows only a small fraction

of GFP+ embryos. In contrast, the FACS-sorted population shows a vast enrichment of *elt-2p*::GFP+ embryos, pointing to a successful sort and collection of gut+ embryos (Fig. 3D).

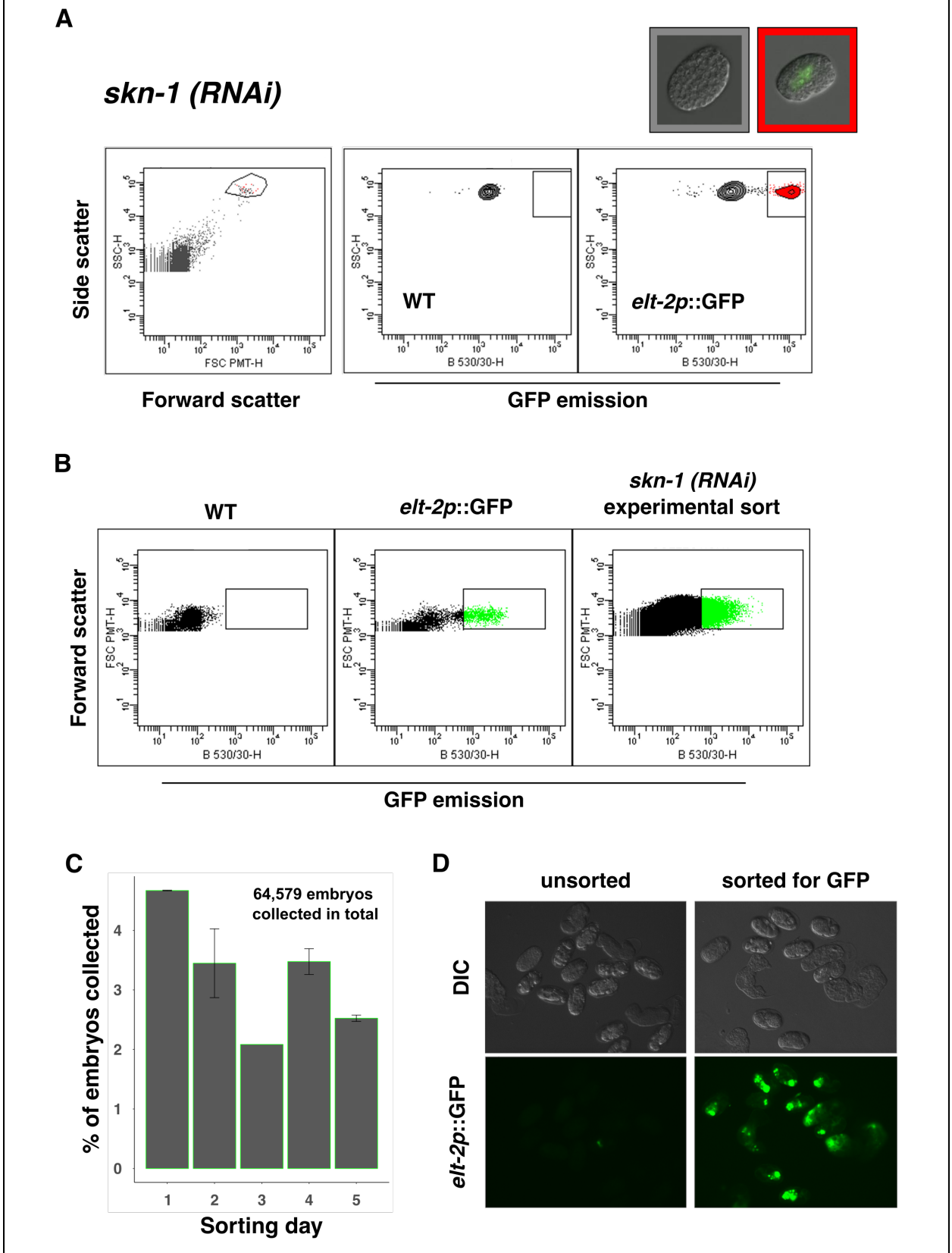
A
skn-1 (RNAi)
Side scatter
SSC-H
FSC PMT-H
WT
elt-2p::GFP
B 530/30-H
Forward scatter
GFP emission
B
WT
elt-2p::GFP
skn-1 (RNAi)
experimental sort
Forward scatter
FSC PMT-H
B 530/30-H
GFP emission
C
64,579 embryos
collected in total
% of embryos collected
0
1
2
3
4
1
2
3
4
5
Sorting day
D
unsorted
sorted for GFP
DIC
elt-2p::GFP

Figure 3 FACS collection of GFP+ embryos

A) The far-right plot shows the gate used to identify embryos in the Aria FACS software. Y-axis shows side scatter, and X-axis shows forward scatter. Other particles coming through the FACS stream represent bacteria and debris. The middle plot shows a single population of GFP- embryos after *skn-1(RNAi)* in a strain without the *elt-2p*::GFP fluorophore as a control, and the far-left plot shows two populations of embryos after *skn-1 (RNAi)* in a strain with the *elt-2p*::GFP fluorophore (GFP- and GFP+). Y-axis shows side scatter, and X-axis shows GFP emission for both of these plots. Representative images of embryos from the gray-colored population and the red-colored population are shown. B) Y-axis shows forward scatter and X-axis shows GFP emission for all three plots. The far-right plot shows a strain without the GFP fluorophore and guided where we should put the gate to collect GFP+ embryos (to the right of the single population). The middle plot shows an *elt-2p*::GFP strain, and we see embryos in the GFP+ gate. The far-left plot shows one of the experimental sorts of the bulk mating population, collecting GFP+ embryos after *skn-1 (RNAi)*. C) Bar graph showing the percent of GFP+ embryos out of all embryos identified by the sort across 5 different sorting days. The number of embryos in the GFP+ gate and the total number of embryos sorted were quantified in the Aria FACS software. D) Representative images of embryos from the bulk mating population after their parents were fed *skn-1 (RNAi).* DIC and GFP images are shown for both the unsorted population and embryos collected in the GFP+ gate.

Sequencing and identification of SNPs for mapping

To identify SNPs between the parent strains that can be used for mapping, whole genome sequencing (WGS) was performed for both parent strains along with the unsorted

and gut+ sorted populations. We pooled our samples such that the unsorted and gut+ groups would receive higher coverage, and after quality control filtering, we find an average of 8.6-fold coverage in the N2 parent, 9.3-fold coverage in the MY16 parent, 53.6-fold coverage in the unsorted sample, and 31.4-fold coverage in the gut+ sample (Fig. 4A). Filtering did not affect the average coverage, but it did reduce the variance in coverage in each sample. Filtering also reduced the percent of the genome covered in both of the parental strains to ~90% in the N2-type parent and ~71% in the MY16 parent (Fig. 4B). This is largely due to the low coverage in the parents and our requirement that each base considered for making the SNP map has greater than 3X coverage in each parent.

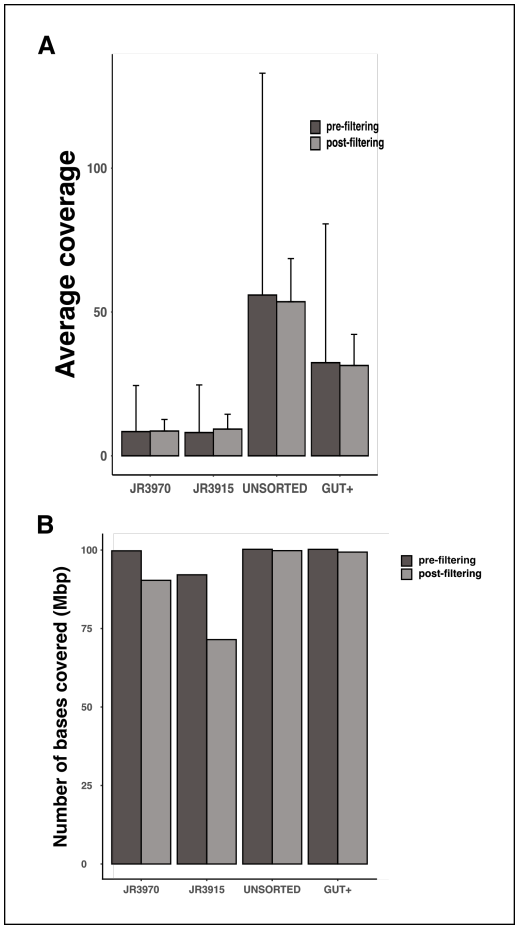

Figure 4 Coverage of sequenced samples

A) Average coverage per base from WGS of both parental strains plus the unsorted and gut+ (GFP+) bulks, both before and after filtering (see *Methods and Materials*). B) Number of bases covered out of the entire genome from all samples both before and after filtering.

To be used as a SNP for mapping, a base has to have been called in both parents, and it has to be different in both parents. In our samples, this yielded 50,306 SNPs across the genome (Fig. 5), or roughly one SNP every 2Kbp. Previous sequencing analysis of RILs from nearly identical parental strains to the ones used in the present study yielded a total of 171,713 SNPs (Torres Cleuren, 2016). One reason for the discrepancy in the number of SNPs is because we did not achieve full coverage of the parental genomes across the 100Mbp genome. Due to low fold-coverage in the parents, some SNPs could not be assessed as they were absent from the data in either one of both of the parents. Intriguingly, when compared to the publicly available MY16 variant map from the *C. elegans* Natural Diversity Resource (Cook et al., 2017), we see 2,012,645 potential SNPs that match the WS235 reference in the MY16 parent. These SNPs are scattered throughout the genome and don't show evidence of strain contamination through mating. One reason for this discrepancy comes from the use of two different reference genomes: CeNDR uses WS263 and we have been using WS235, but that cannot account for these extreme differences. Looking into these sites in the CeNDR variant dataset reveals that some of these sites appear heterozygous, but 1,980,856 homozygous SNPs remain that are not present in our sequencing data.

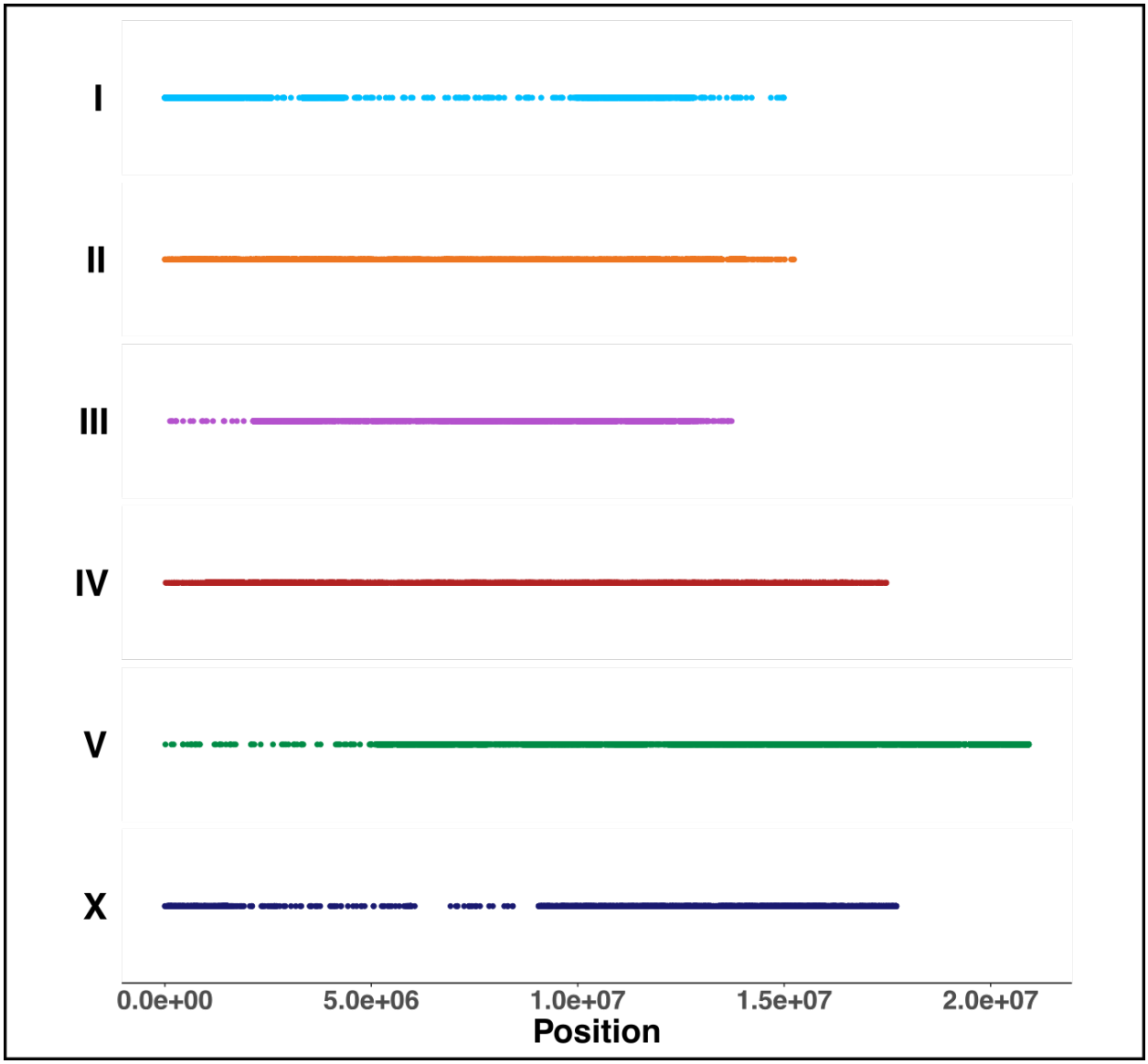

Figure 5 SNPs found between parent strains

Graphic showing the location of all SNPs found between the N2 (JR3970) and MY16 (JR3915) parent strains on all six chromosomes. SNPs are represented as a single dot. Regions with high SNP density appears as continuous lines.

Selection on Chr IV and X in RIL population

First, we looked at the difference in WT allele frequency at each SNP between the gut+ and unsorted samples (Fig. 6). This analysis was useful in showing regions under high selection in the RIL generation and bulk mating. We see that a region on chromosome IV is represented similarly in both populations, and this coincides with the location of *spe-27* at 5.86 Mbp. Additionally, a long stretch on X was similar in both populations, and the sequencing data shows that this is the location of the *elt-2p::*GFP insert (~5.9Mbp).

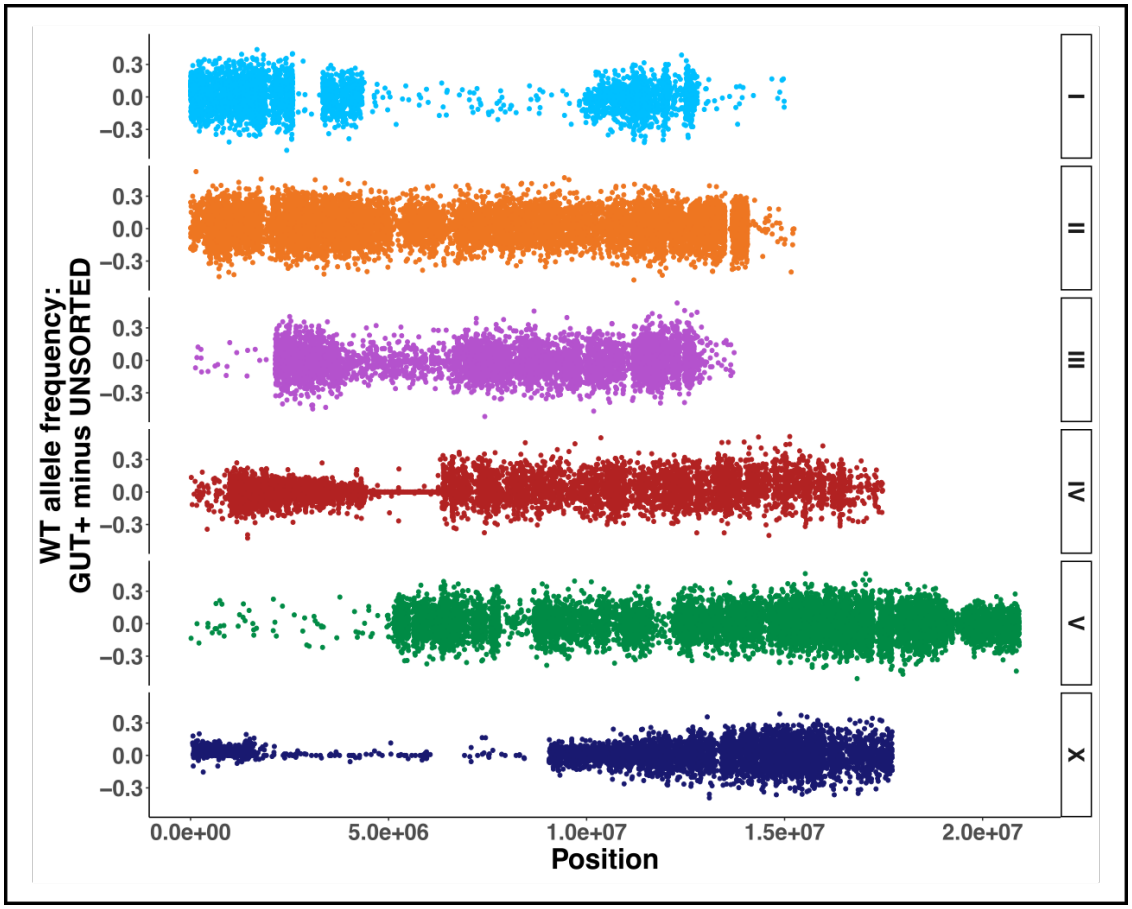

Figure 6 WT allele frequency differences

Graphic showing the difference in WT allele frequency between gut+ and unsorted populations at each SNP across all six chromosomes. Regions showing no difference give continuous SNPs close to 0.0 and signify a selection in the bulk mating population.

Top SNPS identified using *ce*X-QTL

To identify the likelihood of selection for any given SNP, we filtered out sites where the WT allele frequency in the control, unsorted population was less than 1% (3 sites) or greater than 99% (1565 sites), as those SNPs are considered too uniform in the population to be useful for QTL analysis. We used a method first shown to work on bulk sequencing populations in yeast and later further developed for *C. elegans*. This method generates a G-statistic' for each SNP, which takes into consideration the number of WT and MY16 reads at each base in both the unsorted and gut+ samples. This method, titled *ce*X-QTL uses a log normal distribution of the G-statistic' to assign a p-value. To assess the normality of the G-statistic', we looked at -log(G-statistic') in a Q-Q (quantile-quantile) plot, and we noticed an outlier population (Fig. 7A). This population was represented by SNPs where the WT allele frequency was equal in both samples tested, and these are not SNPs we deem useful for analysis and were filtered out (Fig. 7B). After these filtering steps, we kept 48,491 SNPs, or roughly one SNP per every 2Kb.

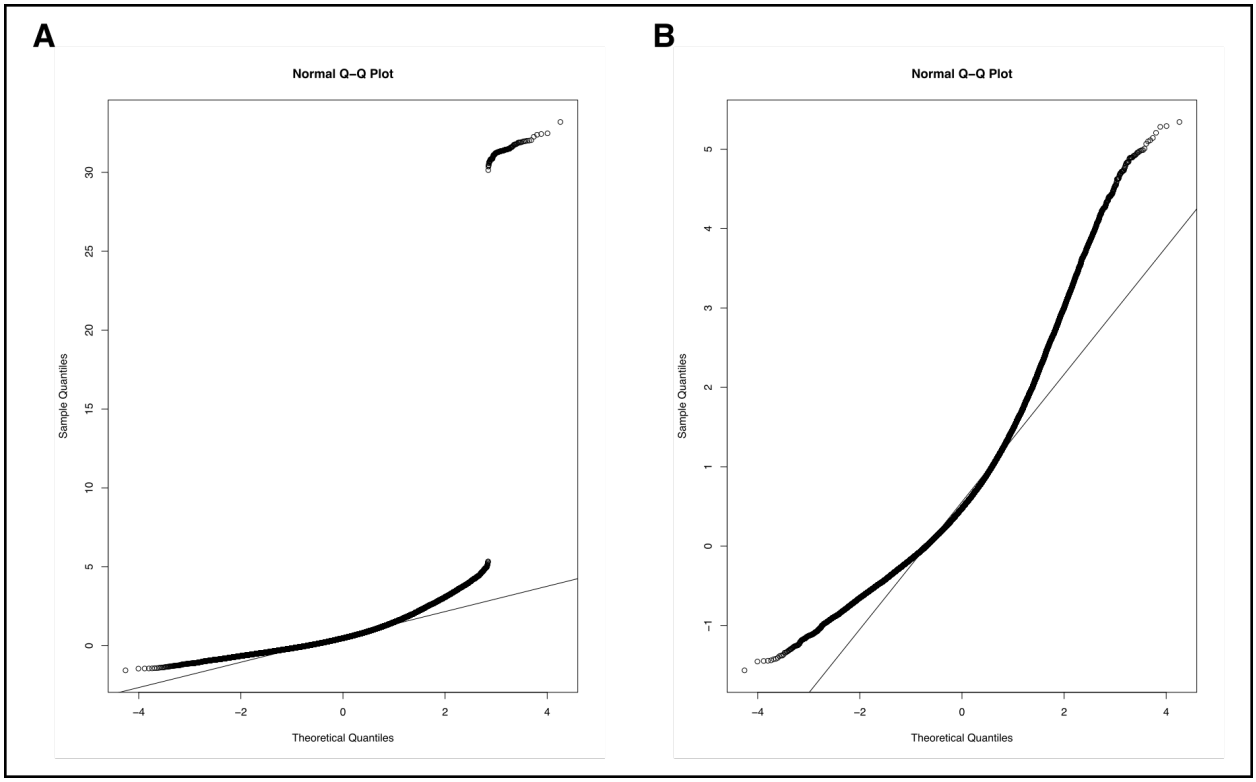

Figure 7 Q-Q plots addressing the normalcy of G-statistic'

A) Q-Q plot with reference line modeled in R showing the -log(G-statistic') of each SNP after filtering for any points where G-statistic' = 0. B) The same Q-Q plot with reference line after further filtering out SNPs where the frequency of the WT allele in the unsorted population was equal to the WT allele frequency in the gut+ population.

Plots of the G-statistic' (Fig. 8) and -log(p-value) (Fig. 9) do reveal some peaks in the data. To view the top SNPs, we collected all SNPs with -log(p-value) > 2. In this collection, we found 18 SNPs, 11 of which were in 11 different genes (Table 1 & Table 2).

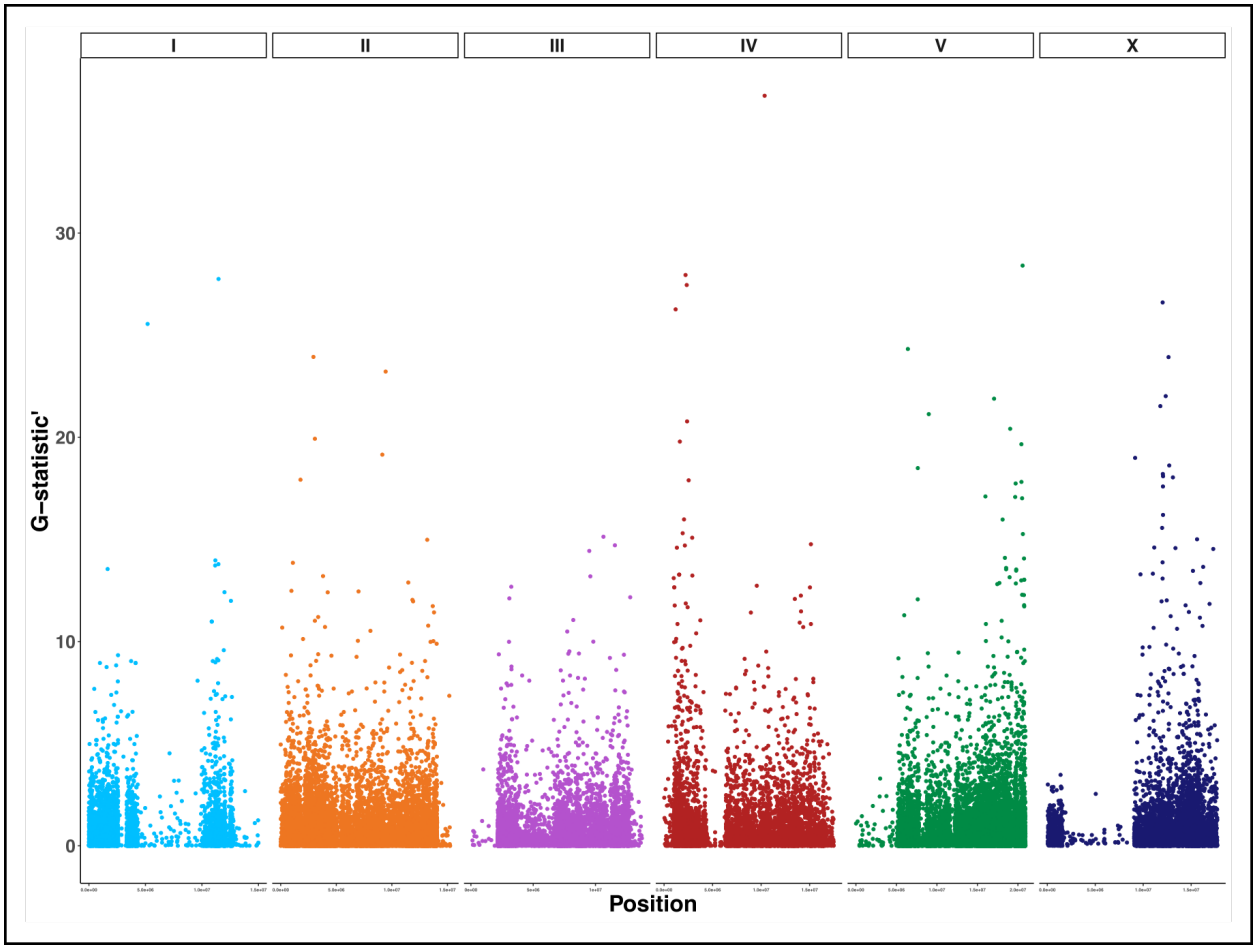

Figure 8 G-statistic' for each SNP

Graphic showing the G-statistic' calculated for each SNP on all chromosomes.

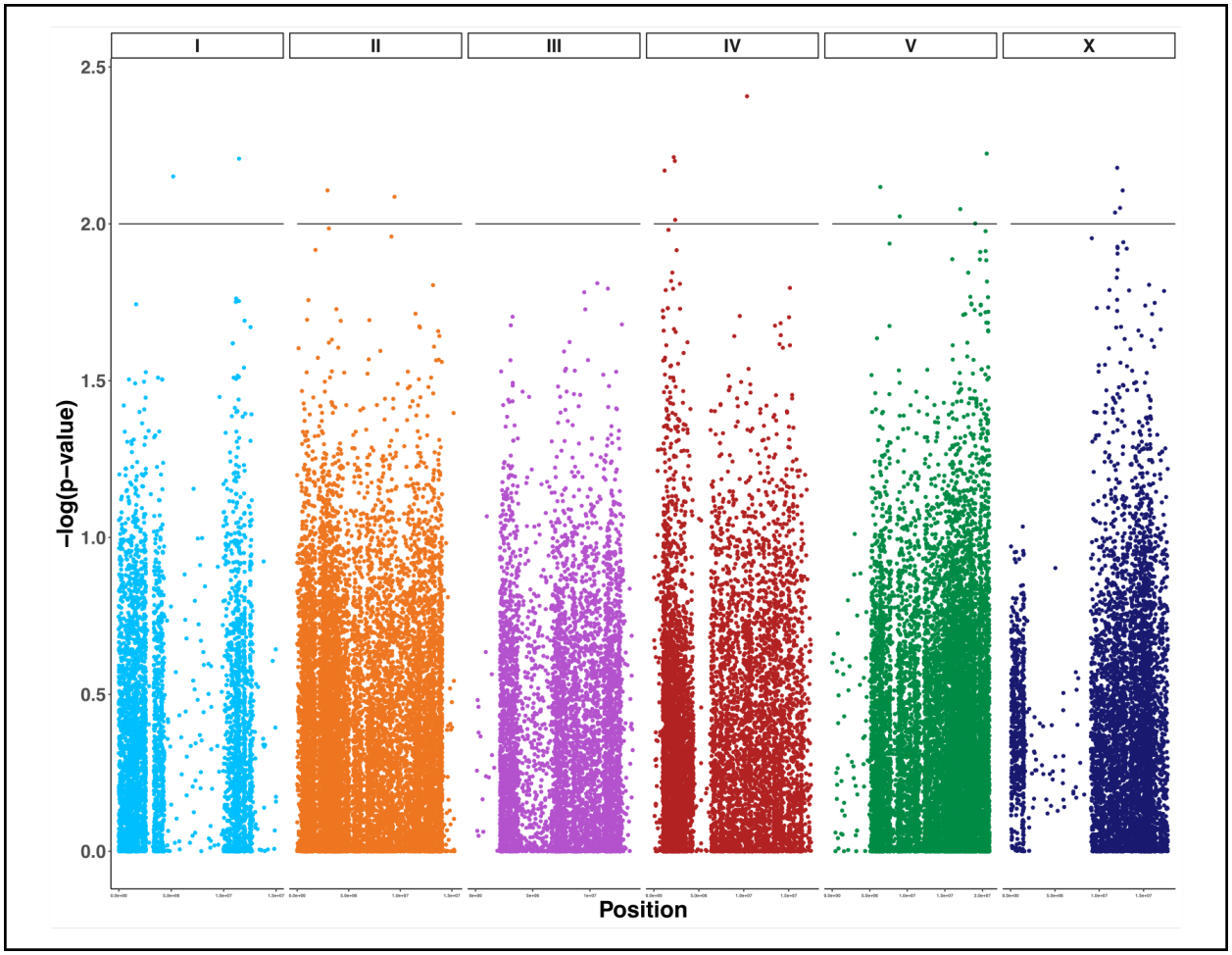

Figure 9 -log(p-value) for each SNP

Graphic showing the -log(p-value) of each SNP calculated using the log-normal distribution of G-statistics'.

Repeating the analysis without filtering any SNPs yields 48 SNPs with -log(p-value) > 2, 12 of which are also found in the top filtered SNPs (Table 3). After applying either a false discovery rate or Bonferroni correction, none of these SNPs are significant. We believe that the significance of any one loci was diluted because multiple loci across the genome control this phenotype. The multigenic nature of this phenotype is visible in the range of *skn-1 (RNAi)* phenotypes seen in the RILs, as we do not simply see one group with

the N2-like phenotype and one group with the MY16-like phenotype (Fig. 1D). Previous work also alludes to multiple genes and complex gene interactions contributing to the diversity of *skn-1 (RNAi)* phenotypes (Torres Cleuren et al., 2019), but these sets of top SNPs, both from the filtered and unfiltered data, provide a good starting point around which to search for candidate genes that may be contributing to the difference in phenotypes from the parent strains.

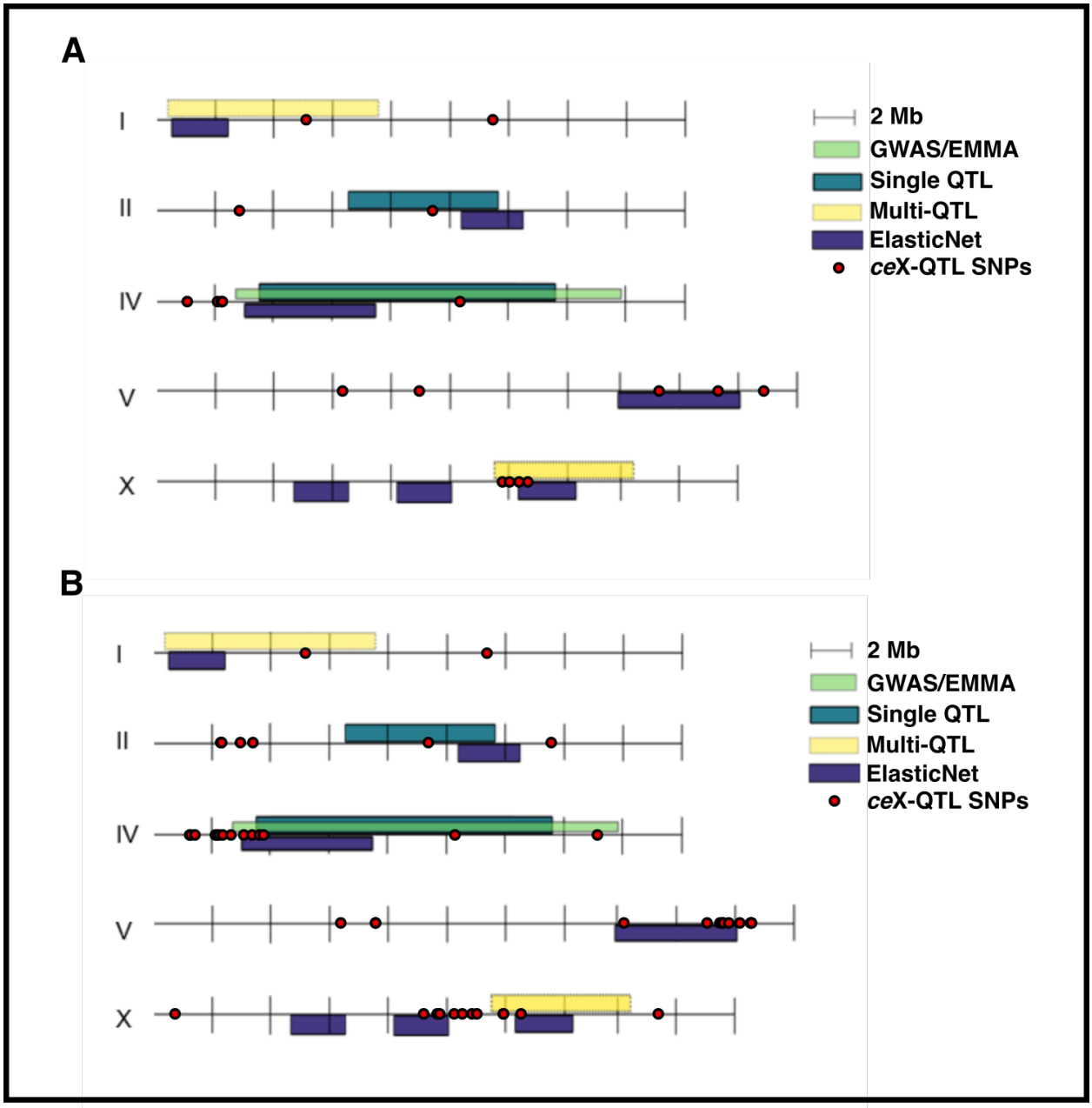

Figure 10 Previously identified QTLs and top SNPs from present analysis

A) Graphic showing the location of previously identified QTLs (Torres Cleuren, 2016, unpublished data from Melissa Alcorn) along with the location of top SNPs identified from the filtered dataset (SNPs whose WT allele frequency was the same in both unsorted and gut+ populations were removed for downstream analysis). B) The same graphic as A) except top SNPs from the unfiltered dataset are shown.

Comparison with previous RIL results

Importantly, both sets of top SNPs (from filtered and unfiltered data), largely land in QTLs revealed from previous RIL and wild isolate analysis (Fig. 10) (Torres Cleuren, 2016). Earlier data as well as the analysis presented here both find QTLs on all chromosomes except for chromosome III. We next used phenotypic and sequencing data from the initial RIL set to assess if any regions found in the present analysis show significant differences in *skn-1 (RNAi)* endoderm formation. The initial RILs were all individually sequenced and phenotyped (Torres Cleuren, 2016), so we identified those SNPs that were closest to the SNPs found in the present analysis. We then looked at the phenotypes of individual lines carrying either the N2/wildtype or MY16 variant at those locations. We found two SNPs that show significantly higher numbers of *skn-1 (RNAi)* embryos that form endoderm if they have the wildtype allele versus the MY16 allele, one on chromosome II (9424346) and one on chromosome IV (3065698) (Fig. 11).

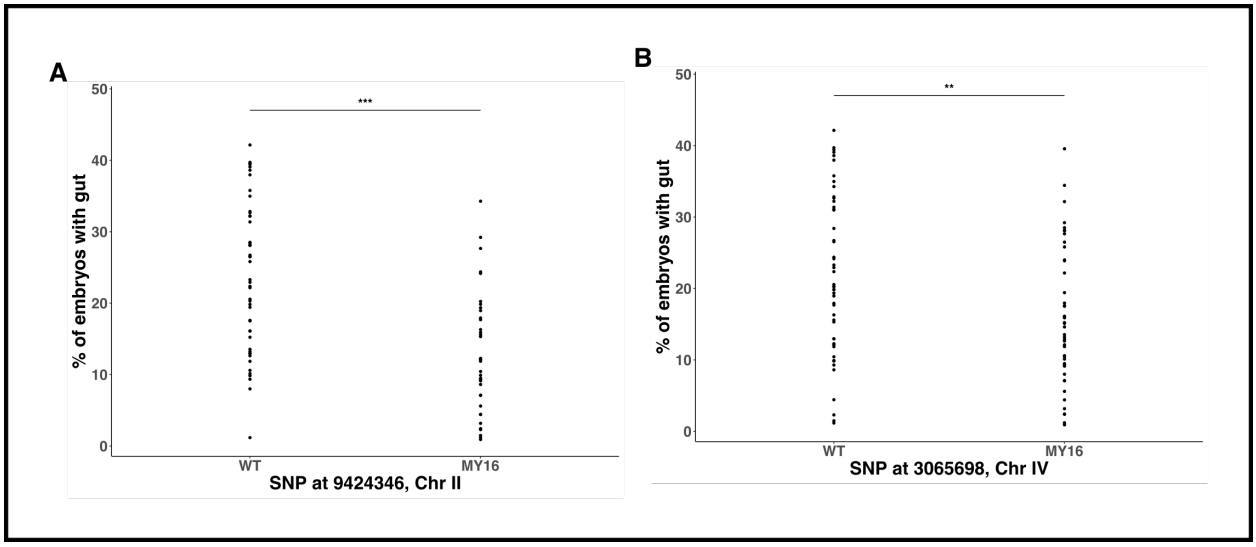

Figure 11 *skn-1 (RNAi)* phenotypes of individual RIL strains

A) Using phenotypic and sequencing data from the original N2/MY16 RILs (Torres Cleuren, 2016), the percent of embryos showing endoderm under *skn-1 (RNAi)* in individual lines is shown on the Y-axis. The X-axis shows whether each strain had the WT or MY16 version of the SNP at position 9424346 on chromosome II. B) The same as A) except strains are divided along the X-axis as to whether they had the WT or MY16 version of the SNP at position 3065698 on chromosome IV. In both graphs, significance was assessed using a Student's t-test.

Genes of interest around top SNPs

To identify the genes around the top SNPs that may be causal, we made a new variant map of our MY16 parent strain using Naïve Variant Caller in Galaxy (Afgan et al., 2018; Blankenberg et al., 2014) in order to gather as many potential candidates as possible. We looked 5Kb upstream and 5Kb downstream of each of the top SNPs (from both filtered and unfiltered datasets) and gathered all variants shown in the MY16 parent strain. We then ran those variants through Variant Effect Predictor (McLaren et al., 2016) and kept all variants

showing a potential moderate or high impact on gene function. These effects include missense, nonsense, and splice region mutations. This analysis leaves us with 53 candidate genes to test using either RNAi or existing mutant strains (Table 4). Using RNAi clones available in lab, we performed a candidate screen of the genes shown in Table 4, knocking them down in worms carrying a point mutation in *skn-1* (Fig. 12). Of the 41 candidate genes tested so far, none have shown significant deviation from the control.

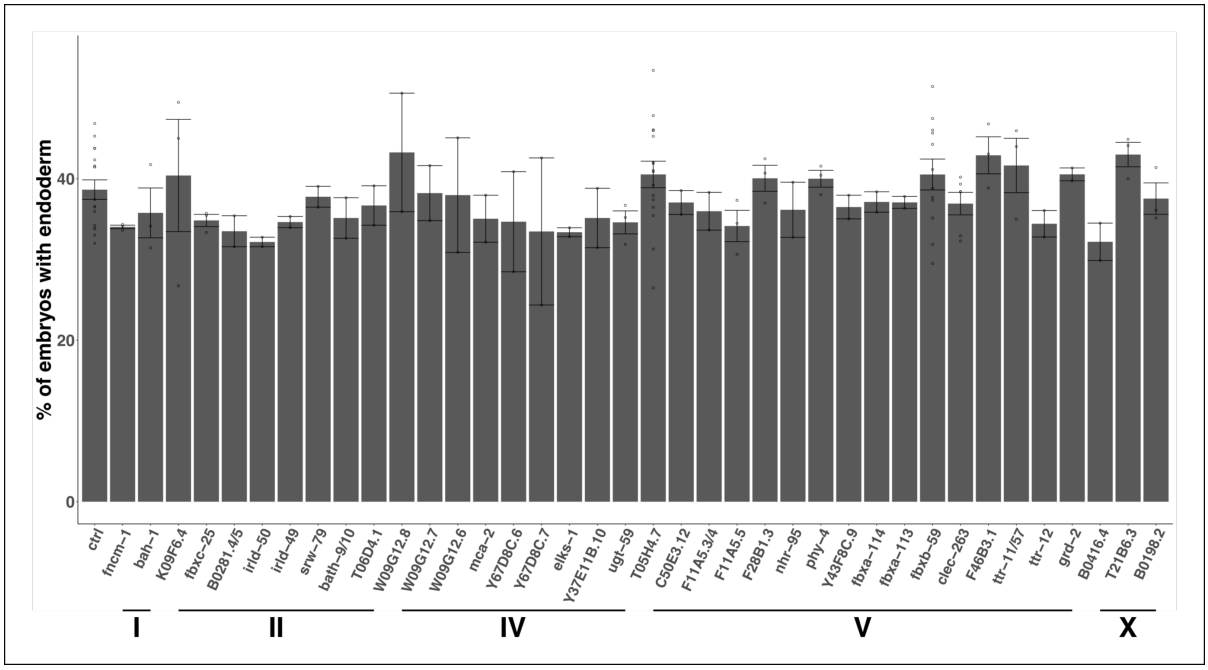

Figure 12 Candidate RNAi screen

RNAi was fed to *skn-1(zu67)* mutant worms and the percent of arrested embryos with gut granules in the next generation was assessed. Out of 37 RNAi constructs screened, none show a significant deviation from the L4440 control. Four constructs (targeting *B0281.4/5*, *bath-9/10*, *F11A5.3/4*, and *ttr-11/57*) target two candidate genes, so 41 genes have been screened using this method.

2.4 Discussion and Conclusions

This method, along with earlier work (Torres Cleuren, 2016), identified regions on all chromosomes except chromosome III that may be contributing to differences in *skn-1 (RNAi)* phenotypes. Top SNPs from the present analysis are found in every previously identified QTL using GWAS, single-QTL, multi-QTL, and machine learning analyses, with the exception of three QTLs found through machine learning/ElasticNet on chromosomes I, II, and X (Torres Cleuren, 2016, unpublished data from Melissa Alcorn). Additionally, one QTL that was only identified by ElasticNet on chromosome V coincides with a cluster of top SNPs presented here. This overlap shows the strength of the *ce*X-QTL method, as this one experiment was able to identity different regions picked up by numerous QTL analyses. It is important to note that none of the SNPs identified here are significant after applying either a false discovery rate or Bonferroni correction, but due to the large coincidence with previously found QTLs, we believe the regions around the top SNPs represent real QTLs where causal genes can be found. Due to the multigenic nature of the phenotype explored, it appears as though many genes were selected for in the FACS sort, diluting the impact of any one SNP. Contrasted with a single-gene trait with a low selection (0.1% selection) (Burga et al., 2019), our analysis selected for ~3% of the population and gave us low peaks across the genome. Previously RIL analysis did produce significant QTLs (Torres Cleuren, 2016), but it is possible that the method presented here was successful in ensuring more recombination and breaking up those significant QTLs into smaller regions. In generating smaller QTLs, we may have selected for more regions, again diluting the significance of any one, larger QTL. While no SNPs were technically significant, we believe they regions they point to are real and will continue screening genes within these regions using RNAi and mutant strains.

Targeting SNPs that result in changes to the protein of a given gene for our candidate screen has so far not yielded any positive hits, as RNAi knockdown of these candidate genes do not change the permissibility of endoderm in *skn-1* mutants. It is possible that the variants causing the diversity of *skn-1 (RNAi)* phenotypes in our RILs lie in regulatory regions as opposed to exons, changing the expression patterns of different genes. In that case, our candidate list based on protein-changing variants would not point us to the correct genes. Further analysis of variants in known regulatory regions may show more promising results. Additionally, due to the multigenic nature of the trait and previously demonstrated interactions between multiple QTLs (Torres Cleuren et al., 2019), it is possible that targeting two or more genes in combination may be necessary to see an effect.

Another interesting note from the analysis is the abundance of SNPs from the CeNDR-MY16 variant map that were found to be wildtype in the MY16 parent used in this study (Cook et al., 2017). These SNPs are spread fairly uniformly across the genome and do not represent an introgression event from strain contamination. It is possible that strain divergence between MY16 samples can account for some of these, but as close to 2 million SNPs are seen as wildtype in our samples, we do not feel that is an adequate explanation. We are confident that our MY16 strain is truly MY16 due to its consistency in *skn-1 (RNAi)* phenotype, so this study is still valid in examining the genetic basis behind its high SKN-1 requirement in the endoderm GRN. Further investigation is warranted to examine the basis of the differences seen between our MY16 and the published sequence from the CeNDR database.

For further genomic analysis to uncover the strongest QTLs, it is possible to redo the experiment without selecting for individual *spe-27*$^{-/-}$ lines at the F2 by using *fog-2*$^{-/-}$ parental

lines. To this end, we have introgressed *fog-2(q71)* into the MY16 parental line to be used for future experiments. One caveat to consider is that *fog-2* is located at 20.2Mbp on chromosome V, which puts it just to the right of the ElasticNet QTL and in the middle of some top SNPs identified in the present study. As such, the suggested experiment may mask the QTL on chromosome V, but it could be helpful to more narrowly define the large QTL on chromosome IV along with others not on the right arm of chromosome V. Additionally, one can redo the initial RIL QTL analysis by making lines out of the last generation from the present experiment and phenotyping and sequencing as many as are feasible possible. With enough sequencing resources, we should see smaller and more targeted QTLs.

The SNP with the highest G-statistic' and lowest p-value landed in the middle of the large QTL identified by GWAS/EMMA on chromosome IV at around 10Mb, and we also see a cluster of top SNPs upstream at the left edge of that QTL around 2Mb. *spe-27* lies within that QTL, and since *spe-27(it32)* was selected for all of the lines used, we see a possible disruption of that region in the analysis. Without the selection, we may have seen a more prominent peak on chromosome IV pointing to a region between 2Mb and 10Mb. Even with the *spe-27* selection, we still see variation in the *skn-1 (RNAi)* phenotype among different RILs (Fig. 1D), suggesting that we did not select out the only QTL. Furthermore, as we know this is a multigenic trait, we may have strengthened our ability to pick up other regions that may have a smaller effect size even if we selected against a main QTL. None of the RILs tested so far show significantly less endoderm development than the wildtype parent under *skn-1 (RNAi)*, so it is possible that we selected against a QTL on chromosome IV that greatly contributes to endoderm suppression, although only about one fifth of RILs have been phenotyped. Even so, we see lines with enhanced ability to make *skn-1*-independent

endoderm (over 60% of embryos with gut, Fig. 1D), meaning this experiment still provides us with the opportunity to study effectors of the endoderm pathway.

Overall, this method showed promising results as the top SNPs identified showed great overlap with previously identified QTLs. As such, we feel this method would be useful to explore other GRNs and can be achieved more quickly than traditional methods, especially if using parental strains that both show the obligate-mating mutation (i.e. *spe-27(it32)* or *fog-2(q71)*) so that singling out individual F2 lines can be avoided. The FACS sort proved to be a powerful selection for phenotypes in the embryo that can be visualized with a fluorescent marker (Fig. 3D). The *ce*X-QTL method using fluorescent markers for late differentiation factors in other germ layers or specific organs could prove to be a powerful tool in exploring many different developmental GRNs in the embryo. Furthermore, coupled with a fluorescence-based worm sorter such as a COPAS machine (Pulak, 2006) opens this method up to phenotypes seen in larvae and adult *C. elegans*. This tool is useful for looking at multigenic traits, but it proves to be more powerful for rapidly exploring monogenic traits (Burga et al., 2019).

The present study further confirmed the highly multigenic nature of the endoderm GRN. QTLs controlling differential requirements for endoderm formation in the natural *C. elegans* population are found throughout the genome and show high dependency on mostly SKN-1-based or Wnt-based signaling systems (Torres Cleuren et al., 2019). The SKN-1 transcription factor was discovered as a master regulator of gut in the N2 reference strain, but if initial GRN dissection had been done using a wild isolate such as CB4852, found only ~285 miles SE of Bristol, UK, in Orsay, France (Cook et al., 2017), we would likely have a different understanding of what is considered the canonical initiation pathway, as CB4852

has a much more lenient requirement for SKN-1 for endoderm formation (Torres Cleuren et al., 2019). We see the effects of this cryptic variation only under knockdown or mutant conditions, as all isolates produce a 20-cell gut upon hatching. The fact that haplotypes within the same species can show large variation in their finely-tuned GRNs to all produce the same, functionally identical organ (the gut) suggests that great redundancy in the endoderm GRN has enabled cryptic evolution while still supporting robust development (Ewe et al., 2020). The logic follows that high redundancy in factors activating gut-specification allows for the buildup of cryptic mutations in one or more activating pathways, eventually selecting for reliance on different inputs, which further allows degradation of less-crucial inputs. The redundancy seen in the endoderm GRN has enabled plasticity of this network within the species and suggests that other wide-range phenotypes within a crucial organ or behavior are underlaid by redundant GRNs.

Table 1- Top SNPs from filtered data (including gene and impact)

Chr	Position	WT	MY16	-log(P-value)	Gene	Effect
I	5190791	C	A	2.150693547	fncm-1	intron variant
I	11466719	G	A	2.207521909	bah-1	intron variant
II	2944836	T	C	2.106451074	irld-49	splice region variant
II	9440171	A	T	2.085982322		
IV	1192440	C	T	2.16954452	mdt-20	intron variant
IV	2212161	G	A	2.212361395		
IV	2336461	T	C	2.200084071	ppgn-1	intron variant
IV	2372024	G	A	2.012365748	tftc-1	intron variant

IV	10361150	T	C	2.40648441	ugt-59	synonymous variant
V	6413543	A	G	2.117286378		
V	8994119	A	G	2.023549412		
V	17055619	C	T	2.046734768	F28B1.3	non-coding exon variant
V	19041273	G	T	2.000835363	asp-16	intron variant
V	20577056	C	A	2.22372447	clec-263	intron variant
X	11789762	G	A	2.035656529		
X	12028925	C	A	2.178332855	B0198.2	5' UTR
X	12349585	A	T	2.050529715		
X	12643973	A	G	2.10616683		

Color key for Tables 2, 3, and 4

FILTERED AND UNFILTERED
UNIQUE TO UNFILTERED
UNIQUE TO FILTERED

Table 2- Top SNPs from filtered data

CHR	POS	REF wb235	JR3970	JR3915	WT frequency in unsorted	WT frequency in gut	G-statistic (smoothed)	pval	neglogP
I	5190791	C	C	A	0.956521739	0.702702703	25.550931	0.00706816	2.15069355
I	11466719	G	G	A	0.965517241	0.5625	27.7531746	0.00620123	2.20752191
II	2944836	T	T	C	0.972972973	0.730769231	23.93694885	0.00782616	2.10645107
II	9440171	A	A	T	0.320754717	0.791666667	23.21885718	0.00820385	2.08598232
IV	1192440	C	C	T	0.981481481	0.78125	26.26502852	0.00676792	2.16954452
IV	2212161	G	G	A	0.98245614	0.766666667	27.9476601	0.00613251	2.21236139
IV	2336461	T	T	C	0.981818182	0.766666667	27.45642702	0.00630835	2.20008407
IV	2372024	G	G	A	0.980392157	0.76	20.77921053	0.00971928	2.01236575
IV	10361150	T	T	C	0.11627907	0.617647059	36.72515379	0.00392207	2.40648441
V	6413543	A	A	G	0.973684211	0.730769231	24.32432432	0.00763332	2.11728638
V	8994119	A	A	G	0.79245283	0.408163265	21.1358119	0.00947219	2.02354941
V	17055619	C	C	T	0.962962963	0.681818182	21.89089482	0.00897977	2.04673477
V	19041273	G	G	T	0.96	0.75	20.41666667	0.00998078	2.00083536
V	20577056	C	C	A	0.983606557	0.774193548	28.40869565	0.00597414	2.22372447
X	11789762	G	G	A	0.974358974	0.78125	21.52742772	0.00921178	2.03565653
X	12028925	C	C	A	0.975609756	0.708333333	26.60346154	0.00663235	2.17833286
X	12349585	A	A	T	0.974358974	0.774193548	22.01654135	0.00890165	2.05052972
X	12643973	A	A	G	0.970588235	0.708333333	23.92685475	0.00783129	2.10616683

Table 3- Top SNPs from unfiltered data

CHR	POS	REF wb235	JR3970	JR3915	WT frequency in unsorted	WT frequency in gut	G-statistic (smoothed)	pval	neglogP
I	5190791	C	C	A	0.956521739	0.702702703	25.550931	0.00816793	2.08788789
I	11466719	G	G	A	0.965517241	0.5625	27.7531746	0.00720932	2.14210574
II	2265018	T	T	G	0.99	0.733333333	38.00087835	0.00440953	2.35560795
II	2282105	C	C	T	0.99	0.761904762	48.3322314	0.0029719	2.52696632
II	2297532	A	A	C	0.99	0.72	91.28099174	0.0009691	3.0136301
II	2944836	T	T	C	0.972972973	0.730769231	23.93694885	0.00900183	2.04566923
II	3377615	T	T	G	0.99	0.8	48.21974462	0.00298351	2.52527274
II	9440171	A	A	T	0.320754717	0.791666667	23.21885718	0.00941599	2.02613417
II	13688709	C	C	T	0.99	0.818181818	34.90414786	0.00505014	2.29669624
IV	1192440	C	C	T	0.981481481	0.78125	26.26502852	0.00783657	2.10587426
IV	1231194	T	T	C	0.99	0.84	25.70062948	0.0080967	2.0916922
IV	1375660	T	T	C	0.99	0.838709677	35.44363636	0.00492876	2.30726245
IV	2082647	C	C	T	0.99	0.866666667	22.81109799	0.00966449	2.01482111
IV	2143168	T	T	A	0.99	0.838709677	32.50493167	0.0056493	2.24800558
IV	2212161	G	G	A	0.98245614	0.766666667	27.9476601	0.00713309	2.14672236
IV	2336461	T	T	C	0.981818182	0.766666667	27.45642702	0.00732807	2.13501026
IV	2617615	C	C	T	0.99	0.827586207	37.44046095	0.00451597	2.34524882
IV	3062462	T	T	C	0.99	0.806451613	48.67523335	0.0029369	2.53211029
IV	3349643	A	A	T	0.99	0.857142857	32.29169746	0.00570774	2.24353605
IV	3604175	A	A	G	0.99	0.814814815	43.66118815	0.00351805	2.45369784
IV	3743804	C	C	T	0.99	0.847826087	56.5698849	0.00227579	2.64286772
IV	10361150	T	T	C	0.11627907	0.617647059	36.72515379	0.00465759	2.33183837
IV	15283274	A	A	G	0.99	0.8	23.70562771	0.0091321	2.03942916
V	6413543	A	A	G	0.973684211	0.730769231	24.32432432	0.00879003	2.05600959
V	7616523	A	A	C	0.931034483	0.80952381	28.63592451	0.00687304	2.16285117
V	7616534	T	T	G	0.99	0.75	28.63593748	0.00687303	2.16285147
V	16199972	T	T	C	0.99	0.833333333	29.27551376	0.00664421	2.17755643
V	19074058	C	C	T	0.99	0.833333333	28.57643098	0.00689494	2.16146964
V	19517783	T	T	C	0.99	0.866666667	24.72727273	0.0085778	2.06662391
V	19589100	A	A	C	0.99	0.814814815	39.24480095	0.00418643	2.37815658
V	19642944	G	G	T	0.99	0.827586207	37.06453423	0.00458957	2.33822773
V	19823476	C	C	A	0.99	0.846153846	27.70011591	0.00723033	2.14084164
V	20207374	C	C	T	0.99	0.862068966	24.04545455	0.00894171	2.04857931
V	20577056	C	C	A	0.983606557	0.774193548	28.40869565	0.00695726	2.15756181
V	20601694	T	T	G	0.99	0.846153846	26.79040404	0.00760583	2.11885322
V	20611787	A	A	G	0.99	0.851851852	24.07862408	0.00892346	2.04946676
X	687011	C	C	T	0.99	0.891891892	25.99486089	0.00795945	2.09911702
X	9279323	A	A	G	0.99	0.851851852	27.75016835	0.00721051	2.14203417
X	9747001	G	G	A	0.99	0.860465116	40.3201581	0.00400704	2.39717677
X	9833395	T	T	C	0.99	0.862068966	24.97029115	0.00845362	2.07295728
X	10333161	G	A	G	0.99	0.838709677	33.87929293	0.00529363	2.27624651
X	10612966	G	G	A	0.99	0.857142857	23.84984985	0.00905054	2.04332547
X	10947277	G	G	T	0.99	0.727272727	38.38552189	0.00433865	2.36264495
X	11119348	T	T	A	0.99	0.88372093	29.36340146	0.00661369	2.17955643
X	12028925	C	C	A	0.975609756	0.708333333	26.60346154	0.00768672	2.11425905
X	12040894	G	G	A	0.99	0.888888889	30.57422325	0.00621411	2.20662127
X	12643973	A	A	G	0.970588235	0.708333333	23.92685475	0.00900745	2.04539796
X	17388910	G	G	A	0.99	0.807692308	54.55645342	0.00242151	2.61591361

Table 4- Moderate and high impact VEP results

CHR	POS	REF	ALT	TYPE	GENE
II	2262520	G	T	missense_variant	K09F6.4
II	2262550	C	T	missense_variant	K09F6.4
II	2262562	T	G	missense_variant	K09F6.4
II	2264189	T	A	splice_acceptor_variant	K09F6.4
II	2265429	G	A	missense_variant	K09F6.4
II	2280452	G	C	missense_variant	K09F6.11
II	2280530	T	A	missense_variant, splice_region_variant	K09F6.11
II	2280576	G	A	missense_variant	K09F6.11
II	2281268	C	T	missense_variant	K09F6.9
II	2282456	T	C	missense_variant	K09F6.9
II	2295395	G	T	missense_variant	fbxc-25
II	2295408	G	A	stop_gained	fbxc-25
II	2298217	C	T	missense_variant	B0281.4
II	2298506	A	T	missense_variant	B0281.4
II	2299965	C	G	missense_variant	B0281.5
II	2299998	C	T	missense_variant	B0281.5
II	2300248	C	T	missense_variant	B0281.5
II	2300274	A	G	missense_variant	B0281.5
II	2300275	T	C	missense_variant	B0281.5

II	2300290	G	A	missense_variant	B0281.5
II	2300322	C	T	missense_variant	B0281.5
II	2300322	C	T	splice_donor_variant	B0281.5
II	2300324	A	C	missense_variant	B0281.5
II	2300324	A	C	missense_variant, splice_region_variant	B0281.5
II	2300432	G	T	missense_variant	B0281.5
II	2943174	C	T	missense_variant	irld-50
II	2944204	T	C	missense_variant	irld-49
II	2944762	T	A	missense_variant	irld-49
II	2944768	G	T	missense_variant	irld-49
II	2947254	A	C	missense_variant	srw-79
II	2948203	T	A	missense_variant	srw-79
II	3375175	C	T	missense_variant	bath-10
II	3375325	A	G	missense_variant	bath-10
II	3375395	T	G	missense_variant	bath-10
II	3375853	T	A	missense_variant	bath-10
II	3376313	T	G	missense_variant	bath-10
II	3376574	G	C	missense_variant	bath-10
II	3377167	G	T	missense_variant	bath-9
II	3377576	C	A	missense_variant	bath-9
II	3377583	A	C	missense_variant	bath-9

II	3377615	T	G	missense_variant	bath-9
II	3378450	A	G	missense_variant	bath-9
II	3381908	C	G	missense_variant	T06D4.1
II	3382005	A	C	missense_variant	T06D4.1
II	3382056	T	C	missense_variant	T06D4.1
II	3382171	C	A	missense_variant	T06D4.1
II	3382173	A	G	missense_variant	T06D4.1
II	3382176	T	C	missense_variant	T06D4.1
II	3382190	G	C	missense_variant	T06D4.1
II	3382218	G	T	missense_variant	T06D4.1
II	3382219	T	A	stop_gained	T06D4.1
II	3382424	G	T	splice_donor_variant	T06D4.1
II	3382546	A	G	missense_variant	T06D4.1
II	3382565	T	G	missense_variant	T06D4.1
II	3382569	T	C	missense_variant	T06D4.1
II	13689713	T	G	missense_variant	rpoa-1
IV	1226624	G	A	missense_variant	W09G12.8
IV	1226765	G	A	missense_variant	W09G12.8
IV	1229775	G	A	missense_variant	W09G12.7
IV	1229784	C	T	missense_variant	W09G12.7
IV	1229885	T	A	missense_variant	W09G12.7
IV	1230254	G	A	missense_variant	W09G12.7

IV	1230271	A	C	missense_variant	W09G12.7
IV	1230456	T	C	missense_variant	W09G12.7
IV	1230506	C	G	missense_variant	W09G12.7
IV	1230552	G	A	missense_variant	W09G12.7
IV	1230569	A	G	missense_variant	W09G12.7
IV	1230570	T	A	missense_variant	W09G12.7
IV	1232399	T	A	missense_variant	W09G12.6
IV	1232402	C	T	missense_variant	W09G12.6
IV	1232682	C	T	missense_variant	W09G12.6
IV	2087099	A	G	missense_variant	mca-2
IV	3062540	A	G	missense_variant	Y67D8C.6
IV	3063086	G	A	missense_variant	Y67D8C.6
IV	3065698	C	T	missense_variant	Y67D8C.7
IV	3066791	G	C	missense_variant	Y67D8C.7
IV	3346274	C	T	missense_variant	elks-1
IV	3604052	T	C	missense_variant	Y37E11B.10
IV	3604175	A	G	missense_variant	Y37E11B.10
IV	3604622	T	C	missense_variant	Y37E11B.10
V	6415176	C	A	missense_variant	T05H4.7
V	6415575	A	G	missense_variant	T05H4.7
V	6416251	A	T	missense_variant	T05H4.7
V	6417362	C	T	missense_variant	T05H4.7
V	7612361	G	A	missense_variant	clec-215

V	7615005	G	T	missense_variant	clec-215
V	7615185	A	G	missense_variant	clec-215
V	7618579	T	G	missense_variant	C50E3.15
V	7618890	G	A	missense_variant	C50E3.15
V	7620102	G	A	missense_variant	C50E3.15
V	7620104	G	A	missense_variant	C50E3.15
V	7620165	C	T	missense_variant	C50E3.15
V	7620781	A	C	missense_variant	C50E3.12
V	7620787	C	T	missense_variant	C50E3.12
V	7620820	T	C	missense_variant	C50E3.12
V	7620961	C	T	missense_variant	C50E3.12
V	7620976	C	T	missense_variant	C50E3.12
V	7621002	C	T	missense_variant	C50E3.12
V	7621009	T	G	missense_variant	C50E3.12
V	7621045	C	A	stop_gained	C50E3.12
V	7621077	C	T	missense_variant	C50E3.12
V	7621099	C	T	missense_variant	C50E3.12
V	7621108	T	C	missense_variant	C50E3.12
V	7621109	T	G	missense_variant	C50E3.12
V	7621113	T	A	missense_variant	C50E3.12
V	7621210	T	A	missense_variant	C50E3.12
V	7621211	A	T	missense_variant	C50E3.12
V	8993048	A	C	missense_variant	F59E11.6

V	16195357	A	T	missense_variant	F11A5.3
V	16195506	T	A	missense_variant	F11A5.3
V	16195992	A	G	missense_variant	F11A5.4
V	16196010	A	G	missense_variant	F11A5.4
V	16196110	A	G	missense_variant	F11A5.4
V	16199790	T	C	missense_variant	F11A5.5
V	16199804	G	A	missense_variant	F11A5.5
V	19039259	C	A	missense_variant	asp-16
V	19039261	T	C	missense_variant	asp-16
V	19039315	G	C	missense_variant	asp-16
V	19044507	G	C	missense_variant	asp-15
V	19076968	T	A	missense_variant	nhr-95
V	19513201	G	C	missense_variant	phy-4
V	19641909	C	T	missense_variant	Y43F8C.7
V	19646139	T	C	missense_variant	Y43F8C.9
V	19646162	A	G	missense_variant	Y43F8C.9
V	19824474	C	G	missense_variant	pcp-4
V	19826838	C	T	missense_variant	pcp-4
V	20202577	A	T	missense_variant	fbxa-114
V	20203574	A	C	missense_variant	fbxa-113
V	20203724	T	A	missense_variant	fbxa-113
V	20203736	C	T	missense_variant	fbxa-113
V	20204058	G	A	missense_variant	fbxa-113

V	20204066	A	C	missense_variant	fbxa-113
V	20204218	A	T	missense_variant	fbxa-113
V	20204223	A	G	missense_variant	fbxa-113
V	20204418	G	A	missense_variant	fbxa-113
V	20204603	C	G	missense_variant	fbxb-59
V	20204734	C	T	missense_variant	fbxb-59
V	20204891	C	A	missense_variant	fbxb-59
V	20204929	A	G	missense_variant	fbxb-59
V	20204932	A	G	missense_variant	fbxb-59
V	20205029	G	A	missense_variant	fbxb-59
V	20207021	C	T	missense_variant	srbc-34
V	20207324	C	G	missense_variant	srbc-34
V	20575071	C	T	missense_variant	clec-263
V	20575087	G	T	missense_variant	clec-263
V	20575248	G	A	missense_variant	clec-263
V	20575273	C	T	missense_variant	clec-263
V	20575332	G	A	missense_variant	clec-263
V	20575355	A	T	missense_variant	clec-263
V	20575383	G	C	missense_variant	clec-263
V	20575416	A	C	missense_variant	clec-263
V	20575652	C	G	missense_variant	clec-263
V	20575885	A	T	missense_variant	clec-263
V	20576452	A	G	missense_variant	clec-263

V	20576745	A	G	missense_variant	clec-263
V	20577553	A	T	missense_variant	clec-263
V	20597052	G	A	missense_variant, splice_region_variant	F46B3.1
V	20597053	C	G	missense_variant, splice_region_variant	F46B3.1
V	20597143	C	A	missense_variant	F46B3.1
V	20600292	A	T	missense_variant	col-163
V	20600309	A	G	missense_variant	col-163
V	20600450	G	C	missense_variant	col-163
V	20600468	A	T	missense_variant	col-163
V	20600477	C	A	missense_variant	col-163
V	20600546	G	A	missense_variant	col-163
V	20600619	C	T	missense_variant	col-163
V	20600678	T	G	missense_variant	col-163
V	20600730	G	A	missense_variant	col-163
V	20600967	G	T	missense_variant	col-163
V	20601191	C	T	missense_variant	col-163
V	20601253	T	G	missense_variant	col-163
V	20601422	T	C	missense_variant	col-163
V	20601482	A	G	missense_variant	col-163
V	20601929	T	C	missense_variant	col-163
V	20603509	G	A	missense_variant	ttr-11

V	20604115	C	G	missense_variant	ttr-11
V	20604305	A	T	missense_variant	ttr-11
V	20604370	T	G	missense_variant	ttr-11
V	20605308	C	G	missense_variant	ttr-57
V	20605513	T	C	missense_variant	ttr-57
V	20605516	C	T	missense_variant	ttr-57
V	20605562	G	A	stop_gained	ttr-57
V	20605567	A	G	missense_variant	ttr-57
V	20605570	A	C	missense_variant	ttr-57
V	20606447	T	G	missense_variant	F46B3.23
V	20606666	A	C	missense_variant	F46B3.23
V	20607967	C	T	missense_variant	ttr-12
V	20608242	A	G	missense_variant	ttr-12
V	20611009	C	G	missense_variant	grd-2
V	20611011	T	G	missense_variant	grd-2
V	20611041	C	A	missense_variant	grd-2
V	20611138	T	G	missense_variant	grd-2
V	20613096	T	A	missense_variant	grd-2
V	20613135	A	G	missense_variant	grd-2
V	20613345	A	G	missense_variant	grd-2
V	20613357	C	T	missense_variant	grd-2
X	9279323	A	G	missense_variant	B0416.4
X	9835368	G	C	missense_variant	ipla-2

X	10947277	G	T	missense_variant	T21B6.3
X	10947621	T	C	missense_variant	T21B6.3
X	12043027	T	C	missense_variant	B0198.3
X	12045180	T	C	missense_variant	B0198.3
X	12352101	C	G	missense_variant	F42F12.4

Chapter 3

BRAP-2 negatively regulates Wnt-mediated endoderm specification in *C. elegans*

(in preparation for publication with minor adjustments)

3.1 Introduction

BRAP-2 is an orthologue of human BRAP (BRCA-1 associated protein, also referred to as BRAP2 and IMP) and is an E3 ubiquitin ligase. BRAP was first identified in a yeast-two hybrid screen for binding partners of the BRCA-1 nuclear localization signal (NLS) (S. Li et al., 1998). BRCA-1 is a DNA damage-response factor involved in homologous recombination repair and is famously associated with high risks of familial breast and ovarian cancers (Burgess & Puhalla, 2014; Hall et al., 1992; Helpman et al., 2017; Rakha et al., 2008). BRAP has been found to regulate NLS-dependent nuclear import of viral proteins (Fulcher et al., 2009), and two yeast-two hybrid screens in testis (human and mouse) showed a range of proteins that show BRAP-dependent nuclear localization (Davies et al., 2013; Fatima et al., 2015). BRAP also regulates NF-κB nuclear translocation through interactions with SCF (Skp1-Cullin1-F-box protein)-type ubiquitin ligase complex and ubiquitin-like molecule Nedd8 (Takashima et al., 2013). Owing to its complex interactions with multiple signaling pathways, polymorphisms in BRAP are associated with numerous human diseases, including metabolic syndrome, schizophrenia, and coronary artery disease (Avery et al., 2011; Chida-Nagai et al., 2019; Hinohara et al., 2009; Imaizumi et al., 2018; Kelly et al., 2020; Kim et al., 2018; Kubo et al., 2017; Liao et al., 2011; Ozaki et al., 2009; Takeuchi et al., 2012; Wu et al., 2013; Y. Yamada et al., 2017; Zhang et al., 2014; Zhao et al., 2017).

In *C. elegans*, mutants that lack *brap-2* undergo larval arrest when exposed to oxidative stress, and that this response requires BRCA1 orthologue BRC-1 (Koon & Kubiseski, 2010). In postembryonic animals, SKN-1 is involved in a wide array of homeostatic functions, including mitigating oxidative stress, a conserved role of Nrf2 (An & Blackwell, 2003). Loss of *brap-2* results in constitutive nuclear localization of phosphorylated SKN-1, which is regulated by p38 MAPK orthologue PMK-1, leading to upregulation of detoxification genes. This transcriptional response is also partially dependent on GATA factor ELT-3 and lipid metabolism regulator NHR-49 (Q. Hu et al., 2017, 2018; Inoue et al., 2005). Additionally, BRAP-2 promotes germline apoptosis in response to DNA damage by inhibiting cell-survival signals from PMK-1/SKN-1 and the Insulin/ Insulin-like growth factor Signaling (IIS) pathway (D'Amora et al., 2018). Given the conserved interactions between BRAP-2 and SKN-1, we sought to explore the role of BRAP-2 in endoderm development. Here, we report that the loss of *brap-2* rescues the loss of endoderm due to the depletion of SKN-1. We showed that this effect depends on the function of POP-1, SYS-1 and PAL-1, and that the expression of *end-1* is upregulated in *brap-2* deletion mutant even in the absence of the immediate transcriptional activator SKN-1. Evidence suggests that BRAP-2 normally works to repress the effect of Wnt signaling on POP-1 as *brap-2* mutants show decreased nuclear/cytoplasmic POP-1 ratio when compared to wildtype. We have thus identified a new role for BRAP-2 in fine-tuning the levels of the Wnt effector to ensure the robust specification of the endoderm during early embryonic development.

3.2 Materials and Methods

C. elegans culture and strains: All strains were cultured under standard conditions, as previously described (Brenner, 1974). Strains were obtained from the Caenorhabditis Genetics Center at the University of Minnesota unless noted otherwise. FX5132 was a generous gift from Dr. Shohei Mitani's lab. Double mutant strains were generated according to standard protocols. The N2 Bristol strain was used as the wild type, and all experiments were conducted at 20 °C. *C. elegans* strains used in this chapter are listed in Table 1 below.

Table 1 List of strains used in Chapter 3

Strain	Genotype	Source
FX5132	brap-2 (tm5132) II	Mitani lab
RB1346	EEED8.16 (ok1492) II	CGC
JR4080	brap-2 (tm5132) II [FX5132 backcrossed 5X with N2]	Rothman lab
JR3936	dpy-13 (e184) skn-1 (zu67) IV / nT1 [qIs51] (IV;V)	Rothman lab
JR3886	brap-2 (tm5132) II ; dpy-13 (e184) skn-1 (zu67) IV / nT1 [qIs51] (IV;V)	Rothman lab
JR1904	wIs117 [pRF4 + pMM414 (med-1::GFP::POP-1::med-1 3'UTR)]	Rothman lab
JR3911	brap-2 (tm5132) II ; wIs117 [pRF4 + pMM414 (med-1::GFP::POP-1::med-1 3'UTR)]	Rothman lab
JR1186	wIs93 [pRF4 + pMM280 (med-1::GFP::MED-1)]	Rothman lab
JR4183	brap-2 (tm5132) II ; wIs93 [pRF4 + pMM280 (med-1::GFP::MED-1)]	Rothman lab
JR996	wIs75 [end-3::NLS::lacZ + end-3::NLS::GFP + pRF4]	Rothman lab
JR4184	brap-2 (tm5132) II ; wIs75 [end-3::NLS::lacZ + end-3::NLS::GFP + pRF4]	Rothman lab
JR4021	itIs37 [pie-1p::mcherry::H2B::pie-1 3'UTR + unc-119 (+)] IV ; stls10226 [his-72p::Itls-24::mcherry::let-858 3'UTR + unc-119 (+)] ; teIs46 [pRL1417; end-1p::GFP::H2B + unc-119 (+)]	Rothman lab
JR4022	brap-2 (tm5132) II ; itIs37 [pie-1p::mcherry::H2B::pie-1 3'UTR + unc-119 (+)] IV ; stls10226 [his-72p::Itls-24::mcherry::let-858 3'UTR + unc-119 (+)] ; teIs46 [pRL1417; end-1p::GFP::H2B + unc-119 (+)]	Rothman lab
AZ212	unc-119 (ed3) ruIs32 [pie-1p::GFP::H2B] III	CGC
JR3913	brap-2 (tm51320) II ; unc-119 (ed3) ruIs32 [pie-1p::GFP::H2B] III	Rothman lab
JR3924	syIs243 [myo-3p::TOM20::mRFP + unc-119 (+) + pBS Sk+] ; kcIs6 [IFB-2::CFP] IV ; rrIs1 [elt-2p::GFP] X	Rothman lab
JR3940	brap-2 (tm5132) II ; syIs243 [myo-3p::TOM20::mRFP + unc-119 (+) + pBS Sk+] ; kcIs6 [IFB-2::CFP] IV ; rrIs1 [elt-2p::GFP] X	Rothman lab
JR3933	dpy-11 (e1180) mom-2(or42) V /nT1 [qIs51] (IV;V)	Rothman lab
JR3934	brap-2 (tm5132) II ; dpy-11 (e1180) mom-2(or42) V /nT1 [qIs51] (IV;V)	Rothman lab
PS4867	syIs146 [mom-2::GFP + unc-119(+)]	CGC

JR4023	brap-2 (tm5132) II ; syIs146 [mom-2::GFP + unc-119(+)]	Rothman lab
TX964	unc-119 (ed3) III ; him-3(e1147) IV ; teIs98 [(pRL1450) pie-1p::GFP::sys-1 + unc-119(+)]	CGC
JR4316	brap-2(tm5132) II ; unc-119 (ed3) III ; him-3 (e1147) IV ; teIs98 [(pRL1450) pie-1p::GFP::sys-1 + unc-119 (+)]	Rothman lab
MS562	unc-119 (ed4) III ; irIs31 [tbx-35::GFP + unc-119 (+)]	Maduro lab
JR4317	brap-2 (tm5132) II ; unc-119 (ed4) III ; irIs31 [tbx-35::GFP + unc-119(+)]	Rothman lab
MY16	*C. elegans* wild isolate	CGC
PD7963	ccIs7963 [hlh-1::GFP + rol]	Fire lab
JR4186	brap-2 (tm5132) II ; ccIs7963 [hlh-1::GFP + rol]	Rothman lab
N2	Domesticated laboratory strain	Rothman lab

RNA interference (RNAi): RNAi clones were obtained from either the Vidal (Rual et al., 2004) or Ahringer (Kamath et al., 2003) libraries. HT115 *E. coli* was transformed with L4440 plasmid expressing desired dsRNA. Bacteria containing the empty L4440 RNAi feeding vector was used the control experiments. The RNAi clones were grown overnight at 37°C in 3mL LB with 50µg/ml ampicillin and 50µg/ml tetracycline. The cultures were diluted 1:10 and grown for an additional 4 hours at 37°C in LB with ampicillin and tetracycline. The cultures were spiked with 1mM IPTG before seeding 80µL onto NGM plates supplemented with 1mM IPTG and 25 µg/ml carbenicillin. Worms were synchronized according to standard practices (Brenner, 1974) and fed E. coli OP50 on nematode growth medium (NMG) plates until the L4 larval stage. L4 worms fed on RNAi for at least 24 hours for all experiments before transfer to a new plate for egg-laying. Egg laying proceeded for 4-12 hours, and adult worms were removed. After another 7-12 hours, embryos were scored for gut granules using polarized light. All embryos were scored within 24 hours of being laid.

RT-PCR: Mixed staged worms were washed with M9 buffer solution, and RNA was collected using the Zymo Direct-zol RNA prep kit (Zymo Research #R2051). cDNA was generated using Superscript III 1st Strand Synthesis System (using the oligo(dT) primer to

target mRNA, Invitrogen #18080051). PCR was performed using KOD Hot Start polymerase enzyme (MilliporeSigma #71086). Primers used for RT-PCR: fwd-CTCAAGCTTGTCTCCACCCT, rv- ATCCGCCAATTGGAGTTATG. PCR products were run on a 1% agarose gel and bands were excised for sequencing. WT and *brap-2(tm5132)* PCR products were Sanger sequenced at UC Berkeley DNA Sequencing Facility using the RT-PCR primers, and sequences were compared using ApE software.

Testing ability to enter the dauer state: NGM plates with worms were left for two weeks after food depletion, and only plates without any contamination were used for the assay. Worms were washed and placed on a rotator in 1% SDS in M9 solution for 30 minutes. After washing again with M9, worms were aliquoted and live and dead worms were counted. Live worms thrash in solution whereas dead worms appear as a straight rod.

Mating experiment testing maternal effect: Crosses were set up according to standard practices (Fay, 2018). In these experiments, fluorescently labeled males (syIs243 [myo-3p::TOM20::mRFP + unc-119 (+) + pBS Sk+]) were mated with non-fluorescent hermaphrodites. Prior to mating, parents were on *skn-1(RNAi)* for 24 hours, and then transferred to mating plates that were also seeded with *skn-1(RNAi)* bacteria. Parents were removed after egg laying, and arrested embryos were scored for the presence of MYO-3::RFP and gut granules the next day.

Microscopy and fluorescent analysis: Images were taken on the Nikon Eclipse Ti using NIS-Elements AR 4.13.05 software and images were processed using FIJI/ImageJ. For early

endoderm marker analysis in Fig. 1C & D, only embryos at stages where GFP was expected were scored. MED-1::GFP was seen from the 8-cell stage to ~100-cell stage; END-3::GFP expression was seen from 2E-8E; END-1::GFP expression started at 2E and lasted until comma stage. For *tbx-35*::GFP analysis, embryos at 2E stage were scored for expression. For analysis of nuclear/cytoplasmic ratio of POP-1 and SYS-1, z-stacks of strain expressing *med-1p::gfp::pop-1* and *pie-1p::GFP::sys-1* were taken in the DIC and GFP channels. Measurements were made on a single plane that was in the best focus for the cell considered. Cell perimeters were drawn based on DIC images. Whole cell fluorescence and nuclear fluorescence were corrected for by taking the raw integrated pixel density and subtracting the area x mean fluorescence intensity of the image background. Cytoplasmic fluorescence was calculated by subtracting the corrected nuclear fluorescence intensity from the corrected whole cell fluorescence intensity. For analysis of MOM-2::GFP expression, max intensity projections of z-stacks were used for fluorescent measurements, and entire embryos were analyzed. To measure HLH-1::GFP, embryos were imaged after gastrulation but before entering the bean stage, when GFP expression was seen around the perimeter of the embryo. Whole embryo fluorescence was measured as described above, except the entire embryo was measured as opposed to one cell. *myo-3*::mRFP was measured in synchronized L1 larvea by taking a measurement of the entire worm and subtracting the mean fluorescence intensity of the image background, as described above.

qPCR: For both N2 and MY16, day 1 adults that had just begun egg laying were washed with M9 buffer solution, and RNA was collected using the Zymo Direct-zol RNA prep kit (Zymo Research #R2051). cDNA was generated using Superscript III 1st Strand Synthesis

System (using the oligo(dT) primer to target mRNA, Invitrogen #18080051). qPCR was performed on a Bio-Rad CFX96 Real-Time PCR machine using the Light Cycler 480 SYBR Green I Master mix (Roche, Cat. #04 707 516 001). Three biological replicates (each with three technical replicates) were used for analysis. Cycle threshold data was converted into linearized values, and *brap-2* transcript levels were normalized to *act-1* transcript levels before calculating fold-over N2 expression. The following primers were used to amplify the targets: *brap-2*F- ACTACCACAAAGGACGACGG, *brap-2*R-TAGGCAACTTGGTCAGGTGGT, *act-1*F- TCCATTGTCGGAAGACCACG, *act-1*R-GGTGACGATACCGTGCTCAA.

Statistics: Standard deviation is used throughout the chapter when reporting quantitative results in text and as error bars on graphs. A Student's two-tailed t-test was used for statistical analysis throughout.

3.3 Results

Loss of BRAP-2 suppresses the endoderm defect in SKN-1 deficient embryos

Eliminating maternal SKN-1 either through RNAi or a strong loss-of-function mutation *skn-1(zu67)* leads to a partial penetrant loss-of-endoderm phenotype, such that only 28.5 ± 6.5% or 34.8 ± 5.1% of embryos contain a differentiated gut, respectively (Fig. 1A), consistent with previous reports (Bowerman et al., 1992; Ewe et al., 2020; Maduro et al., 2005; Torres Cleuren et al., 2019). Knocking out *brap-2* partially rescues *skn-1(RNAi)*-induced loss-of-endoderm phenotypes: 86.0 ± 5.1 and 53.0 ± 6.8% of *brap-2(tm5132)* and *brap-2(ok1492)* embryos produce endoderm, respectively (Fig. 1A & B). RNAi targeting

GFP is equally effective in wild type and *brap-2(tm5132)*, ruling out the possibility that the impenetrate phenotype is the result of reduced RNAi efficacy in *brap-2(-)* mutants (Fig. 2). Confirming our observation, *brap-2(tm5132)* suppresses the gut defect of *skn-1(zu67)* mutant, resulting in 71.2 ± 13.2% of embryos with endoderm (Fig. 1A).

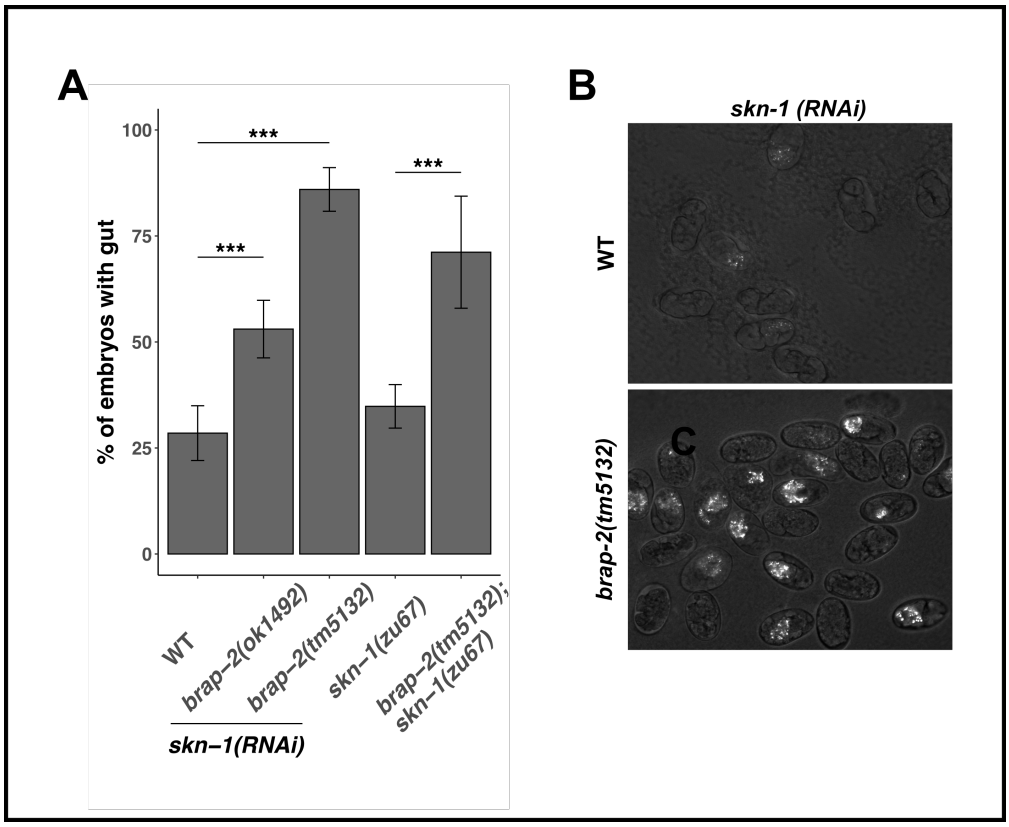

Figure 1 Eliminating *brap-2* suppresses *skn-1*(-) endoderm loss

A) *brap-2* mutants suppress endoderm loss in *skn-1 (RNAi)* and in *skn-1(zu67)* mutants. At least 800 embryos were scored in each condition. B) Image of gut granules using polarized light for WT and *brap-2(tm5132)* embryos after *skn-1(RNAi)*.

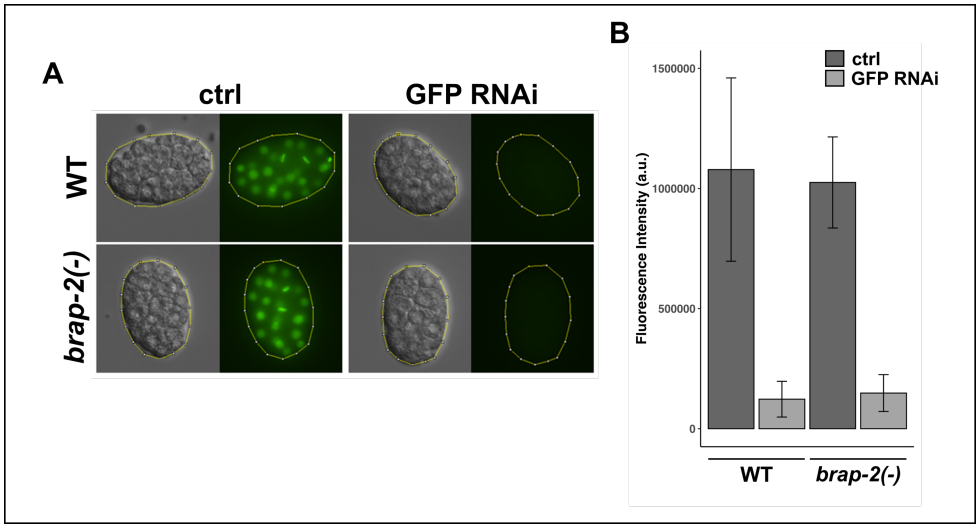

Figure 2 *brap-2(tm5132)* is not RNAi-deficient

A) DIC and GFP images of embryos expressing *pie-1p*::GFP::H2B on control *(RNAi)* and GFP *(RNAi)*. Fluorescence was measured in the embryos in the areas outlined. B) Corrected total embryo fluorescence was measured in both RNAi conditions. We see similar RNAi efficiency between WT and *brap-2(tm5132)*. N > 20 embryos for each condition.

brap-2(-) mutants have in-frame deletions and can enter dauer

It is notable that the *brap-2* allele *ok1492* is a larger C-terminal deletion that removes most of the conserved domains (albeit in frame) (Koon & Kubiseski, 2010), while the *tm5132* allele is a 313 bp deletion with a 7 bp addition and lacks most of the RNA recognition motif in the third and fourth exon. In silico splicing analysis suggested that *tm5132* may result in an in-frame deletion, preserving the C-terminus. Indeed, we detected *brap-2* mRNA in the *tm5132* mutant using RT-PCR, and Sanger sequencing of the product agrees with the in-frame deletion (Fig. 3). An earlier paper focusing on *brap-2(ok1492)* noted that this strain cannot enter the dauer state (D'Amora et al., 2018). Conversely, we found that both *brap-2(ok1492)* and *brap-2(tm5132)* strains can enter dauer, although they

are both slightly dauer defective when compared to wildtype strains (WT- 54.70 ± 30.40%, *brap-2(ok1492)*- 20.16 ± 12.78%, *brap-2(tm5132)*- 19.79 ± 24.56%) (Fig. 4).

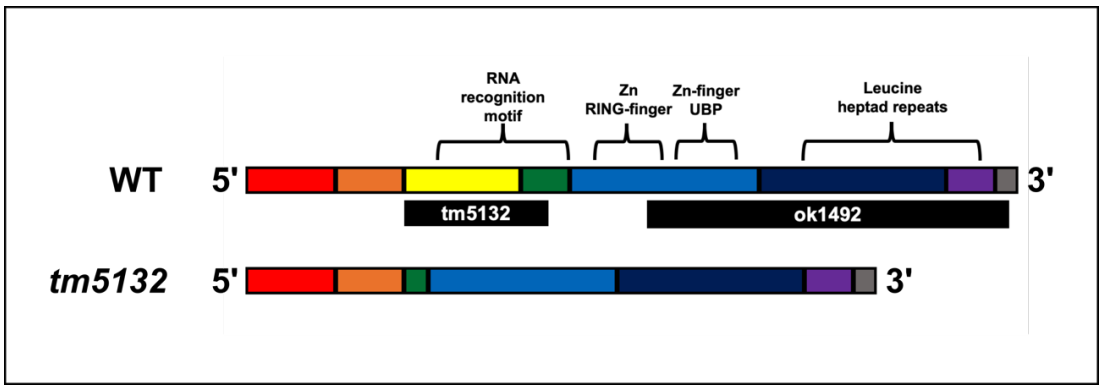

Figure 3 Schematic of *brap-2* mutant alleles

brap-2 mRNA diagram with *tm5132* and *ok1492* deletions shown in black below. *tm5132* mRNA sequencing matches an in-frame deletion removing the third exon and most of the fourth exon, most of the RNA recognition motif.

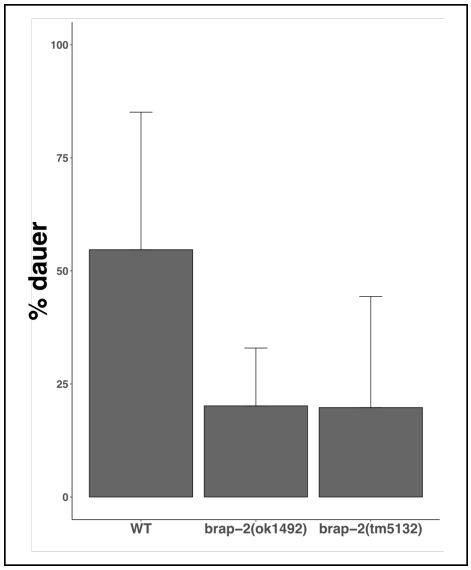

Figure 4 *brap-2(-)* strains can enter dauer

Bar graph showing the percent of worms that survived 1% SDS treatment for 30 minutes, indicating their ability to enter the dauer state. Error bars show standard deviation.

Loss of BRAP-2 does not restore the MS lineage in *skn-1 (RNAi)*

Since SKN-1 is the major transcription factor acting in both MS and E cell lineages (Bowerman et al., 1992), we wanted to test whether loss of BRAP-2 can restore MS gene expression in addition to E gene expression. In MS, SKN-1 activates MED-1/2, which then activate TBX-35 to specify MS, as the E gene cascade is repressed by the transcription factor POP-1 (Broitman-Maduro, 2006; Lin et al., 1995). We examined *tbx-35*::GFP expressing embryos under control *(RNAi)* and *skn-1(RNAi)*, and saw no expression in the early embryo in both the wildtype strain and *brap-2(-)* after *skn-1(RNAi)* (Fig. 5).

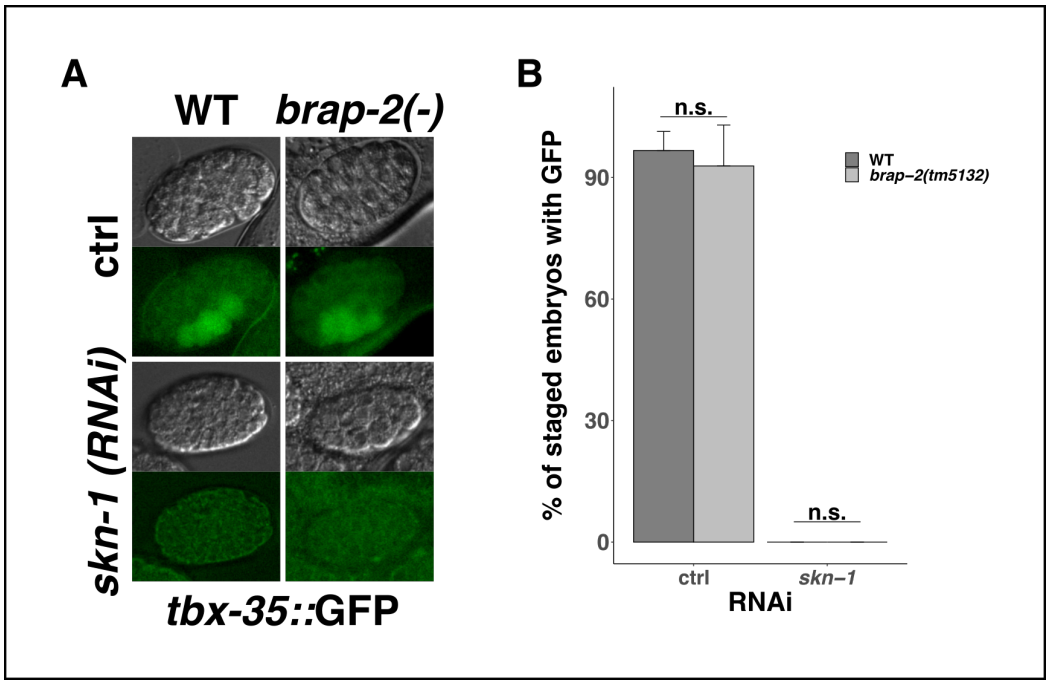

Figure 5 *brap-2(-)* does not rescue MS in *skn-1 (RNAi)*

A) DIC and GFP images of WT and *brap-2(tm5132)* strains expressing *tbx-35*::GFP after control and *skn-1 (RNAi)* treatments. B) Bar graph showing the percent of appropriately staged embryos expressing *tbx-35*:GFP after control and *skn-1 (RNAi)* treatments. Error bars

show standard deviation. As seen in both A and B, neither WT nor *brap-2(tm5132)* strains express the MS-marker *tbx-35*::GFP in early embryos following *skn-1 (RNAi)*.

Loss of maternal BRAP-2 induces END-1 expression

Eliminating *skn-1* causes EMS-to-C transformation (Bowerman et al., 1992). Loss of *brap-2* may restore mesendoderm specification or cause transformation of other progenitor cell(s) to E fate in the absence of *skn-1*. To determine the developmental origin of the gut cells when SKN-1 and BRAP-2 are removed, we examined various early endoderm specification markers in the mutant (Fig. 6). In these experiments, we found no evidence of ectopic END-1::GFP expression in *brap-2(-)*; *skn-1(RNAi)* embryos, suggesting the function of BRAP-2 is restricted to the E lineage (Fig. 6A). While *skn-1 (RNAi)* treatment alone severely impacts END-1 expression (14.3 ± 0.0%), the number of embryos expressing END-1::GFP is significantly increased in *brap-2(-)*; *skn-1(RNAi)* (58.6 ± 2.0%) (Fig. 6B). On the other hand, knocking out *brap-2* failed to restore *med-1* or *end-3* expression in *skn-1(RNAi)* embryos (Fig. 6).

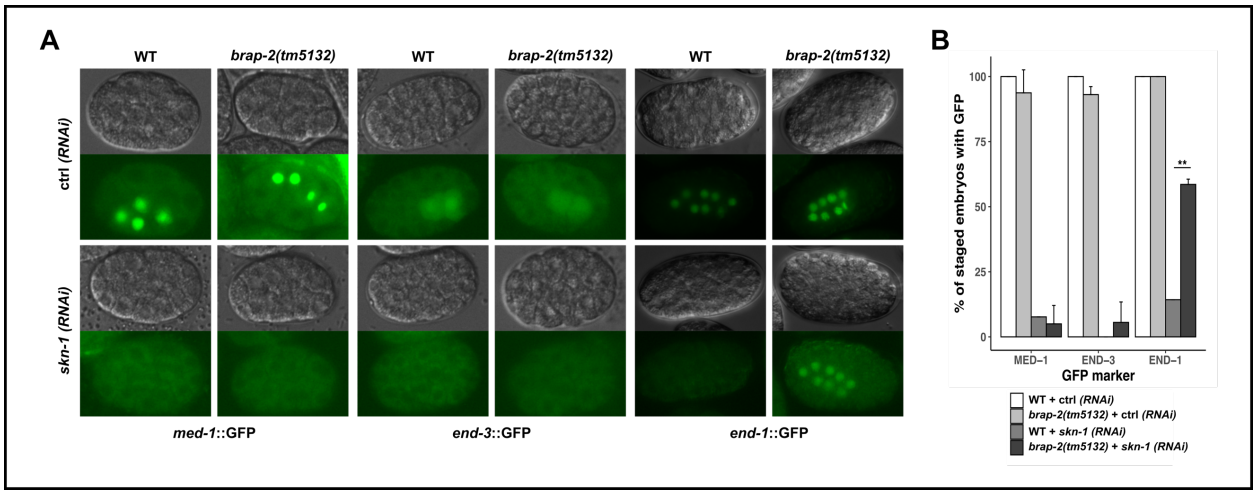

Figure 6 *brap-2(-)* restores *end-1* transcription

A) Images of staged embryos for each GFP marker for WT and *brap-2(tm5132)* on control *(RNAi)* and *skn-1 (RNAi)*. B) Percent of staged embryos showing the GFP marker indicated on the x-axis. WT embryos on control *(RNAi)* (L4440) were used to determine the expression pattern for each marker (See methods and materials). $8 \leq N \leq 43$ for all conditions.

BRAP-2 is a maternal effect gene

To test whether, like SKN-1, BRAP-2 is a maternal factor, we generated *brap-2(-/+)* heterozygotes using either wildtype mothers and *brap-2(-)* fathers or vice versa. Crossed progeny were identified by paternally-derived myo-3-driven RFP. We found that the only *brap-2(-/+)* progeny from *brap-2(-)* mothers can effectively rescue the endoderm defect in *skn-1* deficient embryos, showing 81.7 ± 2.3% of embryos with gut, compared to 44.2 ± 6.7% in heterozygotes from wildtype mothers, showing maternal BRAP-2 is required for its effect in endoderm specification (Fig. 7). In support of this result, *brap-2* transcript is found in the early embryo and is gradually enriched in the germline lineage (Fig. 8). Together, these results show that inactivation of maternally provided BRAP-2 induces endoderm specification in *skn-1(-)* embryos by activating *end-1*.

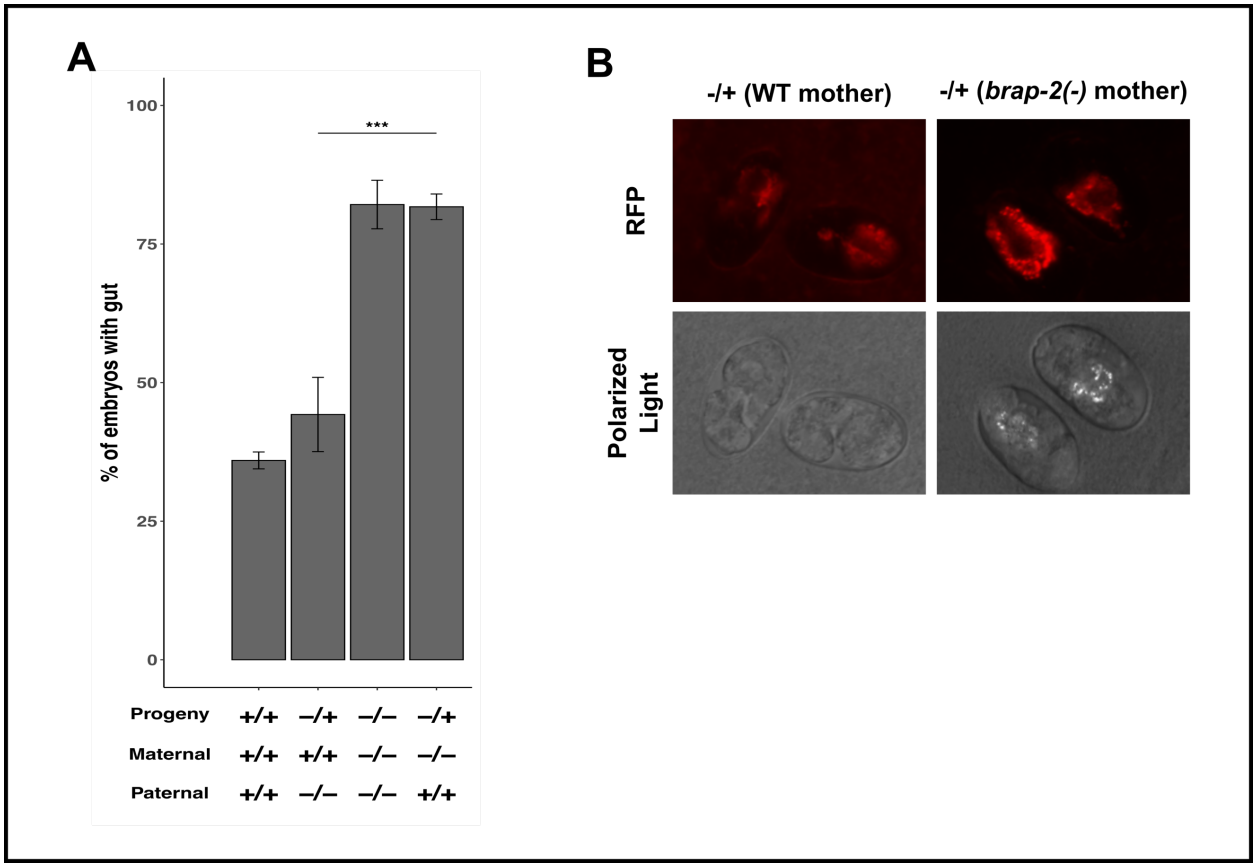

Figure 7 *brap-2(-)* suppression of *skn-1(RNAi)* is a maternal effect

A) Parents were kept on *skn-1(RNAi)*, and only cross progeny were scored (see Methods and Materials). Heterozygotes from *brap-2(tm5132)* mothers showed the rescue effect while those from WT mothers did not. N ≥ 70 for all conditions. B) RFP and gut granules images of heterozygous progeny from E. Only RFP+ embryos were scored for the presence of gut granules. Standard deviation is shown on all bar graphs, and significance was determined using a Student's t-test.

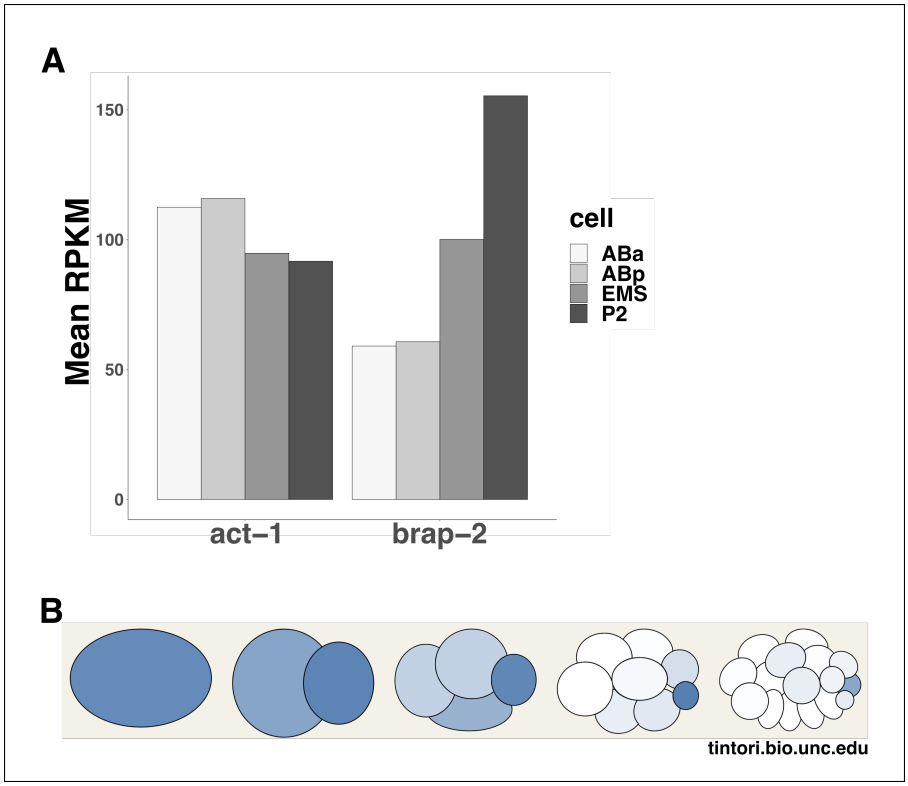

Figure 8 *brap-2* transcript is expressed in the early embryo

A) Bar graph showing the mean RPKM of *act-1* and *brap-2* transcripts found in each cell of the 4-cell embryo. B) Visualization of *brap-2* transcript distribution from the 1-cell stage through the 16-cell stage. Darker cells indicate that more transcript was found in single-cell RNAseq experiments. Data for both A and B was pulled from the RNAseq resource at tinotori.bio.unc.edu. (*Gene Expression Visualization*)

brap-2 expression in increased in a SKN-1-dependent strain

Since it appears that BRAP-2 is normally working to inhibit SKN-1-independent endoderm development, we wanted to test if a wild isolate strain with high SKN-1dependence has altered *brap-2* expression. MY16 can only produce about 2% of embryos with gut under *skn-1(RNAi)* (see Chapter 2). Within the *brap-2* gene, there is only one

variant between N2 and MY16; it is in the first intron and not at a site predicted to affect splicing. The first intron can act as a transcriptional regulatory region and can bind additional transcription factors (Bass et al., 2014), so we hypothesized that *brap-2* expression may be increased in MY16, thereby increasing inhibition of SKN-1-dependent gut development. Indeed, we found that *brap-2* transcripts were increased in MY16 compared to N2 (1.72 ± .23 fold expression over N2) (Fig. 9). This evidence strengthens our hypothesis that BRAP-2 modulates SKN-1-independent endoderm activating inputs.

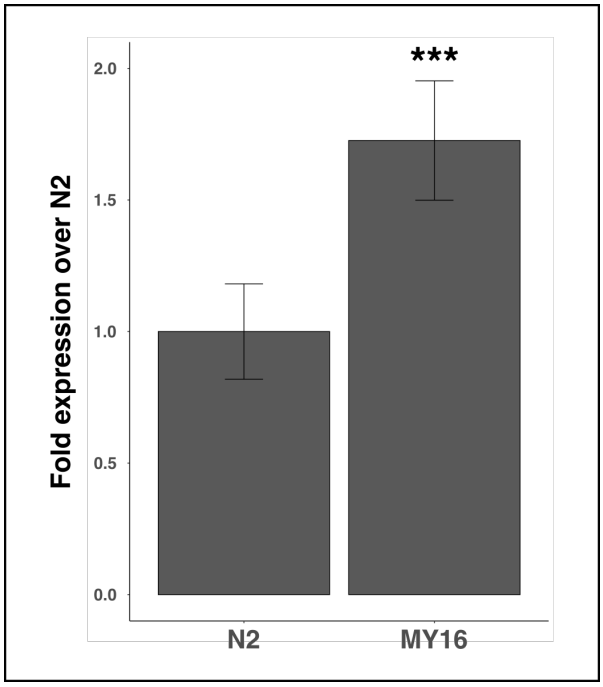

Figure 9 *brap-2* is expressed at higher levels in the wild isolate MY16 than N2

Bar graph showing the fold expression over N2 (WT) of *brap-2* mRNA in the MY16 wild isolate strain, normalized to *act-1* mRNA levels. RNA was collected from day 1 adults, and mRNA levels were measured using qPCR. Error bars show standard deviation, and the asterisks show significance between the linearized and normalized *brap-2* mRNA levels in WT vs. MY16 as determined by a Student's t-test.

BRAP-2 attenuates multiple endoderm-activating inputs

Phosphorylated POP-1, in complex with co-activator SYS-1, together with SKN-1 and PAL-1, account for the major endoderm-activating inputs (Huang et al., 2007; Maduro et al., 2005). We sought to determine the node at which BRAP-2 functions by knocking down *pop-1*, *sys-1*, and *pal-1* alone or in combination in *brap-2*;*skn-1* double mutants. The suppression of loss of endoderm by *brap-2(-)* in *skn-1* deficient embryos is abrogated in the absence of POP-1, SYS-1, or PAL-1 function *(pop-1(RNAi)*: 29.9 ± 1.0%; *sys-1(RNAi)*: 19.6 ± 2.6%; *pal-1(RNAi)*: 26.8 ± 0.1% of embryos with gut). While simultaneous depletion of SKN-1, Wnt signaling and PAL-1 by RNAi completely abolishes endoderm development (Maduro et al., 2005), the additional depletion of *brap-2* results in embryos with gut granules (5.9 ± 3.3% of *brap-2(tm5132)*; *skn-1(zu67*); *pop-1(RNAi)*; *sys-1(RNAi)*; *pal-1(RNAi)* embryos contains gut granules (Fig. 10A), a similar phenotype to that of *brap-2(tm5132);mom-2(or42);skn-1(RNAi)* (8.4 ± 1.4% gut) (Fig. 10B). Our results suggest that loss of *brap-2* restores gut specification in *skn-1* mutants largely by enhancing Wnt signaling activity, which subsequently modulates POP-1 cellular distribution and function (see below); however, the residual endoderm made in mutants lacking SKN-1, PAL-1, POP-1, and SYS-1 suggests that BRAP-2 normally represses at least one additional endoderm activator.

Surprisingly, loss of SYS-1 enhances the gut defect of *brap-2(tm5132);skn-1(zu67);pop-1(RNAi);pal-1(RNAi)* (16.0 ± 2.6% vs. 5.9 ± 3.3% of embryos with gut) (Fig. 10A & B) suggesting incomplete penetrance of *pop-1 (RNAi)* is enhanced by the depletion of Wnt-activated SYS-1 or that the beta-catenin may function independent of POP-1 in E fate activation (Doumpas et al., 2019).

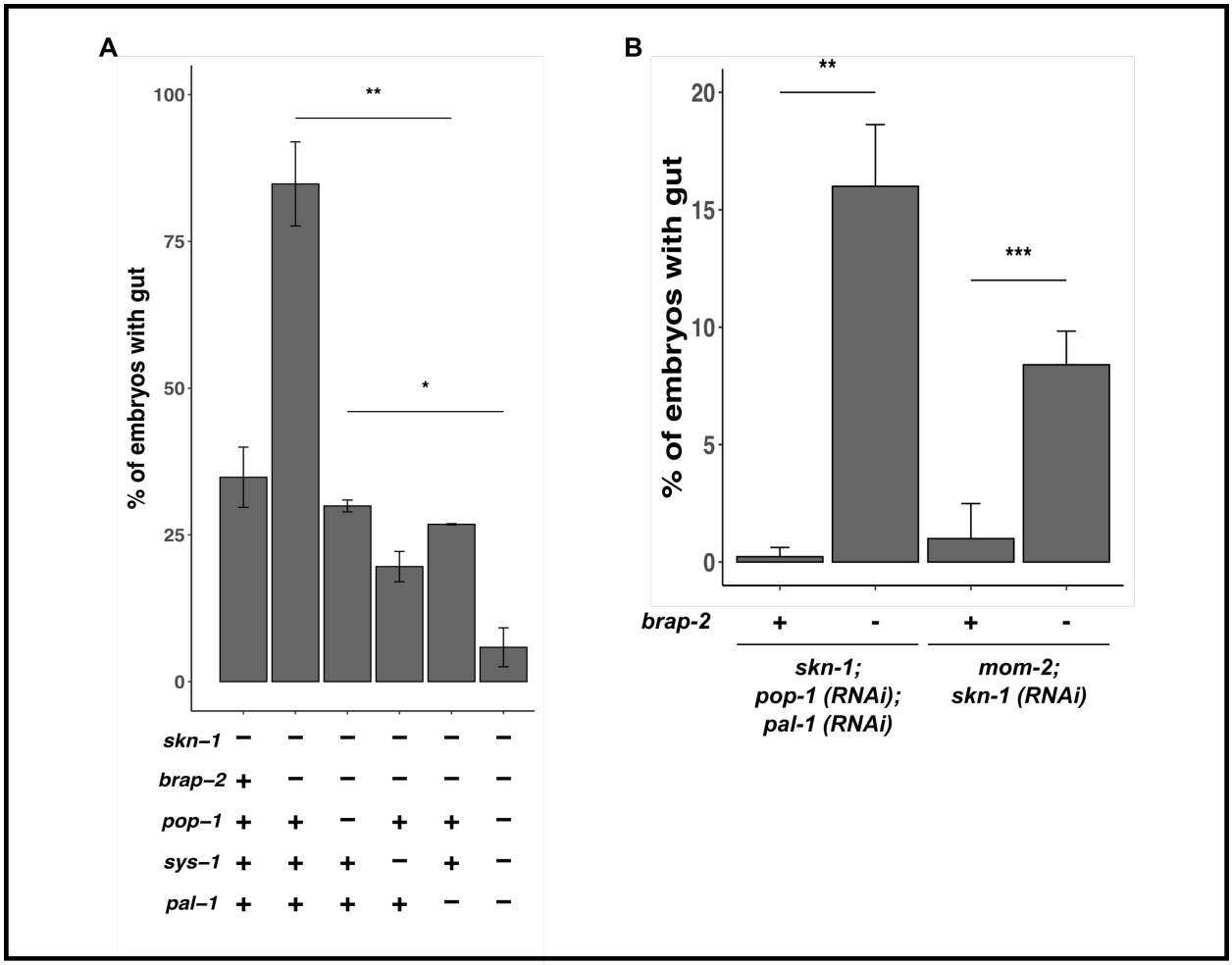

Figure 10 *brap-2(-)* suppression of *skn-1(RNAi)* is dependent on alternate activators

A) *brap-2(tm5132)* rescue is dependent on pop-1, sys-1, and pal-1. Below the graph, *skn-1* and *brap-2* refer to genetic mutants zu67 and tm5132, respectively. '+' refers to the WT copy. For *pop-1, sys-1,* and *pal-1*, '-' refers to RNAi for said gene. For the triple RNAi in the last column, RNAi cultures were mixed in equal proportions after the 4-hour culture and ITPG addition, right before seeding the plates (see Methods and Materials). N > 145 embryos for all conditions. B) *brap-2(tm5132)* can rescue *skn-1*(zu67) on pop-1 + pal-1 *(RNAi)* as well as mom-2(or42) on *skn-1 (RNAi)*. N > 200 embryos for each condition.

BRAP-2 modulates POP-1 asymmetry

To test whether BRAP-2 affects nucleocytoplasmic distribution of POP-1, we crossed *med-1p::POP-1::GFP* into *brap-2(-)* mutants. This construct shows a strong POP-1::GFP signal in the EMS descendants, allowing assessment of POP-1 localization in MSa, MSp, Ea, and Ep cells (Maduro et al., 2002) (Fig. 11A, Materials and Methods). In both wildtype and *brap-2(-)* mutant, the anterior sister cells contain higher total amounts of POP-1::GFP than the posterior sister cells, although *brap-2(-)* mutants appear to have slightly (non-statistically significant) weaker signal across all cells measured (Fig. 11B & C). Importantly, *brap-2(-)* mutants showed significantly decreased nuclear concentrations of POP-1::GFP (Fig. 11E), accompanied by a slight increase in the cytoplasmic signal (Fig. 11F), resulting in an overall lower nuclear/cytoplasmic fluorescence ratio across all four cells analyzed (Fig. 11D). Hence, we conclude that depletion of BRAP-2 leads to enhanced nuclear export of POP-1. Activation of Wnt signaling results in the nuclear export of POP-1. It is possible that depletion of brap-2 results in hyperactivation of the Wnt pathway. To test this, we examined the levels of MOM-2/Wnt in *brap-2(-)* mutants. We see no difference in levels of MOM-2::GFP in *brap-2(-)* mutants compared to WT (Fig. 12A & B). If the Wnt pathway is hyper activated in *brap-2(-)* mutants it would appear to be downstream of mom-2 transcription. To test this, we examined the expression and localization of SYS-1/β-catenin, the Wnt-signaled transcriptional co-activator for POP-1. We also see no difference between *brap-2(-)* and wildtype for SYS-1 expression or localization, suggesting that *brap(-)* does not result in hyperactivation of the Wnt pathway (Fig. 13). Alternatively, BRAP-2 could be regulating nuclear POP-1 levels through an independent mechanism in conjunction with kinases (e.g. LIT-1) analogous to that observed during oxidative stress (D'Amora et al., 2018; Q. Hu et al., 2017).

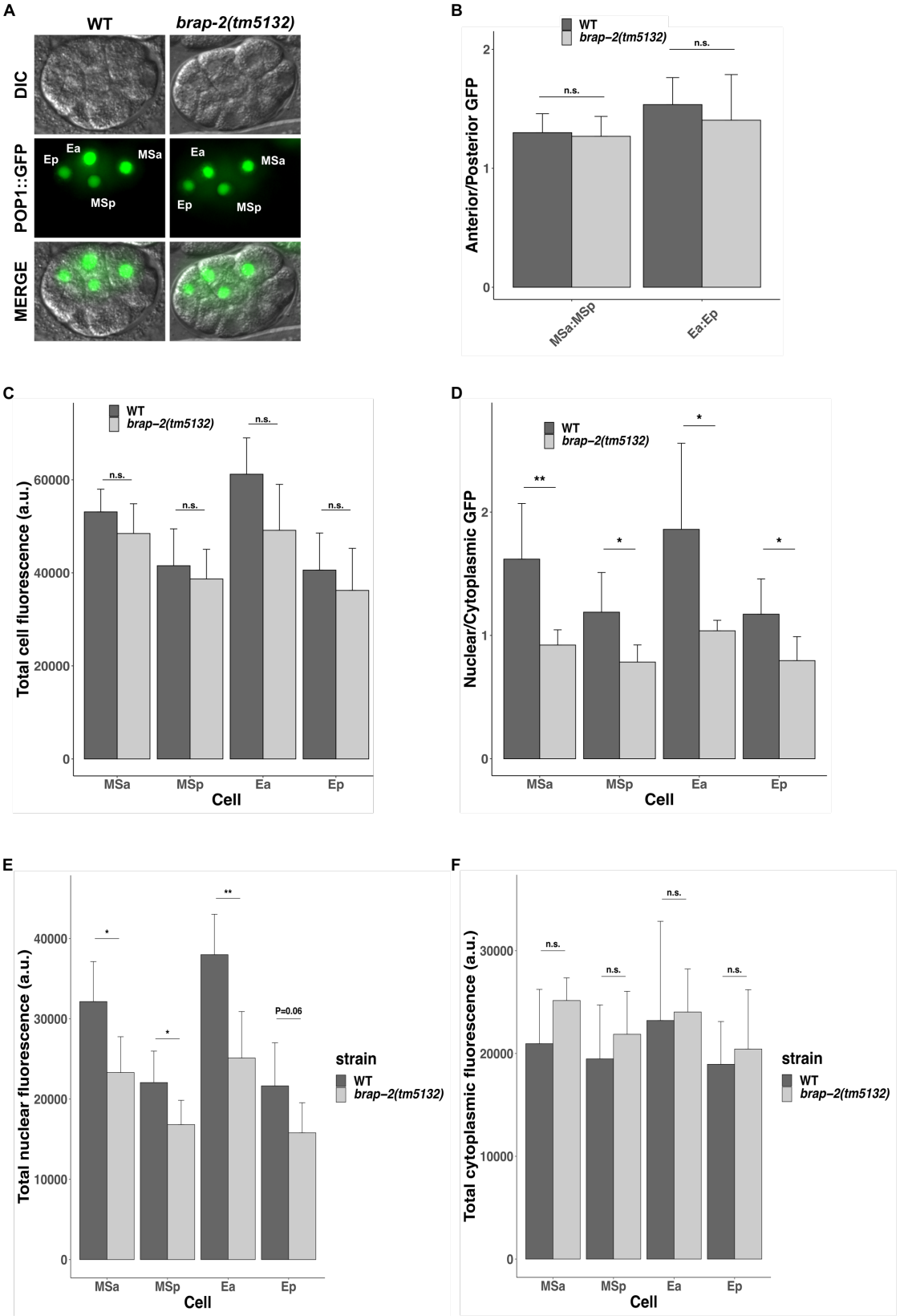
A
WT
brap-2(tm5132)
DIC
POP1::GFP
MERGE
Ea
Ep
MSa
MSp
B
WT
brap−2(tm5132)
Anterior/Posterior GFP
n.s.
MSa:MSp
Ea:Ep
C
Total cell fluorescence (a.u.)
Cell
MSa
MSp
Ea
Ep
D
Nuclear/Cytoplasmic GFP
E
Total nuclear fluorescence (a.u.)
P=0.06
strain
F
Total cytoplasmic fluorescence (a.u.)

Figure 11 *brap-2(-)* affects nuclear localization of POP-1

A) *med-1*-driven GFP::POP-1 expression in WT and *brap-2(tm5132)* embryos. DIC images show a single plane, and GFP images show the max intensity z-stack projection. Contrast levels were adjusted in FIJI/ImageJ to better show the cytoplasmic GFP. B) Anterior total cell GFP/posterior total cell GFP in med-1-driven GFP::POP-1 embryos. N = 5 for the WT strain, and N = 6 for *brap-2(tm5132)*. C) Corrected total cell fluorescence was measured in MSa, MSp, Ea, and Ep (see Methods and Materials). N = 5 for the WT strain, and N = 6 for *brap-2(tm5132)*. D) Ratio of nuclear GFP/cytoplasmic GFP in med-1-driven GFP::POP-1 embryos. N = 5 for the WT strain, and N = 6 for *brap-2(tm5132)*. E) Total nuclear fluorescence in med-1-driven GFP::POP-1 expressing embryos. F) Total cytoplasmic fluorescence in med-1-driven GFP::POP-1 expressing embryos. Cytoplasmic fluorescence intensity was calculated by subtracting nuclear fluorescence intensity from whole cell intensity (see Methods and Materials). N = 5 for WT, and N = 6 for *brap-2(tm5132)*.

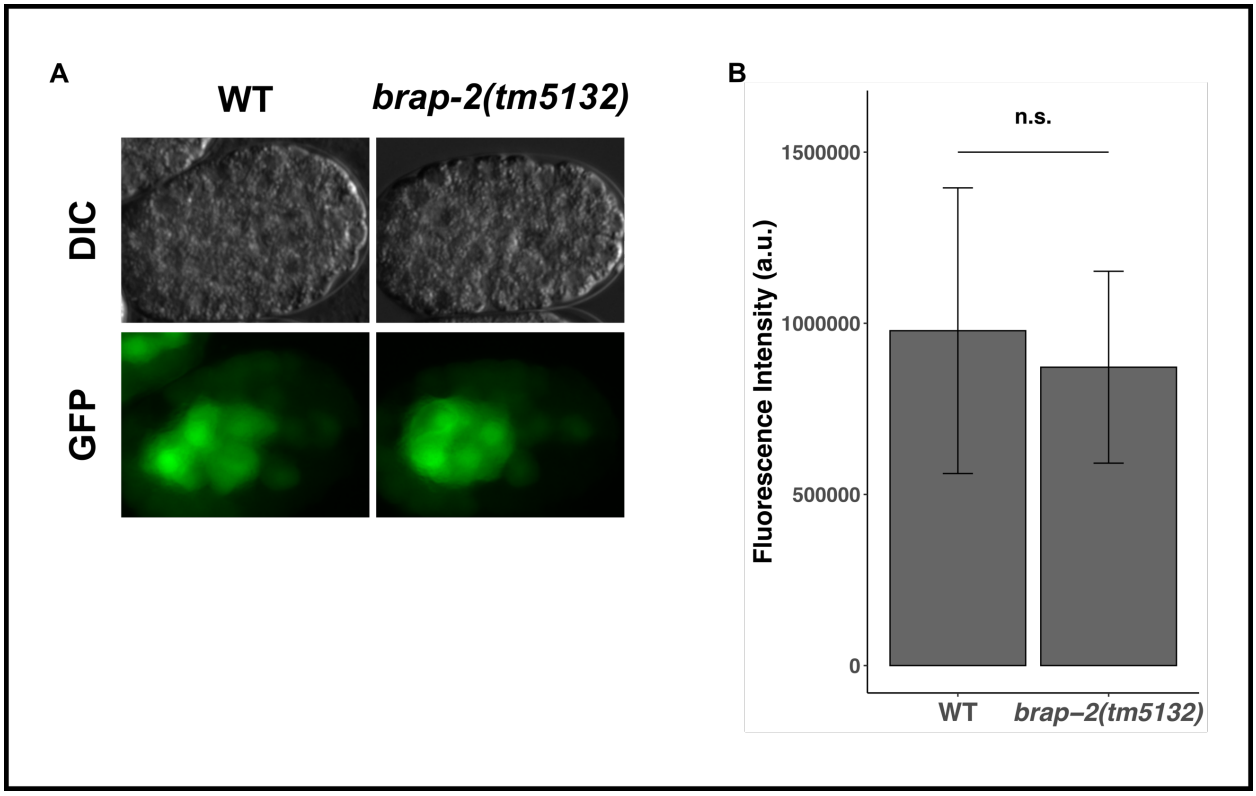

Figure 12 *brap-2(-)* does not affect *mom-2* expression

A) Zygotic mom-2::GFP expression in WT and *brap-2(tm5132)* embryos. B) Quantification of fluorescence pictured in G. Only similarly staged embryos were used for analysis. N=8 for *brap-2(tm5132)*, and N=19 for WT. Standard deviation is shown on all bar graphs, and significance was determined using a Student's t-test.

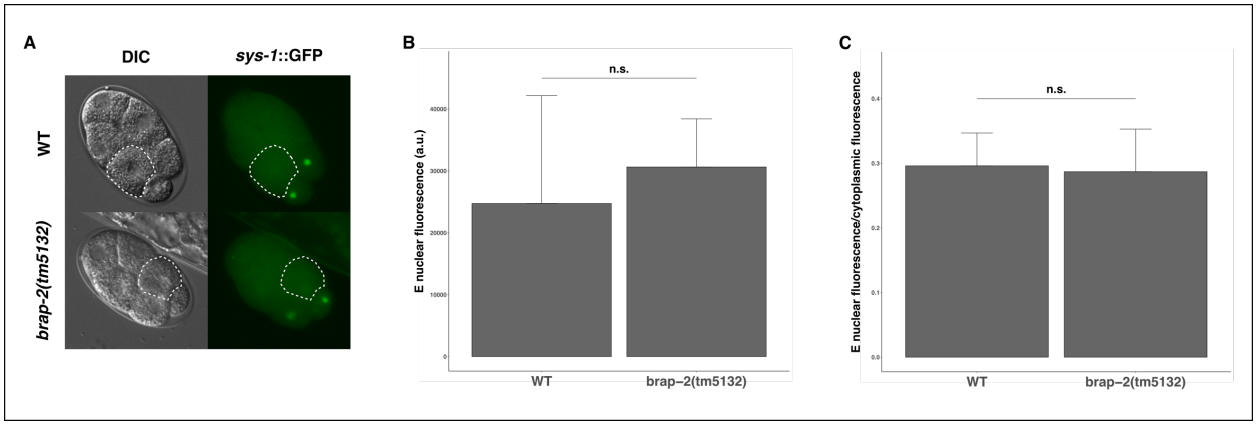

Figure 13 *sys-1*::GFP expression is not altered in *brap-2(-)*

A) DIC and GFP images of WT and *brap-2(tm5132)* embryos expressing *sys-1*::GFP. The E cell is outlined with dashes. The marker is highly dynamic in early embryos and localizes to centrosomes right before and after cell divisions. B) Bar graph showing no significant difference between WT and *brap-2(tm5132)* in the fluorescence intensity of *sys-1*::GFP in the E cell nucleus. Error bars show standard deviation. C) Bar graph showing no significant difference between WT and *brap-2(tm5132)* in the fluorescence intensity of *sys-1*::GFP in the E cell nucleus divided by the fluorescence seen in the cytoplasm. Error bars show standard deviation.

***brap-2(-)* shows contradictory effects on the PAL-1 muscle-specification pathway**

As the *brap-2(-)* suppression of *skn-1(-)* in the endoderm development pathway is also partially dependent on muscle transcription factor PAL-1, we hypothesized that loss of BRAP-2 may increase PAL-1 nuclear localization. To test this, we looked at the expression of two different markers downstream of PAL-1 in muscle development. PAL-1 directly activates *hlh*-1, the only myogenic regulatory factor family-related gene in *C. elegans* (Fukushige & Krause, 2005; Lei et al., 2009). We find that, actually, *hlh-1* expression is significantly decreased in *brap-2(-)* embryos (WT- 145204.1 ± 44058.80 a.u of fluorescence, *brap-2(-)*- 90677.3 ± 26593.57 a.u of fluorescence, p-value<0.001) (Fig. 14A & B). *myo-3*, a myosin structural gene, is downstream of HLH-1 (Fukushige & Krause, 2005), and, in contrast to our previous results, *myo-3* expression is actually significantly increased in *brap-2(-)* L1 larvea (WT- 2472550 ± 1097918 a.u of fluorescence, *brap-2(-)*- 3984436 ± 1728466 a.u of fluorescence, p-value<0.01) (Fig. 14C & D). Further investigation is warranted to analyze if PAL-1 activity is altered in *brap-2(-)* and if this affects some sort of transcriptional negative feedback loop involving HLH-1 and MYO-3.

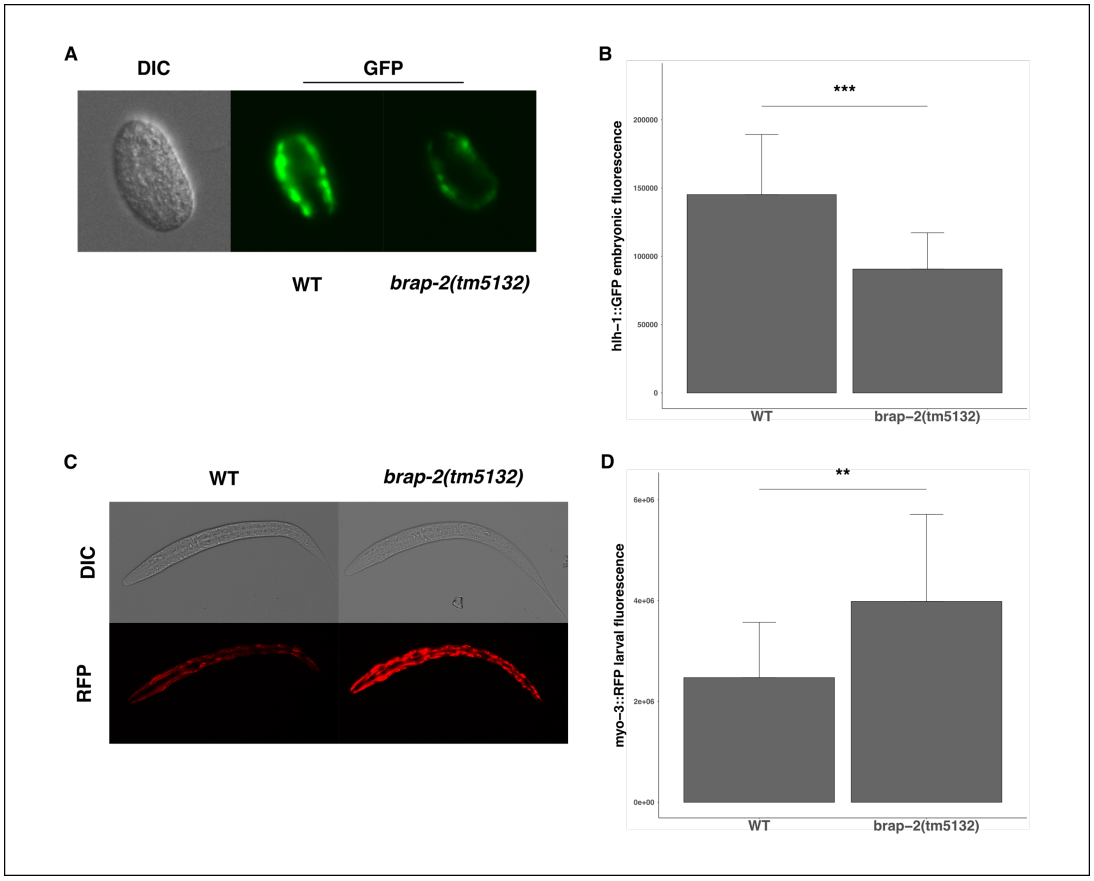

Figure 14 *brap-2(-)* mutants show contradictory *hlh-1* and *myo-3* expression changes

A) DIC image of a WT embryo showing the stage at which all embryos for A were photographed (and subsequently measured for B). GFP images showing embryonic *hlh-1*::GFP expression in WT and *brap-2(tm5132)* strains. B) Bar graph showing the quantification of *hlh-1*::GFP intensity in WT and *brap-2(tm5132)* embryos. Error bars show standard deviation, and asterisks indicate significance found using a Student's t-test. C) DIC and RFP images of WT and *brap-2(tm5132)* L1 larvae expressing *myo-3*::RFP. D) Bar graph showing the quantification of *myo-3*::RFP intensity in WT and *brap-2(tm5132)* L1 larvae. Error bars show standard deviation, and asterisks indicate significance found using a Student's t-test. *myo-3* is downstream of *hlh-1* when specifying muscle, and it is unexpected

that *brap-2* mutants would show increased *hlh-1*::GFP expression yet decreased *myo-3*::RFP expression when compared to WT.

3.4 Discussion and conclusions

In mammals, BRAP has been found to repress Ras-dependent activation of ERK/MAPK signal transduction (Matheny et al., 2004). Additionally, BRAP may function as a cytoplasmic retention factor for many signaling effectors (Asada et al., 2004; Davies et al., 2013; Fatima et al., 2015; Fulcher et al., 2009; Takashima et al., 2013). In *C. elegans*, BRAP-2 inhibits nuclear localization of SKN-1 through PMK-1 phosphorylation, and enhances DAF-16 nuclear localization, likely through inhibition of AKT-1 through PHLP-2 (D'Amora et al., 2018; Q. Hu et al., 2017). In the present study, we show that BRAP-2 additionally negatively regulates endoderm specification by modulating the nuclear localization of the Wnt effector POP-1. Depleting the function of BRAP-2 rescues loss of endoderm in *skn-1* deficient embryos through increased activation of *end-1* in the E lineage, and this rescue is dependent on endoderm activators POP1, SYS-1, and PAL-1. Our study revealed a novel role of BRAP-2 in cell fate specification and provided a new insight into how signaling activities are tuned to ensure a robust developmental outcome.

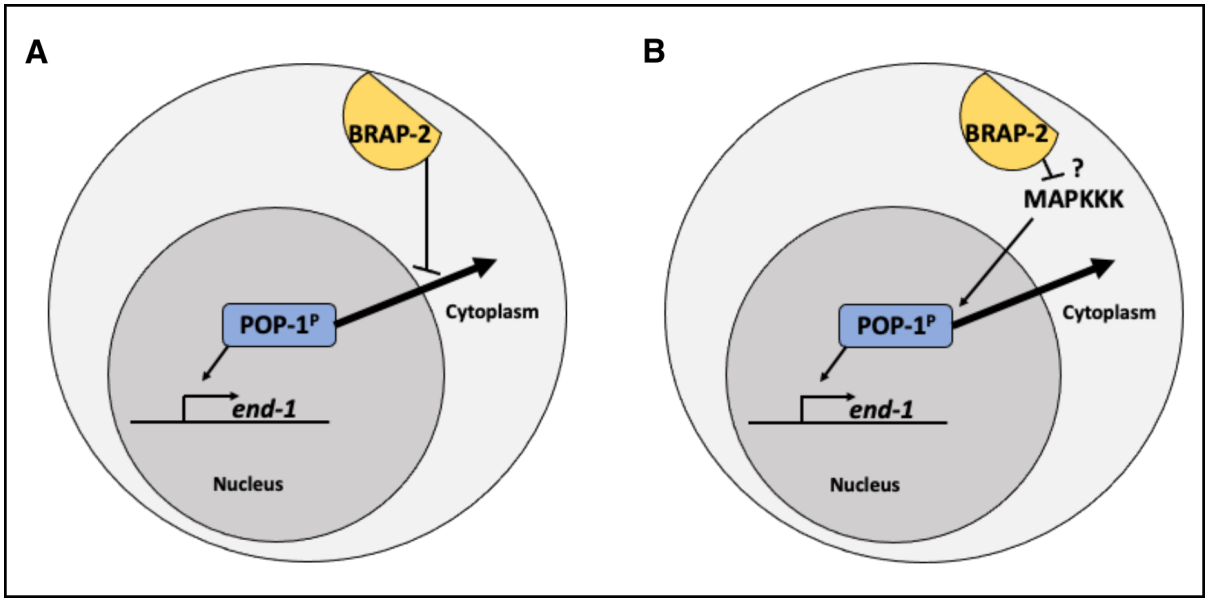

Figure 15 Model of BRAP-2 action in E lineage

A) BRAP-2 prevents nuclear export of POP-1^{P}, allowing for increased activation of POP-1^{P} target, *end-1*. B) BRAP-2 may be acting on factors upstream of POP-1, such as the MAPKKK pathway.

Duplicated GATA factors (MED-1/2, END-1/3 and ELT-2/7) and redundant signaling systems ensure threshold levels of gene expression is achieved and developmental outcome is reproducible (Raj et al., 2010). In addition, gut-inducing signals are tuned by numerous modulators, such as FRK-1/Fer-type nonreceptor kinase, which attenuates the effect of HMP-2/β-catenin (Putzke & Rothman, 2010). We recently reported that βNAC ICD-1 downregulates SKN-1-mediated endoderm specification (Ewe et al., 2019). Similarly, OSM-11 and RICT-1 antagonize SKN-1 activity during embryonic development (Dresen et al., 2015; Ruf et al., 2013). Interestingly, RICT-1 may exert an opposing effect on Wnt signaling, fine-tuning the activating signals in the E cell, although the mechanism remains unclear (Torres Cleuren et al., 2019). In this study, we further show that BRAP-2 negatively

regulates endoderm specification, and depleting BRAP-2 significantly rescues the endoderm defects of *skn-1* deficient embryos. Suboptimal expression of the *end* genes may lead to hypo- or hyperplasia and disruption of gut functions (Maduro et al., 2015). Taspase,TASP-1, and a prion-like glutamine/asparagine rich protein, PQN-82, further modulate the expression level of ELT-2, and TASP-1 can rescue endoderm in *skn-1* knockdown (Wiesenfahrt et al., 2018). These findings reveal the importance of quantitative regulation of developmental signals within a restricted developmental window by multiple partial redundancies acting at multiple levels in a rapidly developing embryo.

Modulation of Wnt signaling has impacts for cancer, especially considering that 90% of all colorectal cancers show hyperactivation of the Wnt pathway (Bienz & Clevers, 2000; Fodde et al., 2001; Giles et al., 2003; Kinzler & Vogelstein, 1996; Segditsas & Tomlinson, 2006). This is perhaps not surprising as there are also reports of the importance of Wnt signaling in the development and homeostasis of the intestine (Fevr et al., 2007; Korinek et al., 1998; Tian et al., 2015; van Es et al., 2005), with Wnt target Lgr5 (leucine-rich-repeat-containing G-protein-coupled receptor 5) serving as a marker for stem cells in the small intestine and colon in mice (Barker et al., 2007). Wnt gene expression helps to define the intestinal stem cell crypt (Van der Flier et al., 2007). Therefore, Wnt signaling thresholds in intestinal stem cells must be finely regulated. Given the conservation of Wnt signaling in gut development across metazoa, it will be interesting to see if BRAP can modulate intestinal Wnt signaling in mammalian systems, as well.

Our results indicate that BRAP-2 modulates nucleocytoplasmic distribution and activity of POP-1: *brap-2* deletion mutants contains lower levels of nuclear levels of POP-1 and slightly higher level of cytoplasmic POP-1. Although not statistically significant, *brap-2*

mutants show slightly lower level of POP-1::GFP in all cells analyzed (Fig. 11C). It is possible that cytoplasmic POP-1 is more prone to degradation, leading to overall lower levels in the whole cell. In canonical Wnt systems, NLK inhibits TCF/LEF signaling through phosphorylation (Ishitani et al., 1999), and it was shown in Xenopus that NLK activity enhanced ubiquitination and subsequent degradation of TCF/LEF (M. Yamada et al., 2006). In *C. elegans* Wnt/β-catenin asymmetry pathway (a divergent canonical Wnt signaling), the NLK orthologue LIT-1 normally works with WRM-1/β-catenin to phosphorylation and cause nuclear export of POP-1. Perhaps increased LIT-1/WRM-1 activity also results in degradation of POP-1 in this pathway. It is currently unclear whether BRAP-2 interacts directly with POP-1 or if it modulates the activity of LIT-1/WRM-1 upstream (Fig. 15).

An atypical MAPKKK activates LIT-1/NLK, and the possibility that BRAP-2 modulates Wnt signaling through repressing MAPKKK activity should be explored. *brap-2(ok1492)* mutants show increased levels of ERK-1/2 orthologue MPK-1 and increased phosphorylation of p38 MAPK PMK-1 (D'Amora et al., 2018; Q. Hu et al., 2017). BRAP can act as a scaffold protein for AKT phosphatase PHLPP1 in mouse testis (Fatima et al., 2015), and *C. elegans* BRAP-2 can interact with PHLPP1/2 orthologue PHLP-2 in vitro (D'Amora et al., 2018). In addition, mammalian BRAP can inhibit ERK signal transduction by binding scaffold KSR1 and preventing interactions with Ras. In stimulated cells however, BRAP can bind Ras and actually help facilitate kinase signaling (Chen et al., 2008; Matheny et al., 2004; Matheny & White, 2009; Ory & Morrison, 2004; Shoji et al., 2017). As evidence suggests that BRAP and its orthologues can modulate chief phosphorylation-based signaling systems, future studies should examine LIT-1 and POP-1 phosphorylation levels in *brap-2(-)* mutants.

Chapter 4

Largescale mutagenesis, sorting, and bulk sequencing of *C. elegans* embryos reveals *act-4* as a novel modulator of the endoderm gene regulatory network.

(in preparation for publication with minor adjustments)

4.1 Introduction

Traditional forward genetic screens in *C. elegans* involve mutagenesis and observation of F2 offspring (Jorgensen & Mango, 2002; Zuryn & Jarriault, 2013). However, achieving saturation with this method is laborious and non-comprehensive, as several candidate mutations inevitably fall through the screen.. To address these issues, we developed a new screen that involves sequencing large numbers of mutants in bulk and applied this strategy to study the gene regulatory network (GRN) governing endoderm development in *C. elegans*. Complex GRNs govern the specification and behavior of cells as they develop from a single-cell embryo. High resolution expression data and an invariant cell lineage make *C. elegans* an ideal model system to study GRNs and their reliable governance of organ development. Through careful dissection and analysis of the endoderm GRN, we provide further insight into the complexity of GRNs generally in both healthy and diseased states.

C. elegans endoderm, consisting of 20 cells, is derived from one blastomere cell—the E blastomere--which is specified at the 4-cell embryo stage (J. E. Sulston et al., 1983). Signals from P_2 blastomere, Wnt, MAPK, and Src tyrosine kinase, converge on the precursor, EMS, instructing it to divide into E and MS along the Anterior-Posterior axis (Bei et al., 2002; Goldstein, 1992, 1993, 1995; Ishitani et al., 1999; Meneghini et al., 1999; Rocheleau et al., 1997; Schierenberg, 1987; Thorpe et al., 1997). These signals relieve repression by TCF/LEF

POP-1 of endoderm GATA factors, *end-1/3*, and turn POP-1 into an activator of endoderm fate (Gina Broitman-Maduro et al., 2005; Lin et al., 1995; Maduro et al., 2002).

The maternally provided SKN-1 transcription factor initiates the mesendoderm GRN, activating GATA factors *med-1/2* which in turn activate GATA factors *end-1/3* in E cell (B. Bowerman et al., 1992; Maduro et al., 2001; Zhu et al., 1997). While SKN-1 is also present in mesectoderm and germ lineages, it is repressed by PAL-1 and PIE-1. Relieving repression of SKN-1 can result in excess endoderm in place of other cell lineages (B. Bowerman et al., 1993; Hunter & Kenyon, 1996; Lin et al., 1995; Maduro et al., 2001; Mello et al., 1992; Seydoux et al., 1996). In addition, several separate factors have been identified that can promote transcription of *end-1/3* in the absence of SKN-1 (Maduro et al., 2005; Ruf et al., 2013; Witze et al., 2009). The plasticity of this network and the modulators that result in the robust development of the gut is an area of active investigation.

In this study, we investigate the mechanisms underlying robust gut development using a high throughput genetic mapping strategy. We selected hundreds of embryos producing excess endoderm from a mutagenized population using fluorescence activated cell sorting (FACS). These samples were pooled and sequenced in bulk using NextGen whole genome sequencing (WGS) followed by bioinformatic analyses. Using this method, we were able to identify candidates that would have been missed with traditional methods. This includes *act-4*, an ortholog of human ACTB that was found to play a role in endoderm development. This rapid mapping strategy of bulk-sequencing in combination with FACS can be readily adapted to identify other genetic regulators of embryogenesis.

4.2 Materials and Methods

Strains: JR3920 [*syIs243 (myo-3p::TOM20::mRFP + unc-119) ; rrIs1 (elt-2p::GFP) X*] was used for the mutagenesis and for the control *(RNAi)* image in Figure 3C. N2 WT and JR1128 [*wIs82 (elt-2::GFP + pRF4 rol-6)*] were used as FACS controls. JR4083 [*itIs(pie-1p::mcherry::H2B::pie-1 3'UTR +unc119(+)) IV ; rrIs1(elt-2p::GFP) X]* was used for the *act*-4 *(RNAi)* image in Figure 3C.

Microscopy: Images were taken on the Nikon Eclipse Ti using NIS-Elements AR 4.13.05 software and images were processed using FIJI/ImageJ. (microscopy descriptions in G3 screens are very minimal, some missing completely, so I'm not inclined to add exposure times).

Mutagenesis: Approximately 200,000 L4 *C. elegans* (JR3920) were mutagenized with 50mM EMS for 4 hours at 20°C with constant rotation. They were re-plated and allowed to lay eggs for 1-2 days. Approximately 200,000 L1s were collected and re-plated while arrested F1 embryos were collected to be sorted on the FACS machine. This was repeated for the F3 generation.

RNAi treatment: L4 worms were fed RNAi clones with the L4440 vector backbone in HT115 cells for at least 24 hours at 20°C. Worms were then either synchronized (as previously described (J. Sulston & Hodgkin, 1988)) to extract embryos or allowed to lay eggs on a new plate before being removed. Embryos were then scraped from the plate using a modified

glass Pasteur pipette. For *act-4* vs control *(RNAi)* endoderm cell counts, significance between the samples was assessed using the Mann-Whitney-Wilcoxon test.

C. elegans embryo sorting for controls: L4s were treated with either L4440 control or *pop*-1 *(RNAi)*. Mixed stage adults were bleached and washed according to standard synchronization protocol (J. Sulston & Hodgkin, 1988), and embryos were left spinning at room temperature in M9 buffer for two hours before being placed on ice. Embryos were strained through a 40μM cell strainer and resuspended in PBST. Embryos were sorted using a BD FACSAria II system using a 100μM nozzle at low sheath fluid pressure (20 psi).

C. elegans embryo sorting from mutagenesis: Arrested embryos were scraped from plates and washed several times with M9 buffer. Embryos were strained through a 40μM cell strainer and resuspended in PBST. Embryos were sorted using a BD FACSAria II system using a 100μM nozzle at low sheath fluid pressure (20 psi).

DNA preparation + sequencing: DNA from sorted embryos was extracted using standard phenol-chloroform extraction protocols (Sambrook & Russell, 2006). Library preparation and next-generation sequencing was carried out by the McDonnell Genome Institute at Washington University. Each generation (F1 and F3) was differentially barcoded. Samples were run on an Illumina NovaSeq and sequenced using 150bp paired-end reads. Collaborators at WashU aligned the sequencing data to the WEBcel235 version of the *C. elegans* genome.

Sequence analysis: Aligned BAM files were received from WashU and filtered for read (Q15) and mapping quality (q15) using BCFtools 1.9 from SAMtools. Coverage was limited to the mean coverage of the sample ±50% of the mean. For 1Kbp and 5Kbp sliding windows, the genome was parsed into bins 1/20th the size of the window, and mutant density, coverage, and coverage variance (Interquartile Range [IQR]/Median) were analyzed in each using a custom python script. Rare variants were called based on representing only 1-2 reads for a given base. Windows were comprised of averages of the bins, and mutant allele densities divided by coverage IQR/Median (to control for variance) were plotted using the middle of a given window. Given that the sequence coverage varies across the genome and that coverage affects the ability to spot rare mutations, we increased the signal to noise ratio by dividing the mutant density for a given window by the variation in coverage depth across that window assessed through the interquartile range divided by the median. We refer to this variable as the normalized mutant density of a window. Regions of interest were called by gathering all windows that were at least 3 standard deviations greater than the mean normalized mutant density. Further analysis looked to at least 8 standard deviations greater than the mean. To select for genes, the number of protein-coding changes were identified among these rare variants using Ensembl Variant Effect Predictor, and the number of protein-coding changes was divided by the longest mRNA length known, according to Wormbase.org. Genes were tested for significance using a 1-sample binomial test and a Bonferroni corrected p-value. GO term analysis was run on The Gene Ontology Resource at Geneontology.org using gene IDs from Wormbase.org.

Gene lists and GO enrichments: All gene lists and GO enrichments mentioned in this chapter are available on the Rothman lab NAS (Network Attached Storage), hosted by University of California, Santa Barbara. The following lists can be found there:

1Kbp window, 3sd above mean, in both generations (410 genes)

1Kbp window, 3sd above mean, in the F3 (1910 genes)

1Kbp window, 5sd above mean, in the F3 (153 genes)

1Kbp window, 8sd above mean, in the F3 (13 genes)

5Kbp window, 3sd above mean, in the F3 (768 genes)

cDNA analysis, significant in both the F1 and F3 (53 genes)

GO term analysis of genes found in cDNA analysis in both samples (53 enrichments)

cDNA analysis, significant in either the F1 or F3 (548 genes)

GO term analysis of genes found in cDNA analysis in either sample (227 enrichments)

cDNA analysis, significant in F1 (388 genes)

4.3 Results and Discussion

High-throughput phenotypic screening using FACS.

It has previously been shown that it is possible to sort *C. elegans* embryos using FACS (Burdick et al., 2016; Stoeckius et al., 2009) . Here we report the ability to sort embryos based on excess endoderm, using a differentiation marker, *elt-2p*::GFP, as a readout. ELT-2 is a direct target of END-1/3, and early specifier for endoderm fate (T. Fukushige et al., 1998; Tetsunari Fukushige et al., 1999; McGhee et al., 2009). Embryos contain autofluorescent gut granules even in the absence of the GFP reporter, and group as one population (Fig 1A). Using GFP expression levels from *elt-2p::GFP* in mixed stage embryos

as a control, we were able to set a conservative gate for any embryos expressing GFP above expected levels, which is greater than 10^4 B 530/30-H (GFP emission height measurement, Fig. 1A x-axis). To confirm the robustness of the 'excess gut' gate, we repeated the experiment with arrested embryos from *pop-1 (RNAi)* in which MS is mis-specified to form E, producing excess endoderm (Lin et al., 1995; Maduro, Hill, et al., 2005; Rocheleau et al., 1997). Mixed stage embryos from these RNAi-treated worms were analyzed on the FACS machine and sorted using the 'excess gut' gate. The mixed population shows many under-developed embryos that do not show any GFP expression (as is expected since *pop-1 (RNAi)* results in a population of embryos arresting before gut development commences?), along with embryos showing various levels of GFP expression. Importantly, we see a portion of the embryos with higher GFP expression than is seen in control *(RNAi)* conditions. The phenotype of the embryos sorted in the 'excess gut' gate were examined manually using fluorescent microscopy and approximately 65% of embryos in that gate were confirmed to have more than the expected 20 *elt-2p::GFP* positive cells (Fig. 1A-C). Why it is that less than 100% of the embryos from the 'excess gut' gate have extra gut cells could possibly be because the process of sorting and collection caused degradation of part of the sample before it could be scored for number of gut cells by hand.

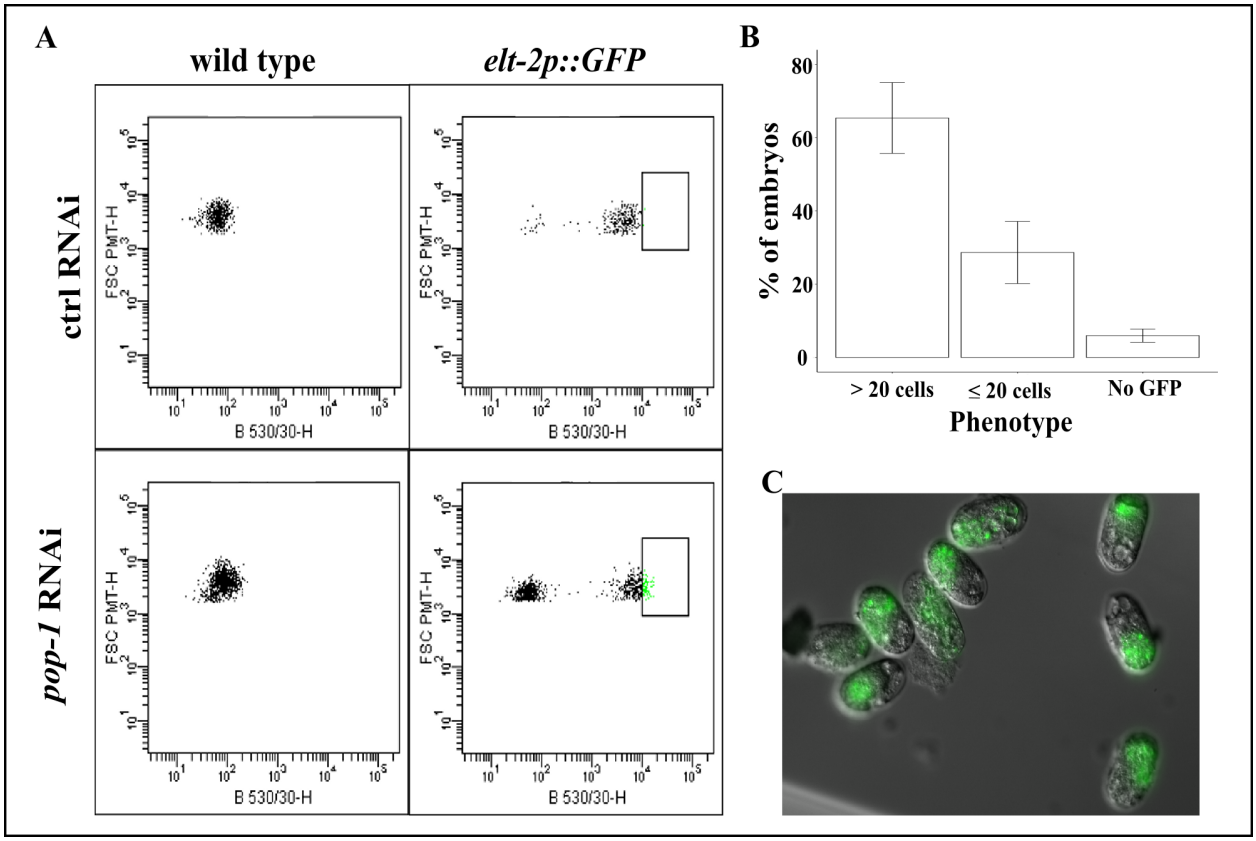

Figure 1 Embryos with extra intestine can be collected via FACS

A) FACS profile of L4440 control and *pop-1 (RNAi)* in N2 WT and *elt-2p::GFP* embryos. X-axis shows 530/570 GFP emission. Y-axis shows forward scatter FSC. The rectangle shows the gate used to collect embryos with excess endoderm. B) Phenotypes of embryos collected in the sort of *elt-p::GFP* embryos under *pop-1 (RNAi)* (n= 222 embryos, error bars show SD). C) *elt-2p::GFP* in embryos collected in the sort of *elt-2p::GFP* embryos under *pop-1 (RNAi)*.

We then mutagenized populations of late L4 (fourth larval stage) worms with ethyl methanesulfonate (EMS) and collected arrested embryos at the F1 and F3 generations that fell within the 'excess gut' gate, above 10^4 B 530/30-H GFP emission (Fig. 2A). At each generation, adults were allowed to lay eggs overnight before being washed off. After the

adults were removed, the plates grew for 24 hours at 15°C, and then larvae were washed from the plates and arrested embryos were collected. In the F1, 18,651 arrested embryos were sorted, and 473 embryos (2.54%) were collected within the 'excess gut' gate. 42,356 arrested embryos were sorted in the F3, and 201 (1.08%) were collected in the 'excess gut' gate (Fig. 2B). We sequenced the genomes of the mutants from each collection in bulk.

Sequencing analysis using sliding windows reveals positive hits

We used ~96% and ~93% of the genome in the F1 and F3, respectively, that remained after quality control (see Methods and Materials). To identify regions of the genome that are associated with the increased gut phenotype selected for in the sorting, we first found what were considered to be rare variants. These variants occurred at low frequency, usually in only one or two reads, and represent unique mutations from the mutagenesis. We found 1,687,091 rare mutations in the F1 bulk and 891,044 rare mutations in the F3 bulk which represents 1.752% and 0.949% of sites in the F1 and F3 pools respectively. Considering our average coverage in each sample, we find 0.0529% and 0.0451% of the genome in the F1 and F3 respectively is mutated on average per embryo in the bulk, giving roughly around 50,000 mutant bases per embryo.

With both 1Kbp and 5Kbp sliding windows, we calculated the mutant density of each window as we scanned the genome and normalized the mutant density by variance in coverage (termed normalized mutant density) (see Methods and Materials) (Fig. 2C). To identify candidates, every window with a normalized mutant density three standard deviations above the mean of that sample was collected and overlapping bins were condensed together. (F1—sliding window 1Kbp: 0.2074763 ± 0.070437; sliding window 5Kbp: 0.1075844 ± 0.02070262; F3—sliding window 1Kbp: 0.09517261 ± 0.04505752;

sliding window 5Kbp: 0.04924972 ± 0.01331044). We gathered genes that are either fully or partially located within each window.

Using a 1Kbp sliding window in our initial analysis, we found a total of 410 significant genes, two of which, *hmp-2* and *apr-1,* were significant in both generations. APR-1 is an orthologue of human APC, a regulator of Wnt signaling, and HMP-2 is one of four *C. elegans* beta-catenins and is canonically involved in cell-to-cell adhesion in epithelial cells (Costa et al., 1998) but has been reported as a modulator of Wnt signaling in the E blastomere downstream of Src signaling (Putzke & Rothman, 2010; Sumiyoshi et al., 2011). Evidence suggests that APR-1 helps to limit HMP-2/Wnt signaling in the E lineage, and excessive Wnt signaling can cause hyperproliferation in the E lineage. Indeed, *apr-1 (RNAi)* modestly increases gut cell number—23.8+/-2.6 compared to 20.0+/-0.3 in WT (Putzke & Rothman, 2010). Additionally, in a mutant lacking the Fer-type nonreceptor tyrosine kinase FRK-1, HMP-2 translates into the gut nucleus and causes hyperplasia (Putzke & Rothman, 2010). We were encouraged in our analysis by the identification of known Wnt effectors that can cause increased gut cell number.

Using the same analysis, we identified *pop-1* and *pie-1* amongst the 1910 significant protein coding genes in the F3 bulk sample. Both *pie-1* and *pop-1 (RNAi)* knockdowns are known to produce excess endoderm in the next generation. *pie-1* R*(RNAi)* relieves repression of SKN-1 in germline and muscle lineages resulting in their differentiation into endoderm (Bowerman et al., 1993; Goldstein, 1995; Hunter & Kenyon, 1996; Mello et al., 1992; Seydoux et al., 1996). *pie-1* is retained in the F3 out of 768 genes when using a 5Kbp sliding window and in the top 153 protein coding genes (5 standard deviations above the mean) using a sliding window of 1Kbp. This genome-wide analysis identified many

intergenic windows as significant, pointing to either regulatory regions that were selected for or mutagenesis hot spots that were more susceptible to insults (Fig. 2D).

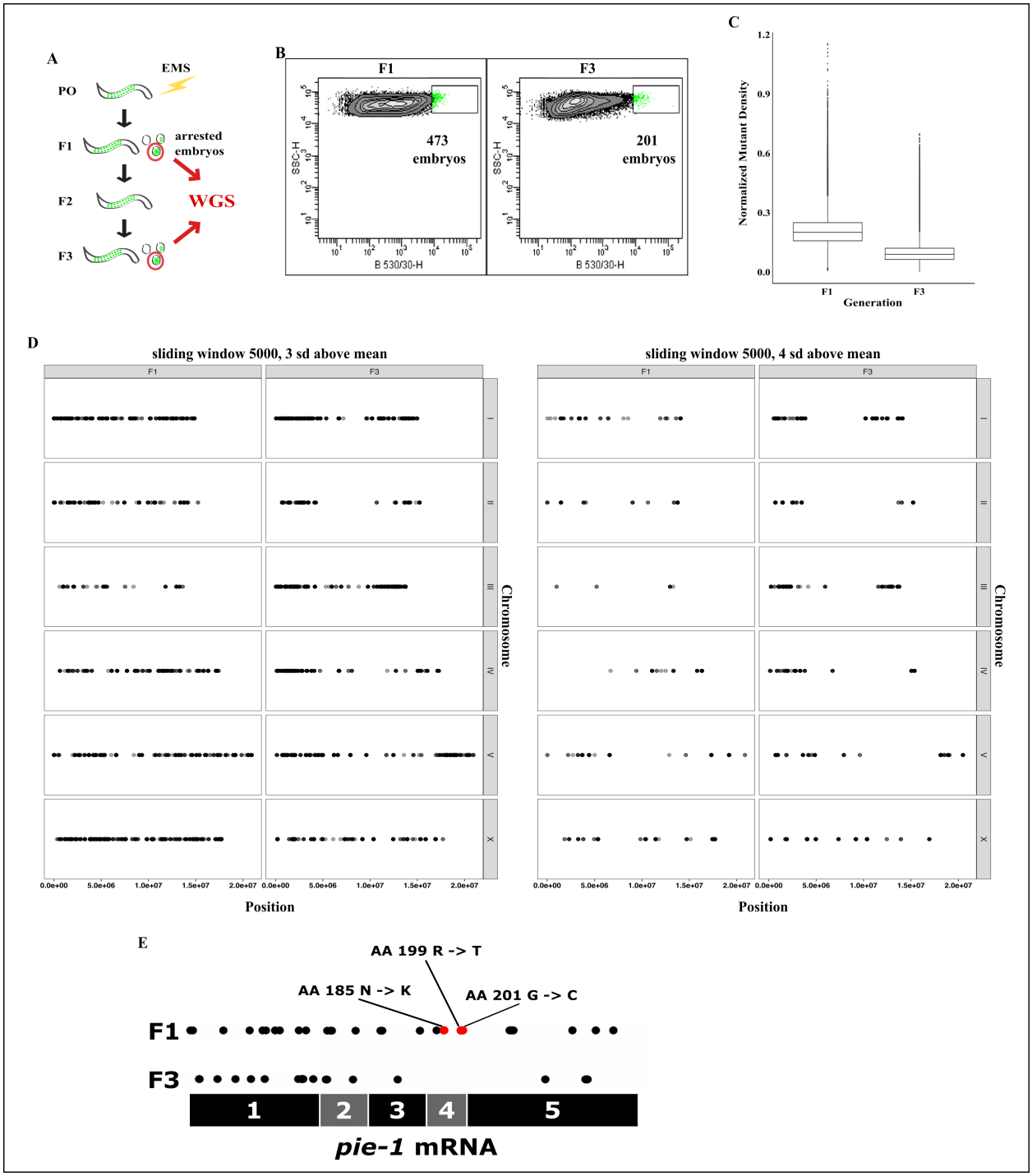

Figure 2 Mutagenesis, collection, and sliding window analysis

A) Schematic of experiment. Arrested embryos showing excess gut are circled in red. These embryos were collected via FACS and sequenced in bulk. B) FACS profiles showing embryo collection after mutagenesis. X-axis shows 530/570 GFP emission, and Y-axis shows side scatter SSC. The rectangles show the gate used to collect embryos with extra intestine, over 10^4 530/570-H GFP emission. Two consecutive collections in the F1 gave a total of 473 embryos, although only one collection is shown here. C) Boxplot showing the normalized mutant density (see Methods and Materials) for each 1Kbp sliding window in each generation. D) Plots showing the locations of 5Kbp sliding windows that were 3 standard deviations (left) and 4 standard deviations (right) above the mean when looking at normalized mutant density. E) Diagram showing the locations of protein-coding mutations in *pie-1* mRNA found in F1 and F3 embryos. Out of three isoforms of *pie*-1, two contain all five exons (Y49E10.14a.1 and Y49E10.14a.2), and one is limited to exon 5 (Y49E10.14b.1). Mutations in the second Zn finger are highlighted in red, and the nature of those three mutations are annotated.

cDNA analysis reveals gene networks involved, known and novel (need to break into sections)

We prioritized genes by filtering for only protein-coding changes in each gene divided by the size of the largest mRNA. A 1-sample binomial test was applied to the mutant allele densities to find genes that were significantly enriched for mutations.

Interestingly, *pie-1* was one of 54 genes found as significant in both the F1 and F3 using this analysis method. As mentioned before, *pie-1* is a maternal effect gene, so we wouldn't expect to see any selection for *pie*-1 in the F1 data. It is possible that dominant

mutants were coupled with early zygotic PIE-1 expression in some arrested embryos displaying extra gut. Closer analysis of the distribution of mutations in *pie-1* revealed that mutations in the second Zinc finger were limited to the F1 pool (Fig. 2E). This Zn finger has been implicated in activating gene expression in the germline, independent of *pie-1*'s repressive role on transcription (Tenenhaus et al., 2001). This Zn finger has also been shown to be sufficient to target PIE-1 to P granules (Reese et al., 2000). It is possible that mutations in this Zn finger lend themselves to dominant effects.

Gene Ontology Enrichment

548 unique protein-coding genes were found among both generations, and GO term analysis revealed an over 5-fold enrichment for axon guidance genes, which is consistent with previous findings that axon guidance genes modulate intestine morphogenesis in *C. elegans* (Asan et al., 2016). While GO term analysis for these 548 genes showed 227 unique GO term enrichments in total (Fig. 3A), only 18 terms were enriched over 10-fold. Among these 18 terms, we find semaphorin-plexin signaling at the top with an > 18-fold enrichment. Semaphorin-plexin signaling is implicated in the axon guidance networks that drive endoderm morphology as well as having conserved roles in cytoskeletal dynamics that drive many cell functions (Jongbloets & Pasterkamp, 2014). Semaphorin-plexin signaling is known to affect actin polymerization in the growth cone as both a repulsive or attractive signal (Goshima et al., 2002; Hinck, 2004; Jongbloets & Pasterkamp, 2014; Polleux et al., 2000). It also has roles in establishing morphology of epithelial and endothelial cells (Tran et al., 2007). Furthermore, it has been shown that plexin signaling in Xenopus and *C. elegans* can drive translation of ADF/cofilin, an actin-depolymerizing enzyme (Chisholm, 2008; Nukazuka et al., 2008; Piper et al., 2006). While semaphorin-plexin signaling is mostly

documented in driving actin regulation in axons, semaphorin signaling has been reported in other diverse cell types regulating a variety of functions, usually initially through regulating morphology (Alto & Terman, 2017; Goshima et al., 2002; S. Hu & Zhu, 2018).

We were further intrigued by an almost 17-fold enrichment for hemidesmosome assembly. Hemidesmosome-like structures in *C. elegans*, called fibrous organelles, help connect developing muscle to the cuticle, which facilitates the transduction of mechanical signals (Francis & Waterston, 1991; Pásti & Labouesse, 2018). The transduction of force signals can also regulate morphology and differentiation during development (Mammoto & Ingber, 2010). Coupled with the enrichment for semaphorin-plexin signaling, we see consistency in the role of signal transduction through the cytoskeleton.

More specifically related to the cytoskeleton, we also see an > 16-fold enrichment for actin-filament based transport, an almost 13-fold enrichment for actin filament-based movement, and a > 10-fold enrichment for regulation of actomyosin structure organization. Furthermore, GO term analysis on the 53 genes found mutually in the F1 and F3 shows an > 100-fold enrichment for both skeletal muscle myosin thick filament assembly and skeletal myofibril assembly (2 genes, *unc-54* and *unc-89*) among 53 enrichments. We also see an > 32-fold enrichment for myofibril assembly, an > 23-fold enrichment for actomyosin structure organization, and an > 11-fold enrichment for both actin cytoskeletal organization and actin filament-based processes (Fig. 3B).

This guided our attention towards the five actin orthologues in *C. elegans* (*act-1-5)*. *act-4* routinely came up in the sliding window analysis. It was one of 13 protein-coding genes whose normalized mutant density was over 8 standard deviations from the mean using a 1Kbp sliding window in the F3, alongside *frl-1*, an orthologue of human FMNL1/2/3

(formalin like), which has predicted actin binding activity, and *ttn-1*, an orthologue of human SPEGNB (SPEG neighbor), which exhibits actin filament binding activity and myosin binding activity. It is notable that *frl-1* and *ttn-1* are very large genes, 27,417 bp and 81,722 bp, respectively. Their large size may bias them for selection in the sliding window analysis, as they are in more of the sliding windows across the genome compared to a smaller gene. *act-4*, however, is only 3,864 bp, making it a highly intriguing candidate.

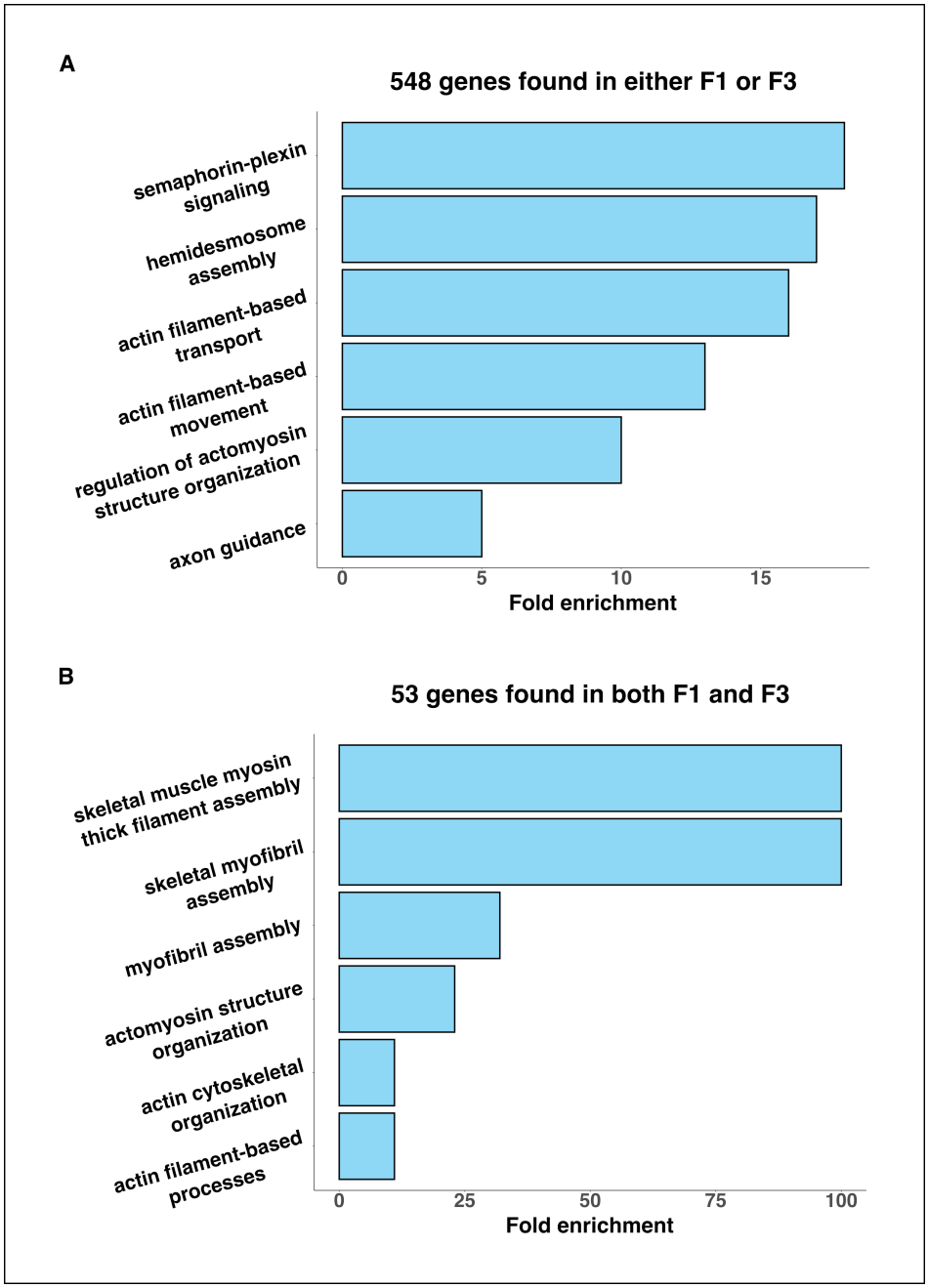

Figure 3 GO enrichment points to actin networks

A) A selection of GO enrichments found in the set of 548 genes found as significant in either the F1 or the F3 using cDNA analysis. B) A selection of GO enrichments found in the set of 53 genes found as significant in both the F1 and the F3 using cDNA analysis.

act-4 and actin networks affect endoderm development

act-4 is an orthologue of human beta-actin and is one of five actin genes in *C. elegans* (MacQueen et al., 2005). Intriguingly, ACT-4 also has predicted interactions with HMP-2 (Zhong & Sternberg, 2006). *act-4 (RNAi)* results in embryonic arrest, and FACS analysis after *act-4 (RNAi)* shows low levels of embryos (1.07 ±0.55%) in the 'excess gut' gate (Fig. 4A & B).

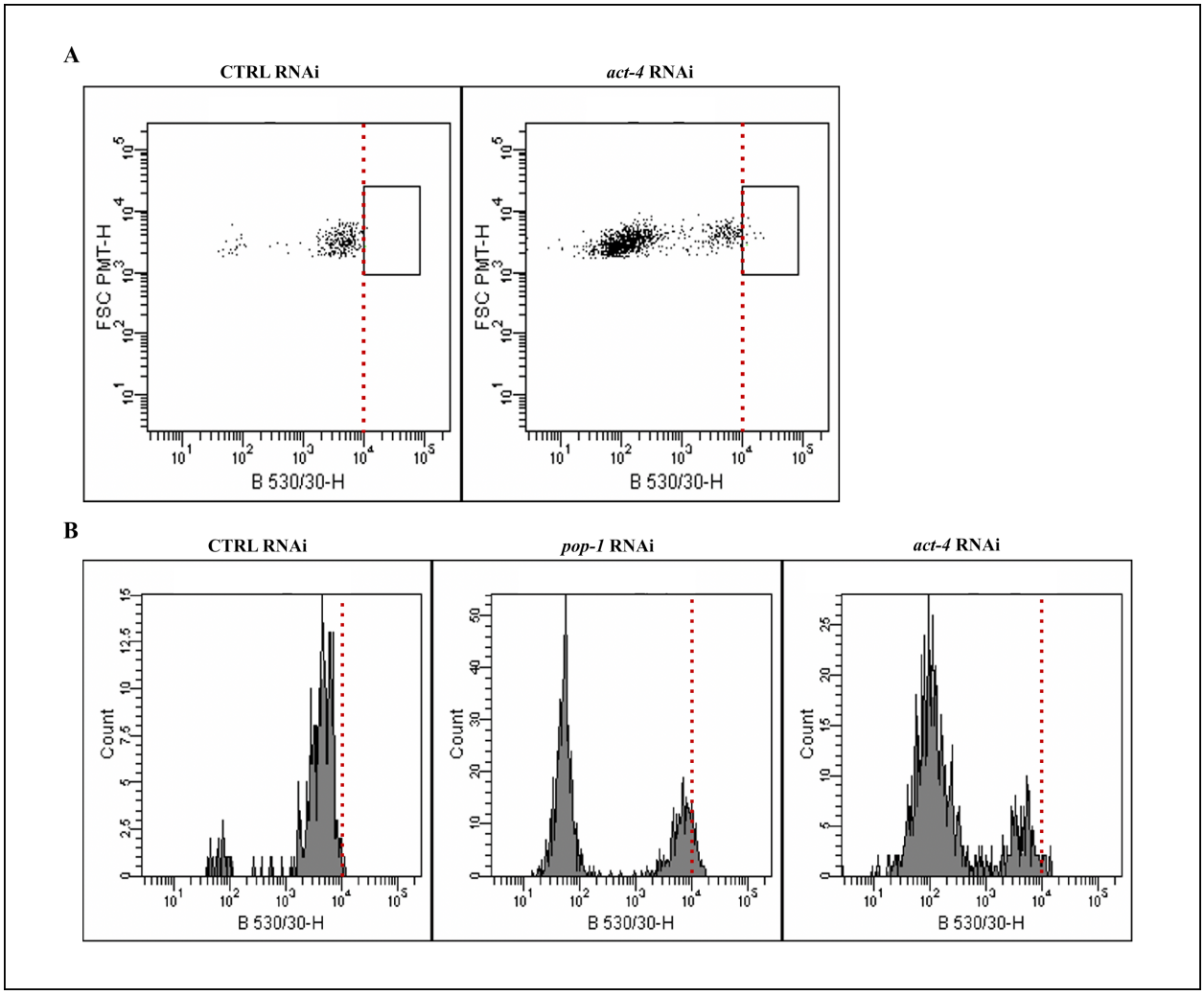

Figure 4 *act-4 (RNAi)* shows embryos with a higher intensity of GFP

A) Embryos from control *(RNAi)* or *act-4 (RNAi)* were analyzed on the FACS machine. X-axis shows 530/570 GFP emission, and Y-axis shows forward scatter FSC. The rectangle shows the excess gut gate, and the red dashed line has been artificially added to mark 10^4 530/570-H GFP emission, which has been the lower limit of GFP emission used for the excess gut gate throughout this study. Any embryos (dots) falling within the gate have been artificially colored black, as the original green was very faint. B) Embryo counts from FACS analysis of control *(RNAi)*, *pop-1 (RNAi)*, and *act-4 (RNAi)*. The red dashed line has been artificially added to mark 10^4 530/570-H GFP emission, which has been the lower limit of GFP emission used for the excess gut gate throughout this study.

Microscopic analysis of *act-4 (RNAi)* embryos shows that a small percentage of embryos (5.01 ± 1.95%, n=766) show more than the 20 expected *elt-2p::GFP* positive cells when scored by hand, up to 25 cells. The distribution of GFP+ cell counts in *act-4* vs control *(RNAi)* embryos was significant (Mann-Whitney-Wilcoxon p-value=1.333e-10, Fig. 5B). It is not known if these extra cells are a result of extra divisions within the E-lineage or inappropriate specification of other cell types. The fact that this method could identify *act-4* even though it has a very mild excess-endoderm phenotype shows its sensitivity and power. Due to the pleiotropic nature of ACT-4, embryos arrested at various stages and took on a variety of morphologies. Some embryos had started elongation before arrest while others remained as a ball of cells. All of the embryos with the extra intestine phenotype have not elongated and show a cluster of GFP+ cells at one pole of the embryo (Fig. 5A), suggesting that dosage and timing of *act-4 (RNAi)* is crucial for this phenotype. Many embryos that appeared similarly staged had variable GFP expression. The majority of embryos had no

GFP expression at all, suggesting that either *elt-2* transcription was never promoted, that it was downregulated early in development, or that ELT-2+ cells died before scoring. We posit that this variability in GFP strengthens the hypothesis that actin is important in controlling specifically endoderm cell number. While the mechanism through which *act-4 (RNAi)* can generate extra intestinal cells in *C. elegans* embryos remains unknown, it may present a new model to study actin's effects on differentiation and gene regulation.

cyk-1 is a dominant actin filament nucleator during cytokinesis, and colocalizes with the actomyosin contractile ring (Chan et al., 2018; Gong et al., 2018). *cyk-1* was one of 388 protein-coding genes found as significant in the F1 alongside *pie-1*, and *cyk-1* has been previously identified in the lab as affecting endoderm cell number (Mengarelli, 2006). *cyk-1 (RNAi)* showed 5.26 ± 1.01% (n=133) of embryos with extra gut cells, ranging from 23 to 34 GFP+ cells (Fig. 5B). Additionally, *tns-1*, a tensin homologue, was found as significant in the F1 with *cyk-1* and *pie-1*. *tns-1* helps mediate actin-filament binding at focal adhesions and has a positive role in motor neuron axon regeneration, and it is also enriched in the intestine (Han et al., 2017; Haynie, 2014; Hisamoto et al., 2019; Le Clainche & Carlier, 2008). Like *cyk-1*, *tns-1* was identified in an earlier screen for genes affecting gut cell number (Mengarelli, 2006). While the mechanism in which actin-effectors *cyk-1* and *tns-1* alter endoderm cell number is unknown, the identification of *act-4* as a modulator of the endoderm GRN may start to bridge these concepts.

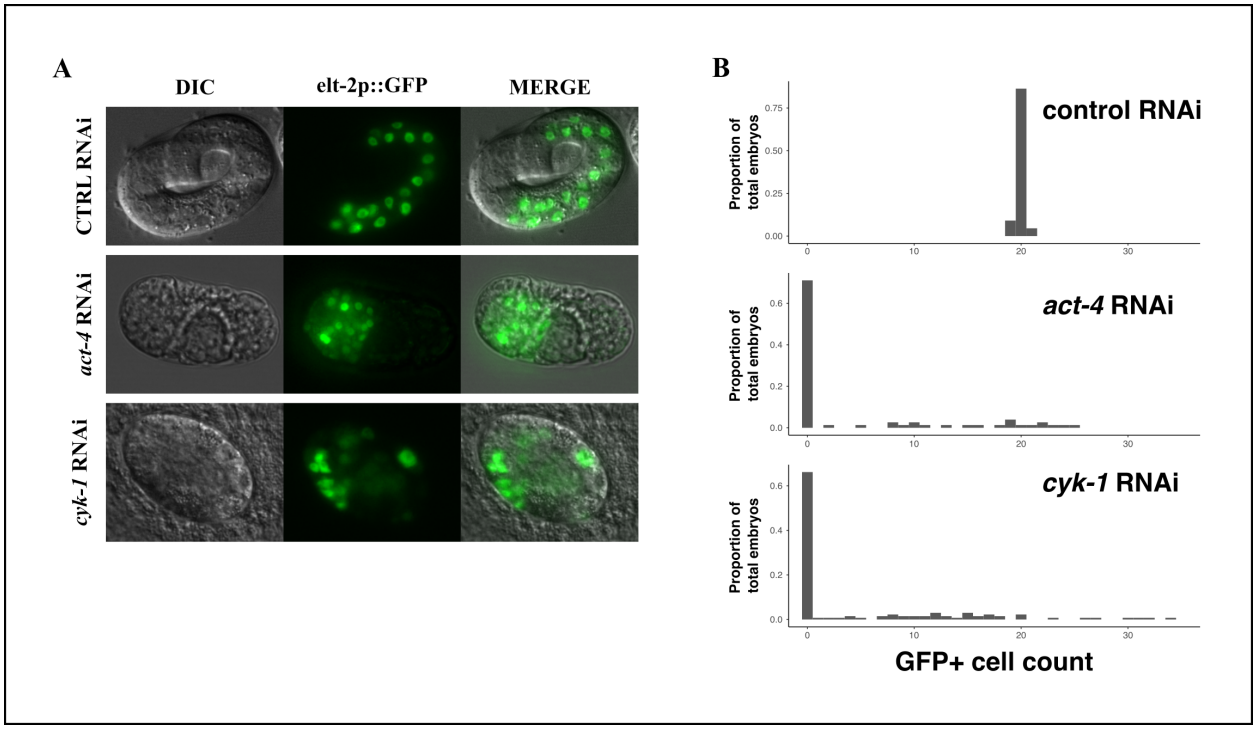

Figure 5 *act-4 (RNAi)* and *cyk-1 (RNAi)* produce embryos with extra gut cells

A) DIC and GFP fluorescent images of an L4440 control *(RNAi)* -treated late-stage embryo showing the expected 20 GFP+ gut cells, an *act-4 (RNAi)* embryo showing excess (24) gut cells, and a *cyk-1 (RNAi)* embryo showing excess (26) gut cells. All images are from a single z-plane except the GFP panel for *act-4 (RNAi)*, which shows a max projection of z-stacks. B) GFP+ cell count of ctrl, *act-4 (RNAi)*, and *cyk-1 (RNAi)* embryos. Mann-Whitney-Wilcoxon tests show that *act-4 (RNAi)* and *cyk-1 (RNAi)* embryo populations are significantly different from the control population (p-values= 1.333e-10 and 4.496e-13, respectively).

Actin is incredibly pleiotropic in *C. elegans* and across eukaryotes (Pollard, 2016; Pollard & Cooper, 2009; Varland et al., 2019). In embryonic development alone, actin plays a role in cytoplasmic streaming, fertilization, cytokinesis dynamics, chromosome segregation, polarity, intracellular transport, and cell morphology, including axon generation

(reviewed in Velarde et al., 2007). Actin has also been implicated in differentiation, and it has been reported that actin inhibition leads to an increase in endodermal differentiation in mouse embryonic stem cells (Boraas et al., 2018; Guo et al., 2014; Jain et al., 2013; Roy et al., 2018). One study suggests that actin-myosin polymerization driven by Rho activity pushes adult stem cells towards specific lineage determinations (McBeath et al., 2004). Another shows that endoderm proliferation persists in embryonic stem cells even if morphological changes induced by actin inhibition are mitigated (Boraas et al., 2018). It has also been shown that increased reprogramming and stemness of different cell cultures showed increased actin tension in generated spherical embryonic bodies (Guo et al., 2014). How cell shape can drive gene expression patterns is under investigation, and it has been reported that actin-myosin contraction can affect histone deacetylase distribution and activity in cultured fibroblasts (Jain et al., 2013).

4.4 Conclusions

This method produced a highly mutagenized sample, and thousands of candidates were pulled from analyses as having high levels of mutagenicity above the background. Many of the candidates were intergenic or, if they were in genes, a knockdown failed to produce arrested embryos, let alone embryos with excess gut. It must be considered that a lot of these regions may be acting in conjunction with each other to produce arrested embryos with extra intestinal cells. In addition to synthetic phenotypes, these mutations might represent either gain of function mutations or be neomorphs that affect protein-protein interactions, which would not be detected by RNAi knockdown experiments. This technique to explore a gene regulatory network may be improved by performing a more conservative

mutagenesis on the samples as to reduce the background mutations. Even with a large number of candidates, we were able to identify positive controls such as *pop-1, pie-1*, and Wnt factors, along with an enrichment for axon guidance genes, which are also known to be involved in endoderm development. We identified a gene, *act-4*, newly discovered to be involved in the endoderm gene regulatory network in *C. elegans.* While we identified *act-4* and *cyk-1* as new modulators of endoderm development, perhaps more exciting was our ability to newly implicate the actin network and actin-based signaling in endoderm cell divisions. The high enrichment for actin-based processes in our samples is accompanied by high enrichment of semaphorin-plexin signaling and hemidesmosome assembly. This new method may be used to investigate the differentiation of other germ layers or any phenotype that can be sorted appropriately, including disease models.

Appendix 1

Experiments concerning *pink-1*

Appendix 1.1

pink-1(w46) CRISPR-generated allele

We used a CRISPR-based system to generate a new allele of *pink-1*. PINK-1 (Pten Induced Kinase) acts as a sensor of mitochondrial membrane potential and assists in initiating mitophagy of defective mitochondria. PINK-1 is constantly cleaved from the membrane of mitochondria with acceptable membrane potential and builds up on the membrane of defective mitochondrial, signaling them for clearance through mitophagy (Fedorowicz et al., 2014; Geisler et al., 2010; Narendra et al., 2008; Sha et al., 2010; Vives-Bauza et al., 2010). We generated a new deletion allele of *pink-1* (Fig. 1) by following the protocol laid out in the supplemental material (File S1) of a 2014 paper from Geraldine Seydoux's lab detailing a CRISPR process for *C. elegans* (Paix et al., 2014). We used two sgRNAs, created by cloning two different oligos (GM30-TGGCCGGAAAACTCGACCGGTTTTAGAGCTAGAAATAGCAAGT and GM31- ACAAAAGAACGCAATAGACGTTTTAGAGCTAGAAATAGCAAGT) into pDD162 constructs using the Q5 Site-Directed Mutagenesis kit from NEB. Following injection of these plasmids, worms were screened for the deletion allele by PCR-amplifying *pink-1* and analyzing the results by gel-electrophoresis. The deletion allele was extracted from the gel and sequenced. In total, we deleted 2013 bp, starting in exon 3 and extending into the 3' UTR.

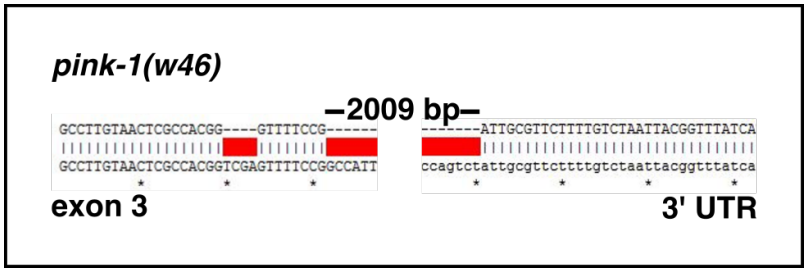

Figure 1 Schematic of *pink-1(w46)*

CRISPR-generated *pink-1(w46)* shows a total deletion of 2013 bp, beginning in exon 3 and ending in the 3' UTR.

Appendix 1.2

PINK-1 in endoderm development

pink-1(w46) is a 2013 bp deletion starting in exon 3 and deleting the rest of the coding sequence (see Appendix 1.1). *pink-1(tm1779)* is a 350 bp deletion that begins upstream of *pink-1*, in the 3' UTR of its neighboring gene, *EEED8.10*, and ends in the beginning of *pink-1*'s second exon. *pink-1(ok3538)* is a roughly 500 bp deletion in exons 6 and 7. It appears that different alleles of *pink-1* have different effects on endoderm development in SKN-1 and MOM-2 knockdowns (Fig. 2). Of the three *pink-1* alleles tested under *skn-1 (RNAi)*, *pink-1(w46)* significantly increases endoderm development (44.61 ± 5.88% of embryos with gut compared with 29.97 ± 6.89% in WT, p-value=0.026), *pink-1(ok3538)* significantly decreases endoderm development (17.62 ± 2.37%, p-value=0.016), and *pink-1(tm1779)* has no effect (33.21 ± 13.08%) (Fig. 2A). *pink-1(tm1779)* does, however, significantly increase endoderm development in *mom-2(or42)* mutant embryos (42.93 ± 3.81% compared with 17.96 ± 6.45% in *mom-2(or42)*, p-value=0.0003) (Fig. 2B). Intriguingly, *pink-1* is just upstream of *brap-2*, and *EEED8.10*, *pink-1*, and *brap-2* are

transcribed together in operon CEOP2176. While none of these genes have been implicated in gut development prior to the present study, our results suggest that this operon may be involved in fine-tuning inputs that control the endoderm GRN.

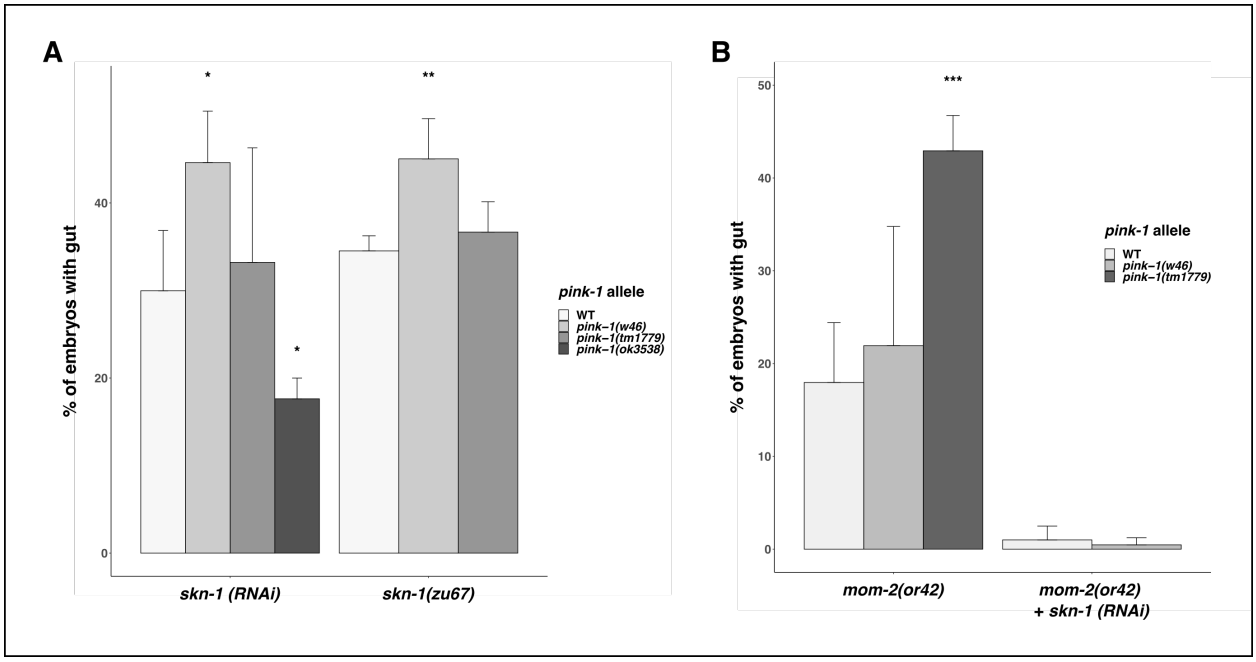

Figure 2 PINK-1 affects endoderm development

A) Bar graph depicting the percent of embryos showing gut granules after *skn-1 (RNAi)* or in *skn-1(zu67)* mutants in different *pink-1* genetic backgrounds. Error bars show standard deviation, and asterisks show the results of a Student's t-test comparing individual *pink-1* alleles with the WT control. B) Bar graph depicting the percent of embryos showing gut granules in *mom-2(or42)* mutants or in *mom-2(or42)* mutants after *skn-1 (RNAi)* in different *pink-1* genetic backgrounds. Error bars show standard deviation, and asterisks show the results of a Student's t-test comparing individual *pink-1* alleles with the WT control.

Appendix 1.2

PINK-1 in cell death

Previous work from the lab identified PINK-1 as an enhancer of cell death (Dey, 2013). In an attempt to replicate and confirm these results, we looked at dopaminergic (DA) neurons in the *C. elegans* pharynx (CEP and ADE neurons, Fig. 3A). Using a *dat-1*-driven GFP marker, which specifically labels the DA neurons, we originally scored for the presence of extra axons extending to the anterior tip of the pharynx (Fig. 3D). In this assay, we see that all three *pink-1* alleles tested show a significant increase in fed, adult worms with an extra GFP-labeled axon when compared with WT worms (WT- 22.86 ± 5.12, *pink-1(tm1779)*- 58.37 ± 9.19, p-value=3.24e-06, *pink-1(ok3538)*- 58.39 ± 4.93, p-value=1.17e-05, *pink-1(w46)*- 92.22 ± 3.14, p-value=7.35e-07). We also found that *pink-1* mutant worms that were either scored as dauers or had experienced the dauer state showed significantly reduced percentages of the extra axon phenotype when compared to their fed counterparts (*pink-1(tm1779)*- 14.81 ± 5.10, p-value=0.0001, *pink-1(ok3538)*- 11.08 ± 3.96, p-value=3.17e-05, *pink-1(w46)*- 53.47 ± 11.37, p-value=0.01). As the extra axon was sometimes very faint, we performed antibody staining against GFP in *pink-1(tm1779)*; *pdat-1*::GFP worms to provide better images of the extra axon running through the middle of the pharynx (Fig. 3C).

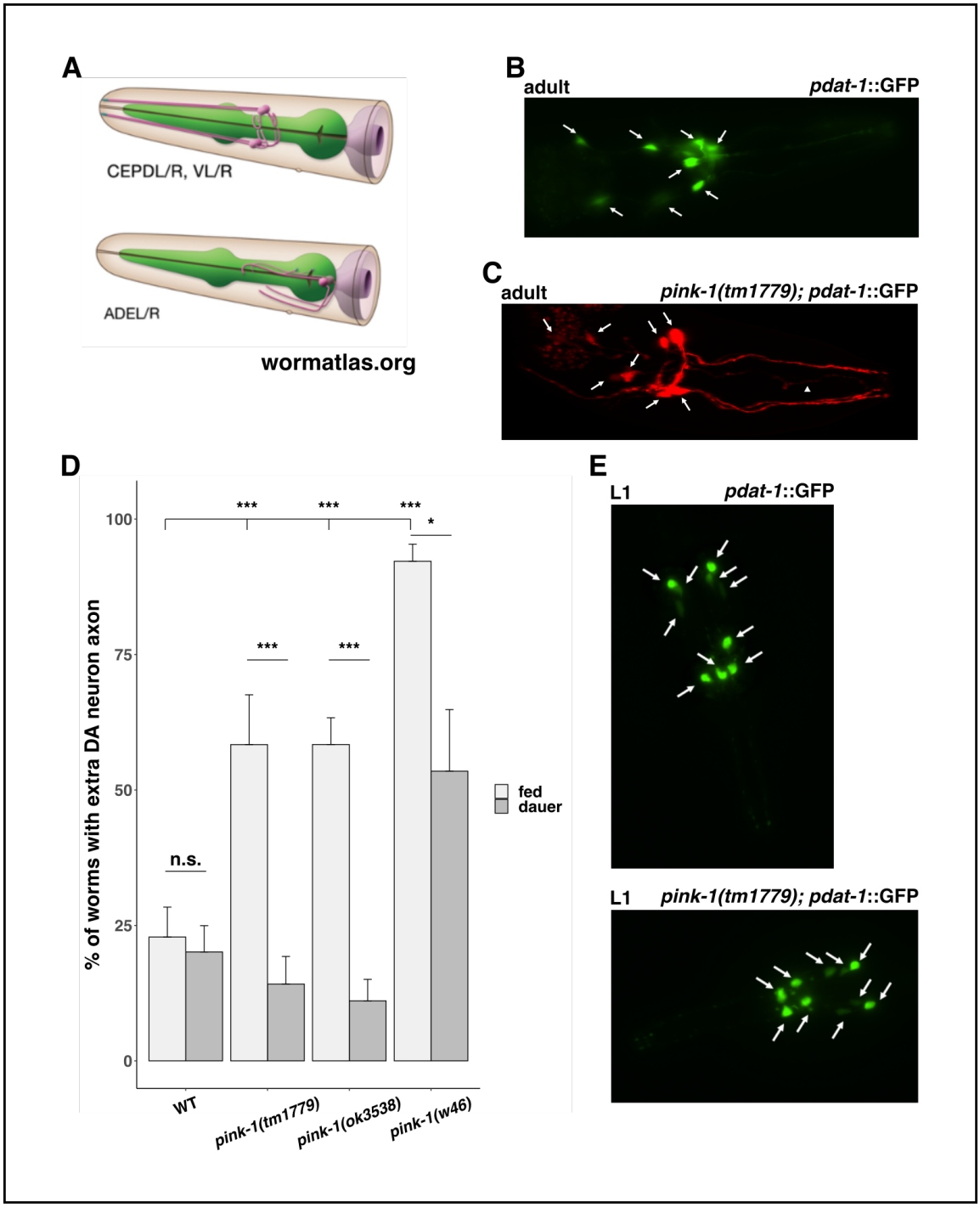

Figure 3 *dat-1* driven GFP shows extra axons in *pink-1* mutants

A) Graphic from wormatlas.org depicting the location of DA neurons in the pharynx. Four CEP cells sit between the two pharynx bulbs and project axons to the anterior end of the nose. Two ADE cells rest at the posterior end of the terminal bulb. B) GFP+ cell bodies in an

adult pharynx are shown with white arrows in *pdat-1*::GFP worm in an otherwise WT background. Eight cell bodies are seen. C) Antibody staining against GFP is shown in a *pink-1(tm1779)*; *pdat-1*::GFP worm. Eight cell bodies are shown with white arrows, and an extra axon extending to the tip of the nose is shown with a white arrowhead. D) The presence of extra axons (more than 4, typically 5) was scored in different *pink-1* genetic backgrounds. Trials were also differentiated based on whether the worms had been fed continuously or whether they had experienced dauer (either scored as dauers or post-dauer). Student's t-tests were performed to test for significance between the fed and dauer data for each strain as well as between the fed WT control allele and individual fed *pink-1* mutant alleles. E) Representative images showing ten GFP+ cell bodies in both a *pdat-1*::GFP L1 larvae and a *pink-1(tm1779)*; *pdat-1*::GFP L1 larvae.

As examination of these strains continued, it became clear that the *pdat-1*::GFP marker was labeling extra cells (eight instead of the expected six) in the WT and mutant adults (Fig. 3B). Looking at synchronized L1 larvae, we saw that worms with this marker hatched with a total of ten GFP+ labeled cells (Fig. 3E). To address this issue, we used a second neuronal marker, *cat-2*::GFP, that should also be specific to the DA neurons. Now scoring for cell bodies instead of extra axons in the pharynx, we saw a significantly increased incidence of an extra cell body in *pink-1(tm1779)*; *cat-2*::GFP L1 larvae when compared with WT (WT- 6.42 ± 3.05%, *pink-1(tm1779)*- 25.21 ± 5.87%, p-value=0.0001) (Fig. 4A+B). We also saw a slight, but significant, decrease in the number of cell corpses seen in the pharynx of *pink-1(tm1779)* L1 larvae (WT- 12.20 ± 5.61, *pink-1(tm1779)*- 8.05 ± 5.85, p-value=0.01) (Fig. 4D). Surprisingly, when *pink-1* mutant strains were scored again a few years later for

cat-2::GFP+ cell bodies, we no longer saw a significant difference in phenotype (Fig. 4C). Before scoring again in 2019, these strains were thawed from our frozen collection, and they were scored ~2 generations after the thaw. There may be some lasting epigenetic effects from the freezing that reduce the prevalence of these extra cell bodies, and future studies should follow these lines for multiple generations after thawing to see if the phenotype manifests.

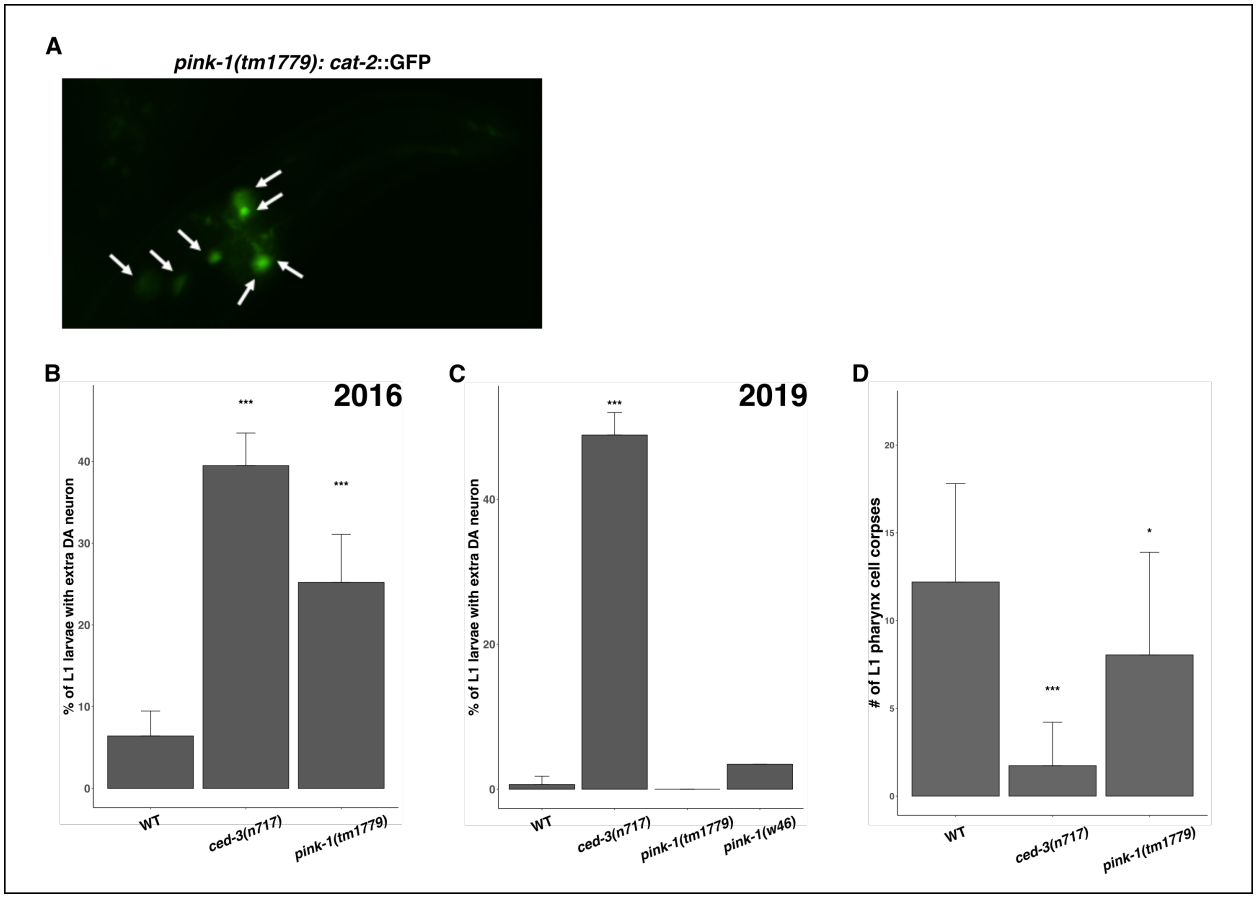

Figure 4 *cat-2*::GFP allows for better comparison of DA cell body number in the pharynx

A) A representative image of a *pink-1(tm1779)*; *cat-2*::GFP L1 larvae pharynx showing the expected six DA neurons (white arrows). B) Bar graph of data collected in 2016 showing the percent of L1 larvae with extra GFP+ cell bodies in the pharynx. All strains used expressed

cat-2::GFP. Error bars show standard deviation, and asterisks show significance determined by using a Student's t-test between mutant strains and the WT control. C) Same as B, except data was collected in 2019, using strains ~2 generations out after thaw from a frozen population. D) Bar graph showing the number of cell corpses seen in L1 pharynxes. All strains were in a *ced-1(e1735)* background to increase visualization of corpses. Error bars show standard deviation, and asterisks show significance determined by using a Student's t-test between mutant strains and the WT control.

Appendix 2

MS-specific markers after *skn-1 (RNAi)*

We discovered that MS-specific markers turn on in late-stage, arrested embryos after *skn-1 (RNAi)* (Fig. 1). This was surprising because the MS transcriptional cascade, similarly to the E cascade, begins with SKN-1 (Broitman-Maduro, 2006; Maduro et al., 2007). The MS-marker *tbx-35*::GFP is typically expressed in the early embryo, and we do not see early embryonic expression after *skn-1 (RNAi)* (see Chapter 3). Looking at late-stage arrested embryos, we see 83.40 ± 6.83% of embryos expressing *tbx-35*::GFP (strain- MS562) and 53.44 ± 12.99% of embryos expressing *ceh-22*::GFP (strain- JR1575).

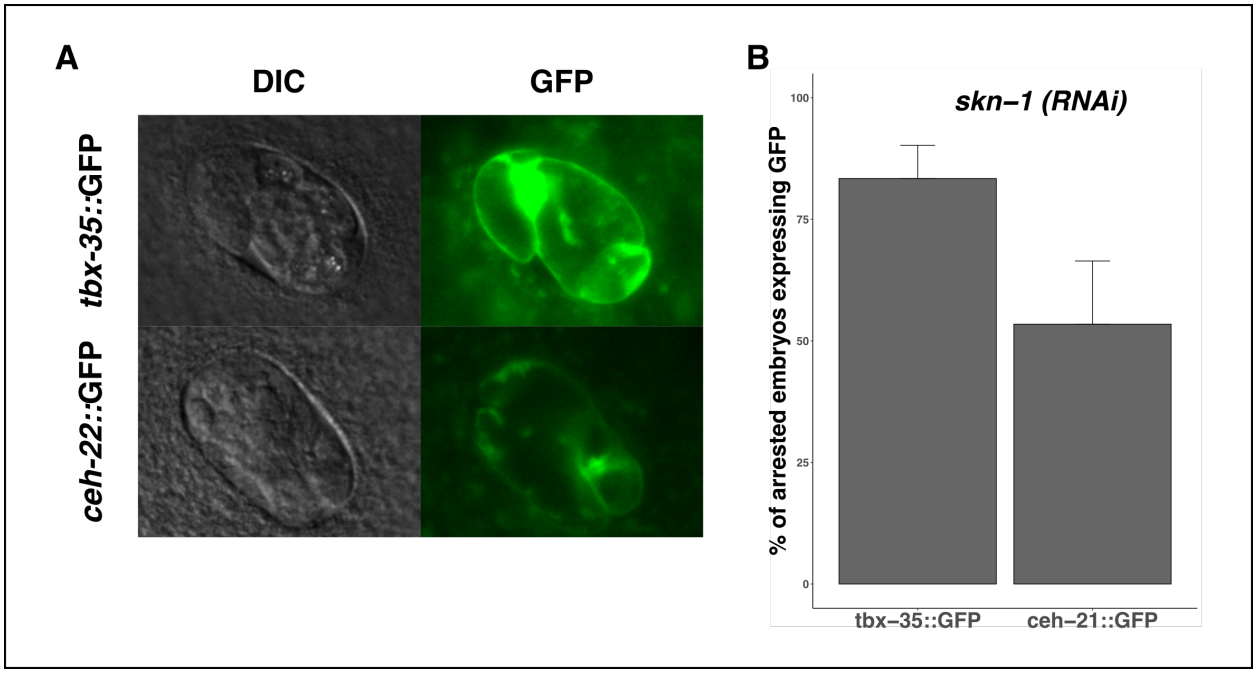

Figure 1 MS markers are expressed in *skn-1 (RNAi) arrested embryos*

A) DIC and GFP images of arrested embryos after *skn-1 (RNAi)* expressing either *tbx-35*::GFP or *ceh-22*::GFP in an otherwise WT genetic background. B) Bar graph showing the percent of arrested embryos after *skn-1 (RNAi)* expressing either *tbx-35*::GFP or *ceh-22*::GFP MS-specific markers. Error bars show standard deviation.

Appendix 3

ced-3 in transorganogenesis

Transorganogensis has been described in *C. elegans* whereby the pharynx and somatic gonad begin to express gut-specific markers and take on a gut-like morphology after heat shock-induced over expression of the GATA factor, ELT-7 (Riddle et al., 2013, 2016). When L1 larvae with a *hs::elt-7* construct are exposed to heat shock conditions (30 minutes at 33°C), the majority of worms arrest at the L1 stage (74.94 ± 17.33% of worms) (Fig. 1). CED-3 is a caspase most notably involved in initiating apoptosis in cells programmed to die

(Ellis & Horvitz, 1986; Sulston & Horvitz, 1977). Here we report that *ced-3(n717)* mutants partially rescue the L1 arrest phenotype after heat shock-induced expression of ELT-7 (33.69 ± 17.14%, p-value=0.04) (Fig. 1). This may indicate that CED-3 plays a role in transorganogenesis, and the whole process is partially blocked in the *ced-3(n717)* strain. Equally plausible is that cell death is necessary for the arrest phenotype.

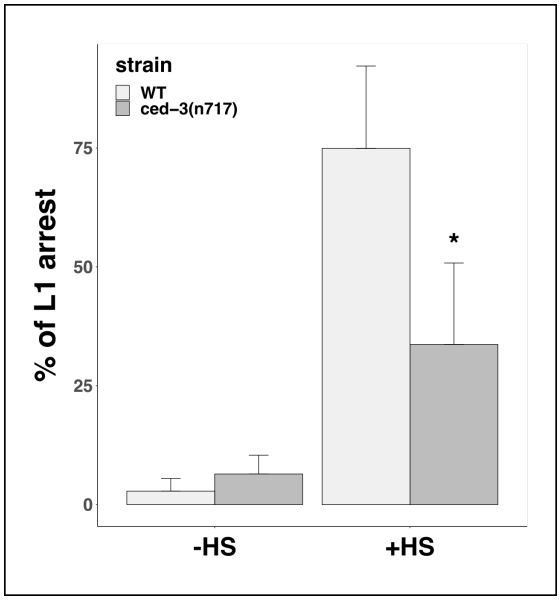

Figure 1 *ced-3(n717)* rescues larval arrest after heat shock

Bar graph showing the percent of worms that arrested as L1 larvae in either WT or *ced-3(n717)* strains expressing the *hs::elt-7* construct wIs125 [*hs::elt-7 + pRF4 (rol-6(su1004))*]. Larval arrest with and without a 30 minute heat shock at 33ºC is shown. Error bars show standard deviation, and the asterisk shows significance determined by using a Student's t-test between *ced-3(n717)* and the WT control.

References

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544. https://doi.org/10.1093/nar/gky379

Alto, L. T., & Terman, J. R. (2017). Semaphorins and their Signaling Mechanisms. *Methods in Molecular Biology (Clifton, N.J.)*, *1493*, 1–25. https://doi.org/10.1007/978-1-4939-6448-2_1

Altun, Z. F., & Hall, D. H. (2006). WormAtas Hermaphrodite Handbook—Introduction. *WormAtlas*. https://doi.org/10.3908/wormatlas.1.1

An, J. H., & Blackwell, T. K. (2003). SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. *Genes & Development*, *17*(15), 1882–1893. https://doi.org/10.1101/gad.1107803

Andersen, E. C., Gerke, J. P., Shapiro, J. A., Crissman, J. R., Ghosh, R., Bloom, J. S., Félix, M.-A., & Kruglyak, L. (2012). Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. *Nature Genetics*, *44*(3), 285–290. https://doi.org/10.1038/ng.1050

Asada, M., Ohmi, K., Delia, D., Enosawa, S., Suzuki, S., Yuo, A., Suzuki, H., & Mizutani, S. (2004). Brap2 Functions as a Cytoplasmic Retention Protein for p21 during Monocyte Differentiation. *Molecular and Cellular Biology*, *24*(18), 8236–8243. https://doi.org/10.1128/MCB.24.18.8236-8243.2004

Asan, A., Raiders, S. A., & Priess, J. R. (2016). Morphogenesis of the C. elegans Intestine Involves Axon Guidance Genes. *PLOS Genetics*, *12*(4), e1005950. https://doi.org/10.1371/journal.pgen.1005950

Avery, C. L., He, Q., North, K. E., Ambite, J. L., Boerwinkle, E., Fornage, M., Hindorff, L. A., Kooperberg, C., Meigs, J. B., Pankow, J. S., Pendergrass, S. A., Psaty, B. M., Ritchie, M. D., Rotter, J. I., Taylor, K. D., Wilkens, L. R., Heiss, G., & Lin, D. Y. (2011). A Phenomics-Based Strategy Identifies Loci on APOC1, BRAP, and PLCG1 Associated with Metabolic Syndrome Phenotype Domains. *PLoS Genetics*, *7*(10). https://doi.org/10.1371/journal.pgen.1002322

Barker, N., van Es, J. H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P. J., & Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, *449*(7165), 1003–1007. https://doi.org/10.1038/nature06196

Barnes, T. M., Kohara, Y., Coulson, A., & Hekimi, S. (1995). Meiotic recombination, noncoding DNA and genomic organization in Caenorhabditis elegans. *Genetics*, *141*(1), 159–179.

Barriere, A. (2005). Natural variation and population genetics of Caenorhabditis elegans. *WormBook*. https://doi.org/10.1895/wormbook.1.43.1

Barrière, A., & Félix, M.-A. (2005). High Local Genetic Diversity and Low Outcrossing Rate in Caenorhabditis elegans Natural Populations. *Current Biology*, *15*(13), 1176–1184. https://doi.org/10.1016/j.cub.2005.06.022

Bass, J. I. F., Tamburino, A. M., Mori, A., Beittel, N., Weirauch, M. T., Reece-Hoyes, J. S., & Walhout, A. J. M. (2014). Transcription factor binding to Caenorhabditis elegans first introns reveals lack of redundancy with gene promoters. *Nucleic Acids Research*, *42*(1), 153. https://doi.org/10.1093/nar/gkt858

Bei, Y., Hogan, J., Berkowitz, L. A., Soto, M., Rocheleau, C. E., Pang, K. M., Collins, J., & Mello, C. C. (2002). SRC-1 and Wnt Signaling Act Together to Specify Endoderm and to Control Cleavage Orientation in Early C. elegans Embryos. *Developmental Cell*, *3*(1), 113–125. https://doi.org/10.1016/S1534-5807(02)00185-5

Berkowitz, L. A., & Strome, S. (2000). MES-1, a protein required for unequal divisions of the germline in early C. elegans embryos, resembles receptor tyrosine kinases and is localized to the boundary between the germline and gut cells. *Development (Cambridge, England)*, *127*(20), 4419–4431.

Bienz, M., & Clevers, H. (2000). Linking Colorectal Cancer to Wnt Signaling. *Cell*, *103*(2), 311–320. https://doi.org/10.1016/S0092-8674(00)00122-7

Blackwell, T. K., Bowerman, B., Priess, J. R., & Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. *Science (New York, N.Y.)*, *266*(5185), 621–628. https://doi.org/10.1126/science.7939715

Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., & Nekrutenko, A. (2010). Manipulation of FASTQ data with Galaxy. *Bioinformatics*, *26*(14), 1783–1785. https://doi.org/10.1093/bioinformatics/btq281

Blankenberg, D., Von Kuster, G., Bouvier, E., Baker, D., Afgan, E., Stoler, N., Taylor, J., Nekrutenko, A., & Galaxy Team. (2014). Dissemination of scientific software with Galaxy ToolShed. *Genome Biology*, *15*(2), 403. https://doi.org/10.1186/gb4161

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boraas, L. C., Pineda, E. T., & Ahsan, T. (2018). Actin and myosin II modulate differentiation of pluripotent stem cells. *PloS One*, *13*(4), e0195588. https://doi.org/10.1371/journal.pone.0195588

Bowerman, B., Draper, B. W., Mello, C. C., & Priess, J. R. (1993). The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos. *Cell*, *74*(3), 443–452. https://doi.org/10.1016/0092-8674(93)80046-h

Bowerman, B., Eaton, B. A., & Priess, J. R. (1992). Skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. *Cell*, *68*(6), 1061–1075. https://doi.org/10.1016/0092-8674(92)90078-Q

Bowerman, Bruce, Draper, B. W., Mello, C. C., & Priess, J. R. (1993). The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos. *Cell*, *74*(3), 443–452. https://doi.org/10.1016/0092-8674(93)80046-H

Brenner, S. (1974). The Genetics of CAENORHABDITIS ELEGANS. *Genetics*, *77*(1), 71–94.

Broitman-Maduro, G. (2006). Specification of the C. elegans MS blastomere by the T-box factor TBX-35. *Development*, *133*(16), 3097–3106. https://doi.org/10.1242/dev.02475

Broitman-Maduro, Gina, Maduro, M. F., & Rothman, J. H. (2005). The Noncanonical Binding Site of the MED-1 GATA Factor Defines Differentially Regulated Target

Genes in the C. elegans Mesendoderm. *Developmental Cell*, *8*(3), 427–433. https://doi.org/10.1016/j.devcel.2005.01.014

Burdick, J., Walton, T., Preston, E., Zacharias, A., Raj, A., & Murray, J. I. (2016). Overlapping cell population expression profiling and regulatory inference in C. elegans. *BMC Genomics*, *17*(1), 159. https://doi.org/10.1186/s12864-016-2482-z

Burga, A., Ben-David, E., Lemus Vergara, T., Boocock, J., & Kruglyak, L. (2019). Fast genetic mapping of complex traits in C. elegans using millions of individuals in bulk. *Nature Communications*, *10*(1), 2680. https://doi.org/10.1038/s41467-019-10636-9

Burgess, M., & Puhalla, S. (2014). BRCA 1/2-Mutation Related and Sporadic Breast and Ovarian Cancers: More Alike than Different. *Frontiers in Oncology*, *4*. https://doi.org/10.3389/fonc.2014.00019

Chan, F.-Y., Silva, A. M., Saramago, J., Pereira-Sousa, J., Brighton, H. E., Pereira, M., Oegema, K., Gassmann, R., & Carvalho, A. X. (2018). The ARP2/3 complex prevents excessive formin activity during cytokinesis. *Molecular Biology of the Cell*, *30*(1), 96–107. https://doi.org/10.1091/mbc.E18-07-0471

Chen, C., Lewis, R. E., & White, M. A. (2008). IMP Modulates KSR1-dependent Multivalent Complex Formation to Specify ERK1/2 Pathway Activation and Response Thresholds. *The Journal of Biological Chemistry*, *283*(19), 12789–12796. https://doi.org/10.1074/jbc.M709305200

Chida-Nagai, A., Shintani, M., Sato, H., Nakayama, T., Nii, M., Akagawa, H., Furukawa, T., Rana, A., Furutani, Y., Inai, K., Nonoyama, S., & Nakanishi, T. (2019). Role of BRCA1-associated protein (BRAP) variant in childhood pulmonary arterial hypertension. *PLoS ONE*, *14*(1). https://doi.org/10.1371/journal.pone.0211450

Chisholm, A. D. (2008). Semaphorin signaling in morphogenesis: Found in translation. *Genes & Development*, *22*(8), 955–959. https://doi.org/10.1101/gad.1669308

Clokey, G. V., & Jacobson, L. A. (1986). The autofluorescent "lipofuscin granules" in the intestinal cells of Caenorhabditis elegans are secondary lysosomes. *Mechanisms of Ageing and Development*, *35*(1), 79–94. https://doi.org/10.1016/0047-6374(86)90068-0

Cook, D. E., Zdraljevic, S., Roberts, J. P., & Andersen, E. C. (2017). CeNDR, the Caenorhabditis elegans natural diversity resource. *Nucleic Acids Research*, *45*(D1), D650–D657. https://doi.org/10.1093/nar/gkw893

Cook, D. E., Zdraljevic, S., Tanny, R. E., Seo, B., Riccardi, D. D., Noble, L. M., Rockman, M. V., Alkema, M. J., Braendle, C., Kammenga, J. E., Wang, J., Kruglyak, L., Félix, M.-A., Lee, J., & Andersen, E. C. (2016). The Genetic Basis of Natural Variation in Caenorhabditis elegans Telomere Length. *Genetics*, *204*(1), 371–383. https://doi.org/10.1534/genetics.116.191148

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. *Journal of Cell Biology*, *141*(1), 297–308. https://doi.org/10.1083/jcb.141.1.297

D'Amora, D. R., Hu, Q., Pizzardi, M., & Kubiseski, T. J. (2018). BRAP-2 promotes DNA damage induced germline apoptosis in C. elegans through the regulation of SKN-1 and AKT-1. *Cell Death & Differentiation*, 1. https://doi.org/10.1038/s41418-017-0038-7

Davies, R. G., Wagstaff, K. M., McLaughlin, E. A., Loveland, K. L., & Jans, D. A. (2013). The BRCA1-binding protein BRAP2 can act as a cytoplasmic retention factor for nuclear and nuclear envelope-localizing testicular proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1833*(12), 3436–3444. https://doi.org/10.1016/j.bbamcr.2013.05.015

Denver, D. R., Morris, K., & Thomas, W. K. (2003). Phylogenetics in Caenorhabditis elegans: An analysis of divergence and outcrossing. *Molecular Biology and Evolution*, *20*(3), 393–400. https://doi.org/10.1093/molbev/msg044

Dey, J. (2013). *Programmed Cell Death and Cellular Stress: Insights from Caenorhabditis elegans pink-1 and icd-1*. University of California, Santa Barbara.

Dickins, B., Rebolledo-Jaramillo, B., Su, M. S.-W., Paul, I. M., Blankenberg, D., Stoler, N., Makova, K. D., & Nekrutenko, A. (2014). Controlling for contamination in re-sequencing studies with a reproducible web-based phylogenetic approach. *BioTechniques*, *56*(3), 134–141. https://doi.org/10.2144/000114146

Doumpas, N., Lampart, F., Robinson, M. D., Lentini, A., Nestor, C. E., Cantù, C., & Basler, K. (2019). TCF/LEF dependent and independent transcriptional regulation of Wnt/β-catenin target genes. *The EMBO Journal*, *38*(2), e98873. https://doi.org/10.15252/embj.201798873

Dresen, A., Finkbeiner, S., Dottermusch, M., Beume, J.-S., Li, Y., Walz, G., & Neumann-Haefelin, E. (2015). Caenorhabditis elegans OSM-11 signaling regulates SKN-1/Nrf during embryonic development and adult longevity and stress response. *Developmental Biology*, *400*(1), 118–131. https://doi.org/10.1016/j.ydbio.2015.01.021

Ehrenreich, I. M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J. A., Gresham, D., Caudy, A. A., & Kruglyak, L. (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature*, *464*(7291), 1039–1042. https://doi.org/10.1038/nature08923

Ellis, H. M., & Horvitz, H. R. (1986). Genetic control of programmed cell death in the nematode C. elegans. *Cell*, *44*(6), 817–829. https://doi.org/10.1016/0092-8674(86)90004-8

Elvin, M., Snoek, L. B., Frejno, M., Klemstein, U., Kammenga, J. E., & Poulin, G. B. (2011). A fitness assay for comparing RNAi effects across multiple C. elegans genotypes. *BMC Genomics*, *12*, 510. https://doi.org/10.1186/1471-2164-12-510

Evans, K. S., Zhao, Y., Brady, S. C., Long, L., McGrath, P. T., & Andersen, E. C. (2017). Correlations of Genotype with Climate Parameters Suggest Caenorhabditis elegans Niche Adaptations. *G3: Genes, Genomes, Genetics*, *7*(1), 289–298. https://doi.org/10.1534/g3.116.035162

Ewe, C. K., Torres Cleuren, Y. N., Alok, G., & Rothman, J. H. (2019). ICD-1/BTF3 antagonizes SKN-1-mediated endoderm specification in Caenorhabditis elegans. *MicroPublication Biology*, *2019*(10). https://www.micropublication.org/journals/biology/micropub-biology-000167/

Ewe, C. K., Torres Cleuren, Y. N., & Rothman, J. H. (2020). Evolution and Developmental System Drift in the Endoderm Gene Regulatory Network of Caenorhabditis and Other Nematodes. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00170

Fatima, S., Wagstaff, K. M., Loveland, K. L., & Jans, D. A. (2015). Interactome of the negative regulator of nuclear import BRCA1-binding protein 2. *Scientific Reports*, *5*. https://doi.org/10.1038/srep09459

Fedorowicz, M. A., de Vries-Schneider, R. L. A., Rüb, C., Becker, D., Huang, Y., Zhou, C., Alessi Wolken, D. M., Voos, W., Liu, Y., & Przedborski, S. (2014). Cytosolic cleaved PINK1 represses Parkin translocation to mitochondria and mitophagy. *EMBO Reports*, *15*(1), 86–93. https://doi.org/10.1002/embr.201337294

Fevr, T., Robine, S., Louvard, D., & Huelsken, J. (2007). Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. *Molecular and Cellular Biology*, *27*(21), 7551–7559. https://doi.org/10.1128/MCB.01034-07

Fodde, R., Smits, R., & Clevers, H. (2001). APC, signal transduction and genetic instability in colorectal cancer. *Nature Reviews. Cancer*, *1*(1), 55–67. https://doi.org/10.1038/35094067

Francis, R., & Waterston, R. H. (1991). Muscle cell attachment in Caenorhabditis elegans. *Journal of Cell Biology*, *114*(3), 465–479. https://doi.org/10.1083/jcb.114.3.465

Fukushige, T., Hawkins, M. G., & McGhee, J. D. (1998). The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. *Developmental Biology*, *198*(2), 286–302.

Fukushige, Tetsunari, Hendzel, M. J., Bazett-Jones, D. P., & McGhee, J. D. (1999). Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. *Proceedings of the National Academy of Sciences*, *96*(21), 11883–11888. https://doi.org/10.1073/pnas.96.21.11883

Fukushige, Tetsunari, & Krause, M. (2005). The myogenic potency of HLH-1 reveals wide-spread developmental plasticity in early C. elegans embryos. *Development*, *132*(8), 1795–1805. https://doi.org/10.1242/dev.01774

Fulcher, A. J., Roth, D. M., Fatima, S., Alvisi, G., & Jans, D. A. (2009). The BRCA-1 binding protein BRAP2 is a novel, negative regulator of nuclear import of viral proteins, dependent on phosphorylation flanking the nuclear localization signal. *The FASEB Journal*, *24*(5), 1454–1466. https://doi.org/10.1096/fj.09-136564

Geisler, S., Holmström, K. M., Treis, A., Skujat, D., Weber, S. S., Fiesel, F. C., Kahle, P. J., & Springer, W. (2010). The PINK1/Parkin-mediated mitophagy is compromised by PD-associated mutations. *Autophagy*, *6*(7), 871–878. https://doi.org/10.4161/auto.6.7.13286

Gene expression visualization. (n.d.). Retrieved November 16, 2020, from http://tintori.bio.unc.edu/

Giles, R. H., van Es, J. H., & Clevers, H. (2003). Caught up in a Wnt storm: Wnt signaling in cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1653*(1), 1–24. https://doi.org/10.1016/S0304-419X(03)00005-2

Goldstein, B. (1992). Induction of gut in Caenorhabditis elegans embryos. *Nature*, *357*(6375), 255–257. https://doi.org/10.1038/357255a0

Goldstein, B. (1993). Establishment of gut fate in the E lineage of C. elegans: The roles of lineage-dependent mechanisms and cell interactions. *Development*, *118*(4), 1267–1277.

Goldstein, B. (1995). An analysis of the response to gut induction in the C. elegans embryo. *Development*, *121*(4), 1227–1236.

Gong, T., Yan, Y., Zhang, J., Liu, S., Liu, H., Gao, J., Zhou, X., Chen, J., & Shi, A. (2018). PTRN-1/CAMSAP promotes CYK-1/formin-dependent actin polymerization during endocytic recycling. *The EMBO Journal*, *37*(9), e98556. https://doi.org/10.15252/embj.201798556

Goshima, Y., Ito, T., Sasaki, Y., & Nakamura, F. (2002). Semaphorins as signals for cell repulsion and invasion. *The Journal of Clinical Investigation*, *109*(8), 993–998. https://doi.org/10.1172/JCI15467

Greene, J. S., Brown, M., Dobosiewicz, M., Ishida, I. G., Macosko, E. Z., Zhang, X., Butcher, R. A., Cline, D. J., McGrath, P. T., & Bargmann, C. I. (2016). Balancing selection shapes density-dependent foraging behaviour. *Nature*, *539*(7628), 254–258. https://doi.org/10.1038/nature19848

Guo, J., Wang, Y., Sachs, F., & Meng, F. (2014). Actin stress in cell reprogramming. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(49), E5252-5261. https://doi.org/10.1073/pnas.1411683111

Haber, M., Schüngel, M., Putz, A., Müller, S., Hasert, B., & Schulenburg, H. (2005). Evolutionary History of Caenorhabditis elegans Inferred from Microsatellites: Evidence for Spatial and Temporal Genetic Differentiation and the Occurrence of Outbreeding. *Molecular Biology and Evolution*, *22*(1), 160–173. https://doi.org/10.1093/molbev/msh264

Hall, J. M., Friedman, L., Guenther, C., Lee, M. K., Weber, J. L., Black, D. M., & King, M. C. (1992). Closing in on a breast cancer gene on chromosome 17q. *American Journal of Human Genetics*, *50*(6), 1235–1242.

Han, S., Schroeder, E. A., Silva-García, C. G., Hebestreit, K., Mair, W. B., & Brunet, A. (2017). Mono-unsaturated fatty acids link H3K4me3 modifiers to C. elegans lifespan. *Nature*, *544*(7649), 185–190. https://doi.org/10.1038/nature21686

Haynie, D. T. (2014). Molecular physiology of the tensin brotherhood of integrin adaptor proteins. *Proteins: Structure, Function, and Bioinformatics*, *82*(7), 1113–1127. https://doi.org/10.1002/prot.24560

Helpman, L., Zidan, O., Friedman, E., Kalfon, S., Perri, T., Ben-Baruch, G., & Korach, J. (2017). Young Israeli women with epithelial ovarian cancer: Prevalence of BRCA mutations and clinical correlates. *Journal of Gynecologic Oncology*, *28*(5), e61. https://doi.org/10.3802/jgo.2017.28.e61

Hinck, L. (2004). The Versatile Roles of "Axon Guidance" Cues in Tissue Morphogenesis. *Developmental Cell*, *7*(6), 783–793. https://doi.org/10.1016/j.devcel.2004.11.002

Hinohara, K., Ohtani, H., Nakajima, T., Sasaoka, T., Sawabe, M., Lee, B.-S., Ban, J., Park, J.-E., Izumi, T., & Kimura, A. (2009). Validation of eight genetic risk factors in East Asian populations replicated the association of BRAP with coronary artery disease. *Journal of Human Genetics*, *54*(11), 642–646. https://doi.org/10.1038/jhg.2009.87

Hisamoto, N., Shimizu, T., Asai, K., Sakai, Y., Pastuhov, S. I., Hanafusa, H., & Matsumoto, K. (2019). C. elegans Tensin Promotes Axon Regeneration by Linking the Met-like SVH-2 and Integrin Signaling Pathways. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(29), 5662–5672. https://doi.org/10.1523/JNEUROSCI.2059-18.2019

Hodgkin, J., & Doniach, T. (1997). Natural Variation and Copulatory Plug Formation in Caenorhabditis Elegans. *Genetics*, *146*(1), 149–164.

Hu, Q., D'Amora, D. R., MacNeil, L. T., Walhout, A. J. M., & Kubiseski, T. J. (2017). The Oxidative Stress Response in Caenorhabditis elegans Requires the GATA Transcription Factor ELT-3 and SKN-1/Nrf2. *Genetics*, *206*(4), 1909–1922. https://doi.org/10.1534/genetics.116.198788

Hu, Q., D'Amora, D. R., MacNeil, L. T., Walhout, A. J. M., & Kubiseski, T. J. (2018). The Caenorhabditis elegans Oxidative Stress Response Requires the NHR-49 Transcription Factor. *G3: Genes|Genomes|Genetics*, *8*(12), 3857–3863. https://doi.org/10.1534/g3.118.200727

Hu, S., & Zhu, L. (2018). Semaphorins and Their Receptors: From Axonal Guidance to Atherosclerosis. *Frontiers in Physiology*, *9*. https://doi.org/10.3389/fphys.2018.01236

Huang, S., Shetty, P., Robertson, S. M., & Lin, R. (2007). Binary cell fate specification during C. elegans embryogenesis driven by reiterated reciprocal asymmetry of TCF POP-1 and its coactivator beta-catenin SYS-1. *Development (Cambridge, England)*, *134*(14), 2685–2695. https://doi.org/10.1242/dev.008268

Hunter, C. P., & Kenyon, C. (1996). Spatial and Temporal Controls Target pal-1 Blastomere-Specification Activity to a Single Blastomere Lineage in C. elegans Embryos. *Cell*, *87*(2), 217–226. https://doi.org/10.1016/S0092-8674(00)81340-9

Imaizumi, T., Ando, M., Nakatochi, M., Yasuda, Y., Honda, H., Kuwatsuka, Y., Kato, S., Kondo, T., Iwata, M., Nakashima, T., Yasui, H., Takamatsu, H., Okajima, H., Yoshida, Y., & Maruyama, S. (2018). Effect of dietary energy and polymorphisms in BRAP and GHRL on obesity and metabolic traits. *Obesity Research & Clinical Practice*, *12*(1, Supplement 1), 39–48. https://doi.org/10.1016/j.orcp.2016.05.004

Inoue, H., Hisamoto, N., An, J. H., Oliveira, R. P., Nishida, E., Blackwell, T. K., & Matsumoto, K. (2005). The C. elegans p38 MAPK pathway regulates nuclear localization of the transcription factor SKN-1 in oxidative stress response. *Genes & Development*, *19*(19), 2278–2283. https://doi.org/10.1101/gad.1324805

Ishitani, T., Ninomiya-Tsuji, J., Nagai, S., Nishita, M., Meneghini, M., Barker, N., Waterman, M., Bowerman, B., Clevers, H., Shibuya, H., & Matsumoto, K. (1999). The TAK1–NLK–MAPK-related pathway antagonizes signalling between β-catenin and transcription factor TCF. *Nature*, *399*(6738), 798–802. https://doi.org/10.1038/21674

Jain, N., Iyer, K. V., Kumar, A., & Shivashankar, G. V. (2013). Cell geometric constraints induce modular gene-expression patterns via redistribution of HDAC3 regulated by actomyosin contractility. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(28), 11349–11354. https://doi.org/10.1073/pnas.1300801110

Jongbloets, B. C., & Pasterkamp, R. J. (2014). Semaphorin signalling during development. *Development*, *141*(17), 3292–3297. https://doi.org/10.1242/dev.105544

Jorgensen, E. M., & Mango, S. E. (2002). The art and design of genetic screens: Caenorhabditis elegans. *Nature Reviews Genetics*, *3*(5), 356–369. https://doi.org/10.1038/nrg794

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*(6920), 231–237. https://doi.org/10.1038/nature01278

Kelly, J., Moyeed, R., Carroll, C., Luo, S., & Li, X. (2020). Genetic networks in Parkinson's and Alzheimer's disease. *Aging (Albany NY)*, *12*(6), 5221–5243. https://doi.org/10.18632/aging.102943

Kim, J. W., Choe, Y. M., Shin, J.-G., Park, B. L., Shin, H.-D., Choi, I.-G., & Lee, B. C. (2018). Associations of BRAP polymorphisms with the risk of alcohol dependence and scores on the Alcohol Use Disorders Identification Test. *Neuropsychiatric Disease and Treatment*, *15*, 83–94. https://doi.org/10.2147/NDT.S184067

Kinzler, K. W., & Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell*, *87*(2), 159–170. https://doi.org/10.1016/s0092-8674(00)81333-1

Knight, C. G., Azevedo, R. B., & Leroi, A. M. (2001). Testing life-history pleiotropy in Caenorhabditis elegans. *Evolution; International Journal of Organic Evolution*, *55*(9), 1795–1804. https://doi.org/10.1111/j.0014-3820.2001.tb00828.x

Koon, J. C., & Kubiseski, T. J. (2010). Developmental Arrest of Caenorhabditis elegans BRAP-2 Mutant Exposed to Oxidative Stress Is Dependent on BRC-1. *The Journal of Biological Chemistry*, *285*(18), 13437–13443. https://doi.org/10.1074/jbc.M110.107011

Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P. J., & Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genetics*, *19*(4), 379–383. https://doi.org/10.1038/1270

Korswagen, H. C., Herman, M. A., & Clevers, H. C. (2000). Distinct beta-catenins mediate adhesion and signalling functions in C. elegans. *Nature*, *406*(6795), 527–532. https://doi.org/10.1038/35020099

Kubo, Y., Imaizumi, T., Ando, M., Nakatochi, M., Yasuda, Y., Honda, H., Kuwatsuka, Y., Kato, S., Kikuchi, K., Kondo, T., Iwata, M., Nakashima, T., Yasui, H., Takamatsu, H., Okajima, H., Yoshida, Y., & Maruyama, S. (2017). Association between kidney function and genetic polymorphisms in atherosclerotic and chronic kidney diseases: A cross-sectional study in Japanese male workers. *PLoS ONE*, *12*(10). https://doi.org/10.1371/journal.pone.0185476

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Le Clainche, C., & Carlier, M.-F. (2008). Regulation of Actin Assembly Associated With Protrusion and Adhesion in Cell Migration. *Physiological Reviews*, *88*(2), 489–513. https://doi.org/10.1152/physrev.00021.2007

Lei, H., Liu, J., Fukushige, T., Fire, A., & Krause, M. (2009). Caudal-like PAL-1 directly activates the bodywall muscle module regulator hlh-1 in C. elegans to initiate the embryonic muscle gene regulatory network. *Development (Cambridge, England)*, *136*(8), 1241–1249. https://doi.org/10.1242/dev.030668

Levine, M., & Davidson, E. H. (2005). Gene regulatory networks for development. *Proceedings of the National Academy of Sciences*, *102*(14), 4936–4942. https://doi.org/10.1073/pnas.0408031102

Li, H. (2011a). Improving SNP discovery by base alignment quality. *Bioinformatics*, *27*(8), 1157–1158. https://doi.org/10.1093/bioinformatics/btr076

Li, H. (2011b). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Subgroup, 1000 Genome Project Data Processing. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., Ruan, J., & Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Research*, *18*(11), 1851–1858. https://doi.org/10.1101/gr.078212.108

Li, S., Ku, C.-Y., Farmer, A. A., Cong, Y.-S., Chen, C.-F., & Lee, W.-H. (1998). Identification of a Novel Cytoplasmic Protein That Specifically Binds to Nuclear Localization Signal Motifs. *Journal of Biological Chemistry*, *273*(11), 6183–6189. https://doi.org/10.1074/jbc.273.11.6183

Liao, Y.-C., Wang, Y.-S., Guo, Y.-C., Ozaki, K., Tanaka, T., Lin, H.-F., Chang, M.-H., Chen, K.-C., Yu, M.-L., Sheu, S.-H., & Juo, S.-H. H. (2011). BRAP Activates Inflammatory Cascades and Increases the Risk for Carotid Atherosclerosis. *Molecular Medicine*, *17*(9–10), 1065–1074. https://doi.org/10.2119/molmed.2011.00043

Lin, R., Hill, R. J., & Priess, J. R. (1998). POP-1 and anterior-posterior fate decisions in C. elegans embryos. *Cell*, *92*(2), 229–239. https://doi.org/10.1016/s0092-8674(00)80917-4

Lin, R., Thompson, S., & Priess, J. R. (1995). Pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. *Cell*, *83*(4), 599–609. https://doi.org/10.1016/0092-8674(95)90100-0

Little, S. C., Tikhonov, M., & Gregor, T. (2013). Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. *Cell*, *154*(4), 789–800. https://doi.org/10.1016/j.cell.2013.07.025

Liu, J., Phillips, B. T., Amaya, M. F., Kimble, J., & Xu, W. (2008). The C. elegans SYS-1 protein is a bona fide beta-catenin. *Developmental Cell*, *14*(5), 751–761. https://doi.org/10.1016/j.devcel.2008.02.015

MacQueen, A. J., Baggett, J. J., Perumov, N., Bauer, R. A., Januszewski, T., Schriefer, L., & Waddle, J. A. (2005). ACT-5 Is an Essential Caenorhabditis elegans Actin Required for Intestinal Microvilli Formation. *Molecular Biology of the Cell*, *16*(7), 3247–3259. https://doi.org/10.1091/mbc.E04-12-1061

Maduro, M. F. (2017). Gut development in C. elegans. *Seminars in Cell & Developmental Biology*, *66*, 3–11. https://doi.org/10.1016/j.semcdb.2017.01.001

Maduro, M. F., Broitman-Maduro, G., Choi, H., Carranza, F., Wu, A. C.-Y., & Rifkin, S. A. (2015). MED GATA factors promote robust development of the C. elegans endoderm. *Developmental Biology*, *404*(1), 66–79. https://doi.org/10.1016/j.ydbio.2015.04.025

Maduro, M. F., Broitman-Maduro, G., Mengarelli, I., & Rothman, J. H. (2007). Maternal deployment of the embryonic SKN-1—>MED-1,2 cell specification pathway in C. elegans. *Developmental Biology*, *301*(2), 590–601. https://doi.org/10.1016/j.ydbio.2006.08.029

Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R., & Rothman, J. H. (2005). Genetic redundancy in endoderm specification within the genus Caenorhabditis. *Developmental Biology*, *284*(2), 509–522. https://doi.org/10.1016/j.ydbio.2005.05.016

Maduro, M. F., Kasmir, J. J., Zhu, J., & Rothman, J. H. (2005). The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. *Developmental Biology*, *285*(2), 510–523. https://doi.org/10.1016/j.ydbio.2005.06.022

Maduro, M. F., Lin, R., & Rothman, J. H. (2002). Dynamics of a developmental switch: Recursive intracellular and intranuclear redistribution of Caenorhabditis elegans POP-1 parallels Wnt-inhibited transcriptional repression. *Developmental Biology*, *248*(1), 128–142. https://doi.org/10.1006/dbio.2002.0721

Maduro, M. F., Meneghini, M. D., Bowerman, B., Broitman-Maduro, G., & Rothman, J. H. (2001). Restriction of Mesendoderm to a Single Blastomere by the Combined Action of SKN-1 and a GSK-3β Homolog Is Mediated by MED-1 and -2 in C. elegans. *Molecular Cell*, *7*(3), 475–485. https://doi.org/10.1016/S1097-2765(01)00195-2

Maduro, M. F., & Rothman, J. H. (2002). Making Worm Guts: The Gene Regulatory Network of the Caenorhabditis elegans Endoderm. *Developmental Biology*, *246*(1), 68–85. https://doi.org/10.1006/dbio.2002.0655

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology*, *7*(11), e1002255. https://doi.org/10.1371/journal.pcbi.1002255

Mammoto, T., & Ingber, D. E. (2010). Mechanical control of tissue and organ development. *Development*, *137*(9), 1407–1420. https://doi.org/10.1242/dev.024166

Matheny, S. A., Chen, C., Kortum, R. L., Razidlo, G. L., Lewis, R. E., & White, M. A. (2004). Ras regulates assembly of mitogenic signalling complexes through the effector protein IMP. *Nature*, *427*(6971), 256–260. https://doi.org/10.1038/nature02237

Matheny, S. A., & White, M. A. (2009). Signaling Threshold Regulation by the Ras Effector IMP. *The Journal of Biological Chemistry*, *284*(17), 11007–11011. https://doi.org/10.1074/jbc.R800082200

McBeath, R., Pirone, D. M., Nelson, C. M., Bhadriraju, K., & Chen, C. S. (2004). Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Developmental Cell*, *6*(4), 483–495. https://doi.org/10.1016/s1534-5807(04)00075-9

McGhee, J. D., Fukushige, T., Krause, M. W., Minnema, S. E., Goszczynski, B., Gaudet, J., Kohara, Y., Bossinger, O., Zhao, Y., Khattra, J., Hirst, M., Jones, S. J. M., Marra, M. A., Ruzanov, P., Warner, A., Zapf, R., Moerman, D. G., & Kalb, J. M. (2009). ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. *Developmental Biology*, *327*(2), 551–565. https://doi.org/10.1016/j.ydbio.2008.11.034

McGhee, J. D., Sleumer, M. C., Bilenky, M., Wong, K., McKay, S. J., Goszczynski, B., Tian, H., Krich, N. D., Khattra, J., Holt, R. A., Baillie, D. L., Kohara, Y., Marra, M. A., Jones, S. J. M., Moerman, D. G., & Robertson, A. G. (2007). The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. *Developmental Biology*, *302*(2), 627–645. https://doi.org/10.1016/j.ydbio.2006.10.024

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122. https://doi.org/10.1186/s13059-016-0974-4

Mello, C. C., Draper, B. W., Krause, M., Weintraub, H., & Priess, J. R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. *Cell*, *70*(1), 163–176. https://doi.org/10.1016/0092-8674(92)90542-k

Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., & Priess, J. R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. *Nature*, *382*(6593), 710–712. https://doi.org/10.1038/382710a0

Meneghini, M. D., Ishitani, T., Carter, J. C., Hisamoto, N., Ninomiya-Tsuji, J., Thorpe, C. J., Hamill, D. R., Matsumoto, K., & Bowerman, B. (1999). MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in Caenorhabditis elegans. *Nature*, *399*(6738), 793–797. https://doi.org/10.1038/21666

Mengarelli, I. (2006). *Investigation of Genes that Coordinate Cell Cycle and Development in Caenorhabditis elegans*. University of California, Santa Barbara.

Minniti, A. N., Sadler, C., & Ward, S. (1996). Genetic and molecular analysis of spe-27, a gene required for spermiogenesis in Caenorhabditis elegans hermaphrodites. *Genetics*, *143*(1), 213–223.

Narendra, D., Tanaka, A., Suen, D.-F., & Youle, R. J. (2008). Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *Journal of Cell Biology*, *183*(5), 795–803. https://doi.org/10.1083/jcb.200809125

Nicholas, W. L., Dougherty, E. C., & Hansen, E. L. (1959). Axenic Cultivation of Caenorharditis Briggsae (nematoda: Rhabditidae) with Chemically Undefined Supplements; Comparative Studies with Related Nematodes*. *Annals of the New York Academy of Sciences*, *77*(2), 218–236. https://doi.org/10.1111/j.1749-6632.1959.tb36902.x

Nigon, V. M., & Félix, M.-A. (2017). History of research on C. elegans and other free-living nematodes as model organisms. *WormBook: The Online Review of C. Elegans Biology*, *2017*, 1–84. https://doi.org/10.1895/wormbook.1.181.1

Nukazuka, A., Fujisawa, H., Inada, T., Oda, Y., & Takagi, S. (2008). Semaphorin controls epidermal morphogenesis by stimulating mRNA translation via eIF2α in Caenorhabditis elegans. *Genes & Development*, *22*(8), 1025–1036. https://doi.org/10.1101/gad.1644008

Ory, S., & Morrison, D. K. (2004). Signal Transduction: IMPlications for Ras-Dependent ERK Signaling. *Current Biology*, *14*(7), R277–R278. https://doi.org/10.1016/j.cub.2004.03.023

Ozaki, K., Sato, H., Inoue, K., Tsunoda, T., Sakata, Y., Mizuno, H., Lin, T.-H., Miyamoto, Y., Aoki, A., Onouchi, Y., Sheu, S.-H., Ikegawa, S., Odashiro, K., Nobuyoshi, M., Juo, S.-H. H., Hori, M., Nakamura, Y., & Tanaka, T. (2009). SNPs in BRAP associated with risk of myocardial infarction in Asian populations. *Nature Genetics*, *41*(3), 329–333. https://doi.org/10.1038/ng.326

Paix, A., Wang, Y., Smith, H. E., Lee, C.-Y. S., Calidas, D., Lu, T., Smith, J., Schmidt, H., Krause, M. W., & Seydoux, G. (2014). Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in Caenorhabditis elegans. *Genetics*, *198*(4), 1347–1356. https://doi.org/10.1534/genetics.114.170423

Pásti, G., & Labouesse, M. (2018). Epithelial junctions, cytoskeleton, and polarity. In *WormBook: The Online Review of C. elegans Biology [Internet]*. WormBook. https://www.ncbi.nlm.nih.gov/books/NBK19677/

Peter, I. S., & Davidson, E. H. (2011). Evolution of Gene Regulatory Networks that Control Embryonic Development of the Body Plan. *Cell*, *144*(6), 970–985. https://doi.org/10.1016/j.cell.2011.02.017

Phillips, B. T., Kidd, A. R., King, R., Hardin, J., & Kimble, J. (2007). Reciprocal asymmetry of SYS-1/β-catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, *104*(9), 3231–3236. https://doi.org/10.1073/pnas.0611507104

Piper, M., Anderson, R., Dwivedy, A., Weinl, C., van Horck, F., Leung, K. M., Cogill, E., & Holt, C. (2006). Signaling Mechanisms Underlying Slit2-Induced Collapse of Xenopus Retinal Growth Cones. *Neuron*, *49*(2), 215–228. https://doi.org/10.1016/j.neuron.2005.12.008

Pollard, T. D. (2016). Actin and Actin-Binding Proteins. *Cold Spring Harbor Perspectives in Biology*, *8*(8). https://doi.org/10.1101/cshperspect.a018226

Pollard, T. D., & Cooper, J. A. (2009). Actin, a Central Player in Cell Shape and Movement. *Science (New York, N.Y.)*, *326*(5957), 1208–1212. https://doi.org/10.1126/science.1175862

Polleux, F., Morrow, T., & Ghosh, A. (2000). Semaphorin 3A is a chemoattractant for cortical apical dendrites. *Nature*, *404*(6778), 567–573. https://doi.org/10.1038/35007001

Pulak, R. (2006). Techniques for Analysis, Sorting, and Dispensing of C. elegans on the COPAS^{TM} Flow-Sorting System. In K. Strange (Ed.), *C. elegans: Methods and Applications* (pp. 275–286). Humana Press. https://doi.org/10.1385/1-59745-151-7:275

Putzke, A. P., & Rothman, J. H. (2010). Repression of Wnt signaling by a Fer-type nonreceptor tyrosine kinase. *Proceedings of the National Academy of Sciences*, *107*(37), 16154–16159. https://doi.org/10.1073/pnas.1006600107

Raj, A., Rifkin, S. A., Andersen, E., & van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature*, *463*(7283), 913–918. https://doi.org/10.1038/nature08781

Rakha, E. A., El-Sheikh, S. E., Kandil, M. A., El-Sayed, M. E., Green, A. R., & Ellis, I. O. (2008). Expression of BRCA1 protein in breast cancer and its prognostic significance. *Human Pathology*, *39*(6), 857–865. https://doi.org/10.1016/j.humpath.2007.10.011

Reese, K. J., Dunn, M. A., Waddle, J. A., & Seydoux, G. (2000). Asymmetric Segregation of PIE-1 in C. elegans Is Mediated by Two Complementary Mechanisms that Act through Separate PIE-1 Protein Domains. *Molecular Cell*, *6*(2), 445–455. https://doi.org/10.1016/S1097-2765(00)00043-5

Riddle, M. R., Spickard, E. A., Jevince, A., Nguyen, K. C. Q., Hall, D. H., Joshi, P. M., & Rothman, J. H. (2016). Transorganogenesis and transdifferentiation in C. elegans are dependent on differentiated cell identity. *Developmental Biology*, *420*(1), 136–147. https://doi.org/10.1016/j.ydbio.2016.09.020

Riddle, M. R., Weintraub, A., Nguyen, K. C. Q., Hall, D. H., & Rothman, J. H. (2013). Transdifferentiation and remodeling of post-embryonic C. elegans cells by a single transcription factor. *Development (Cambridge, England)*, *140*(24), 4844–4849. https://doi.org/10.1242/dev.103010

Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y.-H., Ali, M., Priess, J. R., & Mello, C. C. (1997). Wnt Signaling and an APC-Related Gene Specify

Endoderm in Early C. elegans Embryos. *Cell*, *90*(4), 707–716. https://doi.org/10.1016/S0092-8674(00)80531-0

Rocheleau, C. E., Yasuda, J., Shin, T. H., Lin, R., Sawa, H., Okano, H., Priess, J. R., Davis, R. J., & Mello, C. C. (1999). WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in C. elegans. *Cell*, *97*(6), 717–726. https://doi.org/10.1016/s0092-8674(00)80784-9

Roy, B., Venkatachalapathy, S., Ratna, P., Wang, Y., Jokhun, D. S., Nagarajan, M., & Shivashankar, G. V. (2018). Laterally confined growth of cells induces nuclear reprogramming in the absence of exogenous biochemical factors. *Proceedings of the National Academy of Sciences*, *115*(21), E4741–E4750. https://doi.org/10.1073/pnas.1714770115

Rual, J.-F., Ceron, J., Koreth, J., Hao, T., Nicot, A.-S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., Heuvel, S. van den, & Vidal, M. (2004). Toward Improving Caenorhabditis elegans Phenome Mapping With an ORFeome-Based RNAi Library. *Genome Research*, *14*(10b), 2162–2168. https://doi.org/10.1101/gr.2505604

Ruf, V., Holzem, C., Peyman, T., Walz, G., Blackwell, T. K., & Neumann-Haefelin, E. (2013). TORC2 signaling antagonizes SKN-1 to induce C. elegans mesendodermal embryonic development. *Developmental Biology*, *384*(2), 214–227. https://doi.org/10.1016/j.ydbio.2013.08.011

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protocols*, *2006*(1). https://doi.org/10.1101/pdb.prot4455

Schierenberg, E. (1987). Reversal of cellular polarity and early cell-cell interaction in the embryo of Caenorhabditis elegans. *Developmental Biology*, *122*(2), 452–463. https://doi.org/10.1016/0012-1606(87)90309-5

Schulenburg, H., & Müller, S. (2004). Natural variation in the response of Caenorhabditis elegans towards Bacillus thuringiensis. *Parasitology*, *128*(Pt 4), 433–443. https://doi.org/10.1017/s003118200300461x

Segditsas, S., & Tomlinson, I. (2006). Colorectal cancer and genetic alterations in the Wnt pathway. *Oncogene*, *25*(57), 7531–7537. https://doi.org/10.1038/sj.onc.1210059

Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., Priess, J. R., & Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. *Nature*, *382*(6593), 713–716. https://doi.org/10.1038/382713a0

Sha, D., Chin, L.-S., & Li, L. (2010). Phosphorylation of parkin by Parkinson disease-linked kinase PINK1 activates parkin E3 ligase function and NF-kappaB signaling. *Human Molecular Genetics*, *19*(2), 352–363. https://doi.org/10.1093/hmg/ddp501

Shetty, P., Lo, M.-C., Robertson, S. M., & Lin, R. (2005). C. elegans TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. *Developmental Biology*, *285*(2), 584–592. https://doi.org/10.1016/j.ydbio.2005.07.008

Shin, T. H., Yasuda, J., Rocheleau, C. E., Lin, R., Soto, M., Bei, Y., Davis, R. J., & Mello, C. C. (1999). MOM-4, a MAP kinase kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in C. elegans. *Molecular Cell*, *4*(2), 275–280. https://doi.org/10.1016/s1097-2765(00)80375-5

Shoji, S., Hanada, K., Ohsawa, N., & Shirouzu, M. (2017). Central catalytic domain of BRAP (RNF52) recognizes the types of ubiquitin chains and utilizes oligo-ubiquitin

for ubiquitylation. *Biochemical Journal*, *474*(18), 3207–3226. https://doi.org/10.1042/BCJ20161104

Shook, D. R., Brooks, A., & Johnson, T. E. (1996). Mapping quantitative trait loci affecting life history traits in the nematode Caenorhabditis elegans. *Genetics*, *142*(3), 801–817.

Shook, D. R., & Johnson, T. E. (1999). Quantitative trait loci affecting survival and fertility-related traits in Caenorhabditis elegans show genotype-environment interactions, pleiotropy and epistasis. *Genetics*, *153*(3), 1233–1243.

Sivasundar, A., & Hey, J. (2003). Population genetics of Caenorhabditis elegans: The paradox of low polymorphism in a widespread species. *Genetics*, *163*(1), 147–157.

Smit, L., Baas, A., Kuipers, J., Korswagen, H., van de Wetering, M., & Clevers, H. (2004). Wnt activates the Tak1/Nemo-like kinase pathway. *The Journal of Biological Chemistry*, *279*(17), 17232–17240. https://doi.org/10.1074/jbc.M307801200

Sommermann, E. M., Strohmaier, K. R., Maduro, M. F., & Rothman, J. H. (2010). Endoderm development in C. elegans: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. *Developmental Biology*, *347*(1), 154–166. https://doi.org/10.1016/j.ydbio.2010.08.020

Stoeckius, M., Maaskola, J., Colombo, T., Rahn, H.-P., Friedländer, M. R., Li, N., Chen, W., Piano, F., & Rajewsky, N. (2009). Large scale sorting of C. elegans embryos reveals the dynamics of small RNA expression. *Nature Methods*, *6*(10), 745–751. https://doi.org/10.1038/nmeth.1370

Sulston, J. E., & Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology*, *56*(1), 110–156. https://doi.org/10.1016/0012-1606(77)90158-0

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology*, *100*(1), 64–119. https://doi.org/10.1016/0012-1606(83)90201-4

Sulston, J., & Hodgkin, J. (1988). Methods. In *The Nematode Caenorhabditis elegans* (W.B. Wood, ed., p. 587). Cold Spring Harbor Laboratory Press.

Sumiyoshi, E., Takahashi, S., Obata, H., Sugimoto, A., & Kohara, Y. (2011). The β-catenin HMP-2 functions downstream of Src in parallel with the Wnt pathway in early embryogenesis of C. elegans. *Developmental Biology*, *355*(2), 302–312. https://doi.org/10.1016/j.ydbio.2011.04.034

Takashima, O., Tsuruta, F., Kigoshi, Y., Nakamura, S., Kim, J., Katoh, M. C., Fukuda, T., Irie, K., & Chiba, T. (2013). Brap2 Regulates Temporal Control of NF-κB Localization Mediated by Inflammatory Response. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0058911

Takeuchi, F., Yokota, M., Yamamoto, K., Nakashima, E., Katsuya, T., Asano, H., Isono, M., Nabika, T., Sugiyama, T., Fujioka, A., Awata, N., Ohnaka, K., Nakatochi, M., Kitajima, H., Rakugi, H., Nakamura, J., Ohkubo, T., Imai, Y., Shimamoto, K., … Kato, N. (2012). Genome-wide association study of coronary artery disease in the Japanese. *European Journal of Human Genetics*, *20*(3), 333–340. https://doi.org/10.1038/ejhg.2011.184

Tenenhaus, C., Subramaniam, K., Dunn, M. A., & Seydoux, G. (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of Caenorhabditis elegans. *Genes & Development*, *15*(8), 1031–1040. https://doi.org/10.1101/gad.876201

Thorpe, C. J., Schlesinger, A., Carter, J. C., & Bowerman, B. (1997). Wnt Signaling Polarizes an Early C. elegans Blastomere to Distinguish Endoderm from Mesoderm. *Cell*, *90*(4), 695–705. https://doi.org/10.1016/S0092-8674(00)80530-9

Tian, H., Biehs, B., Chiu, C., Siebel, C., Wu, Y., Costa, M., de Sauvage, F. J., & Klein, O. D. (2015). Opposing activities of Notch and Wnt signaling regulate intestinal stem cells and gut homeostasis. *Cell Reports*, *11*(1), 33–42. https://doi.org/10.1016/j.celrep.2015.03.007

Torres Cleuren, Y. N. (2016). *Dissecting complex quantitative traits in Caenorhabditis elegans endoderm development*. The University of Auckland.

Torres Cleuren, Y. N., Ewe, C. K., Chipman, K. C., Mears, E. R., Wood, C. G., Al-Alami, C. E. A., Alcorn, M. R., Turner, T. L., Joshi, P. M., Snell, R. G., & Rothman, J. H. (2019). Extensive intraspecies cryptic variation in an ancient embryonic gene regulatory network. *ELife*, *8*. https://doi.org/10.7554/eLife.48220

Tran, T. S., Kolodkin, A. L., & Bharadwaj, R. (2007). Semaphorin Regulation of Cellular Morphology. *Annual Review of Cell and Developmental Biology*, *23*(1), 263–292. https://doi.org/10.1146/annurev.cellbio.22.010605.093554

Van der Flier, L. G., Sabates-Bellver, J., Oving, I., Haegebarth, A., De Palo, M., Anti, M., Van Gijn, M. E., Suijkerbuijk, S., Van de Wetering, M., Marra, G., & Clevers, H. (2007). The Intestinal Wnt/TCF Signature. *Gastroenterology*, *132*(2), 628–632. https://doi.org/10.1053/j.gastro.2006.08.039

van Es, J. H., Jay, P., Gregorieff, A., van Gijn, M. E., Jonkheer, S., Hatzis, P., Thiele, A., van den Born, M., Begthel, H., Brabletz, T., Taketo, M. M., & Clevers, H. (2005). Wnt signalling induces maturation of Paneth cells in intestinal crypts. *Nature Cell Biology*, *7*(4), 381–386. https://doi.org/10.1038/ncb1240

Varland, S., Vandekerckhove, J., & Drazic, A. (2019). Actin Post-translational Modifications: The Cinderella of Cytoskeletal Control. *Trends in Biochemical Sciences*, *44*(6), 502–516. https://doi.org/10.1016/j.tibs.2018.11.010

Velarde, N., Gunsalus, K. C., & Piano, F. (2007). Diverse roles of actin in C. elegans early embryogenesis. *BMC Developmental Biology*, *7*, 142. https://doi.org/10.1186/1471-213X-7-142

Vives-Bauza, C., Zhou, C., Huang, Y., Cui, M., Vries, R. L. A. de, Kim, J., May, J., Tocilescu, M. A., Liu, W., Ko, H. S., Magrané, J., Moore, D. J., Dawson, V. L., Grailhe, R., Dawson, T. M., Li, C., Tieu, K., & Przedborski, S. (2010). PINK1-dependent recruitment of Parkin to mitochondria in mitophagy. *Proceedings of the National Academy of Sciences*, *107*(1), 378–383. https://doi.org/10.1073/pnas.0911187107

Walston, T., Tuskey, C., Edgar, L., Hawkins, N., Ellis, G., Bowerman, B., Wood, W., & Hardin, J. (2004). Multiple Wnt Signaling Pathways Converge to Orient the Mitotic Spindle in Early C. elegans Embryos. *Developmental Cell*, *7*(6), 831–841. https://doi.org/10.1016/j.devcel.2004.10.008

Wiesenfahrt, T., Berg, J. Y., Osborne Nishimura, E., Robinson, A. G., Goszczynski, B., Lieb, J. D., & McGhee, J. D. (2016). The function and regulation of the GATA factor ELT-2 in the C. elegans endoderm. *Development (Cambridge, England)*, *143*(3), 483–491. https://doi.org/10.1242/dev.130914

Wiesenfahrt, T., Duanmu, J., Snider, F., Moerman, D., Au, V., Li-Leger, E., Flibotte, S., Parker, D. M., Marshall, C. J., Nishimura, E. O., Mains, P. E., & McGhee, J. D.

(2018). A Strategy To Isolate Modifiers of Caenorhabditis elegans Lethal Mutations: Investigating the Endoderm Specifying Ability of the Intestinal Differentiation GATA Factor ELT-2. *G3 (Bethesda, Md.)*, *8*(5), 1425–1437. https://doi.org/10.1534/g3.118.200079

Witze, E. S., Field, E. D., Hunt, D. F., & Rothman, J. H. (2009). C. elegans pur alpha, an activator of end-1, synergizes with the Wnt pathway to specify endoderm. *Developmental Biology*, *327*(1), 12–23. https://doi.org/10.1016/j.ydbio.2008.11.015

Wodarz, A., & Nusse, R. (1998). Mechanisms of Wnt signaling in development. *Annual Review of Cell and Developmental Biology*, *14*, 59–88. https://doi.org/10.1146/annurev.cellbio.14.1.59

WormBase: Nematode Information Resource. (n.d.). Retrieved October 27, 2020, from https://wormbase.org/#012-34-5

Wu, L., Xi, B., Hou, D., Zhao, X., Liu, J., Cheng, H., Shen, Y., Wang, X., & Mi, J. (2013). The single nucleotide polymorphisms in BRAP decrease the risk of metabolic syndrome in a Chinese young adult population. *Diabetes and Vascular Disease Research*, *10*(3), 202–207. https://doi.org/10.1177/1479164112455535

Yamada, M., Ohnishi, J., Ohkawara, B., Iemura, S., Satoh, K., Hyodo-Miura, J., Kawachi, K., Natsume, T., & Shibuya, H. (2006). NARF, an Nemo-like Kinase (NLK)-associated Ring Finger Protein Regulates the Ubiquitylation and Degradation of T Cell Factor/Lymphoid Enhancer Factor (TCF/LEF). *Journal of Biological Chemistry*, *281*(30), 20749–20760. https://doi.org/10.1074/jbc.M602089200

Yamada, Y., Sakuma, J., Takeuchi, I., Yasukochi, Y., Kato, K., Oguri, M., Fujimaki, T., Horibe, H., Muramatsu, M., Sawabe, M., Fujiwara, Y., Taniguchi, Y., Obuchi, S., Kawai, H., Shinkai, S., Mori, S., Arai, T., & Tanaka, M. (2017). Identification of polymorphisms in 12q24.1, ACAD10, and BRAP as novel genetic determinants of blood pressure in Japanese by exome-wide association studies. *Oncotarget*, *8*(26), 43068–43079. https://doi.org/10.18632/oncotarget.17474

Zhang, F., Liu, C., Xu, Y., Qi, G., Yuan, G., Cheng, Z., Wang, J., Wang, G., Wang, Z., Zhu, W., Zhou, Z., Zhao, X., Tian, L., Jin, C., Yuan, J., Zhang, G., Chen, Y., Wang, L., Lu, T., … Zhang, D. (2014). A Two-Stage Association Study Suggests BRAP as a Susceptibility Gene for Schizophrenia. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0086037

Zhao, Y., Wei, L., Shao, M., Huang, X., Chang, J., Zheng, J., Chu, J., Cui, Q., Peng, L., Luo, Y., Tan, W., Tan, W., Lin, D., & Wu, C. (2017). BRCA1-Associated Protein Increases Invasiveness of Esophageal Squamous Cell Carcinoma. *Gastroenterology*, *153*(5), 1304-1319.e5. https://doi.org/10.1053/j.gastro.2017.07.042

Zhong, W., & Sternberg, P. W. (2006). Genome-Wide Prediction of C. elegans Genetic Interactions. *Science*, *311*(5766), 1481–1484. https://doi.org/10.1126/science.1123287

Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., & Rothman, J. H. (1997). End-1 encodes an apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. *Genes & Development*, *11*(21), 2883–2896. https://doi.org/10.1101/gad.11.21.2883

Zuryn, S., & Jarriault, S. (2013). Deep sequencing strategies for mapping and identifying mutations from genetic screens. *Worm*, *2*(3). https://doi.org/10.4161/worm.25081