DISSERTATION

INVESTIGATION OF TRANSCRIPTIONAL DYNAMICS IN THE *CAENORHABDITIS ELEGANS* INTESTINE GENE REGULATORY NETWORK

Submitted by

Robert Thomas Patton Williams

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2022

Doctoral Committee:

Advisor: Erin Osborne Nishimura

Carol Wilusz
Jeffrey Hansen
Tom Santangelo

ABSTRACT

INVESTIGATION OF TRANSCRIPTIONAL DYNAMICS IN THE *CAENORHABDITIS ELEGANS* INTESTINE GENE REGULATORY NETWORK

ELT-2 is the major transcription factor required for the activation of *Caenorhabditis elegans* intestinal development. ELT-2 expression initiates in embryos to promote development and persists after hatching through larval and adult stages. Though the sites of ELT-2 binding have been defined and the transcriptional changes that result from ELT-2 depletion described, the intestine-specific transcriptome profile over developmental time has not been characterized, in part because of the difficulty in isolating intestine from other tissues. To address this knowledge gap, we used Fluorescence Activated Cell Sorting (FACS) to enrich intestine cells and performed RNA-seq analysis at distinct developmental stages. By linking the transcriptome profiles to previous ELT-2 studies, we were able to gain new insight into the role of ELT-2 in the intestinal regulatory network throughout development. Correlation of ELT-2 binding to the intestine transcriptome data, revealed that only 33% of intestine-enriched genes were direct targets of ELT-2 binding in embryos, but that number increased to 75% by the L3 stage. This suggests additional transcription factors may promote intestine-specific transcription early in development. Consistent with this possibility, half of the ELT-2 direct target genes were not transcriptionally dependent on ELT-2 for appropriate expression (i.e. their expression was not impacted following ELT-2 depletion) implying that other factors may compensate in the absence of ELT-2.

Among direct target genes that were affected by ELT-2 depletion, equal proportions were over and under-expressed thus ELT-2 can both activate and repress direct target genes. Both activated and repressed sets of ELT-2 target genes were enriched for defense response

genes reinforcing recent findings demonstrating ELT-2 participating in mediating the immune response upon pathogen exposure.

Fluorescent reporter assays demonstrated that expression of two direct targets of ELT-2 *cebp-1 and ets-4* are indeed repressed by ELT-2. Moreover, we observed that ELT-2 repressed its own promoter in a negative feedback loop that regulates *elt-2* gene expression. Together, our findings illustrate that ELT-2 contributes directly to roughly 20 – 50% of intestine-specific gene expression, that ELT-2 exerts both positive and negative regulatory control on its direct targets, and that our overall picture of the intestinal regulatory network is incomplete with more intestine specific transcription factors and mechanisms remaining to be discovered.

DEDICATION

This dissertation is dedicated to:

My father, Jimmie T. Williams

My uncle, Beto Patton

My friends, Christopher P. Allen and Yaw S. Owusu-Boaitey

To everyone that helped me along the way,

and to everyone I had the pleasure of helping.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Multicellular organisms begin life as a single cell which rapidly proliferates and diversifies into cells with unique forms and functions. Understanding the mechanisms that control the diversity of cells has been a long term goal in the field of developmental biology. This mystery is amplified with the knowledge that most cells in an organism share a common genome. How does cellular diversity arise when cells share a common genome?

To understand the molecular basis of cellular diversity for a developmental system, two key questions must be asked. First, what are the tissue-specific factors that make each cell type unique? In other words, what is the "parts list" of a cell? Second, what are the regulatory relationships between these factors that dictate how and when the genome is utilized. Recent advances in experimental and computational tools have allowed for perturbation and measurement of complex regulatory relationships to gain new insights.

This introduction will begin with a general overview of concepts from the fields of molecular gene expression, developmental biology, and systems biology because this dissertation sits at the intersection of these fields. Next, I will provide background on GATA transcription factors as key developmental regulators and an overview of the *C. elegans* intestine as a model for organogenesis. Finally, I will provide the detailed scientific premise for studying the developmental transcriptome of the *C. elegans* intestine and the GATA transcription factor ELT-2.

1.1 Basic concepts in gene expression, developmental biology, and systems biology

1.1.1 Gene expression

In the 1800s, Gregor Mendel investigated the characteristics of discrete inherited units with pea plants [1]. Many decades later, Hershey and Chase demonstrated that DNA was the molecular repository of genetic information linking Mendel's units of inheritance to physical structures [2]. After the molecular structure of DNA was identified in 1953, the central dogma of molecular biology proposed the transfer of genetic information from DNA to RNA and to proteins: DNA can be copied to DNA (replication), DNA can be copied into RNA (transcription), and proteins can be synthesized using the information in RNA (translation) [3]. Since then, the field of molecular biology has exploded with novel insight in the molecular mechanisms of gene expression.

The typical genetic structural organization is now well characterized, and the features of this organization give evidence of how gene transcription is carried out and controlled [4]. Genes are sequences of DNA nucleotides that are transcribed to produce a functional RNA molecule by the RNA Polymerase protein complex. Genes are a linear sequence of DNA nucleotides contained within a chromosome and are flanked with regulatory sequences of DNA that are proximal to the gene (promoters) or distal to the gene (enhancers). Both promoters and enhancers serve as landing sites for sequence-specific DNA binding proteins which recruit additional machinery necessary for transcription. Protein complexes such as RNA polymerase, general transcription factors, and the mediator complex are recruited to these regulatory sequences to copy the DNA template into an RNA molecule termed messenger RNA (mRNA) or transcript.

The regulated expression of genes underlies cell differentiation, the process that gives rise to diverse cell and tissue types over the course of development. Major orchestrators of cellular diversity include a class of proteins known as transcription factors (TFs), which serve to regulate gene transcription through activation or repression. TFs are typically composed of two

protein domains; A structured domain that directly binds a specific DNA sequence, and an unstructured domain that serves to recruit additional transcription machinery. Analysis of TFs has revealed a multitude of DNA binding domain structures that each are associated with a specific DNA sequence. The specificity between protein structure and DNA sequence is driven by non-covalent interactions between the protein, the DNA bases, and the DNA backbone.

1.1.2 Developmental biology

Developmental biology is the study of the cellular and molecular processes that facilitate plant and animal growth [5]. There are three key steps in development of a multicellular organism: 1) cell proliferation, 2) cell differentiation, and 3) morphogenesis. When an egg is fertilized, that single cell must divide to produce the multitude of cells that constitutes a multicellular organism's body. As the cells divide, they begin to differentiate and take on specialized roles for the body's tissues. Once specialized cells are formed, morphogenesis, the process of cell movement into distinct germ layers, initiates. Germ layers serve as the foundation for organ formation and include the ectoderm (skin and neural tissue), mesoderm (blood and muscle tissue), and endoderm (gastrointestinal and respiratory tract).

Of the three steps of development, this dissertation is centered in the process of cell differentiation. Examples of differentiated cells include neurons, epithelial cells, and hematopoietic cells. Differentiated cells typically utilize a fraction of the genes encoded in their genome which are required for their specific function and utilize their genome differently based on the designated cell type. The mechanism by which differentiated cells utilize their genomes is driven by complex networks of gene regulation – termed gene regulatory networks (GRNs) [6]. Overall, GRNs describe the interconnected networks between gene expression regulators and target genes which guide cells to take on unique functions through integration of intrinsic and extrinsic signals. GRNs lead to cellular diversity through the combinatorial control of transcription, RNA processing, localization and decay, and translation [7,8].

In addition to differentiation, this dissertation explores the concept of organogenesis. Organogenesis describes the process after morphogenesis of the germ layers where cells are organized into organ systems. Organs perform specific roles, such as digestion or cognition, and constitute a functional unit of an animal body. Organogenesis involves further differentiation of cells and requires signaling between germ layers. A major goal in the field is to determine what GRN components are required to drive organogenesis. By determining the GRN components and their interactions that guide normal organogenesis, we can better understand the mechanisms that lead to developmental abnormalities and diseases such as cancer.

1.1.3 Systems biology

Systems biology is a scientific framework designed to understand complex biological systems by integrating multiple sources of complex data. Systems biology arose during the 1990s and 2000s with advances in scientific methodology and computer technology [9]. Systems biology approaches are well suited for studying the complexities of GRNs. "-omics" datasets are typically utilized to answer questions in systems biology, which quantitatively measure the totality of biological molecules such as transcriptomics (RNA), proteomics (proteins), or metabolomics (metabolites) [10]. In systems biology, hypotheses are evaluated by integrating relevant datasets. Findings are then experimentally validated, and conclusions from these findings then refine the initial hypothesis and typically lead to new hypotheses.

Systems biology is dependent on the collection of genome-wide data. Two genome-wide assays utilized in this dissertation include RNA sequencing (RNA-seq) and chromatin immunoprecipitation with sequencing (ChIP-seq) [11–13]. Both assays are dependent on advances in high-throughput DNA sequencing technology such as the Illumina sequencing platform. RNA-seq involves extracting RNA from cell populations and measuring the RNA abundance for all genes in the genome. RNA-seq is particularly useful for determining the transcriptome response

after a treatment or identifying how a transcriptome differs between cell types. RNA-seq can be performed on bulk populations of cells or on many single cells (termed scRNA-seq) [14,15].

In contrast, ChIP-seq provides a map of where a chromatin-associated protein localizes throughout the genome. ChIP-seq is used to identify the global binding sites of proteins such as TFs, or histones with specific post-translational modifications. ChIP-seq requires crosslinking protein to DNA, shearing the genome and isolating DNA bound to a protein of interest utilizing an antibody specific to the protein. Together, RNA-seq and ChIP-seq have become essential tools for investigating GRNs [16,17]. They allow the identification of the factors within the network, mapping of where key factors are located throughout the genome, and evaluation of how the genome is regulated providing an unprecedented opportunity to dissect GRNs that guide cell differentiation and organogenesis.

1.2 GATA transcription factors and the *C. elegans* intestine

1.2.1 GATA transcription factors

GATA TFs are evolutionarily conserved among animals, plants and fungi and play key roles in regulation of cell differentiation and specification during tissue development and organogenesis. They are also essential for the development of tissues derived from all three germ layers [18]. GATA TFs are named after the consensus DNA sequence (A/T)GATA(A/G) which is recognized by their zinc-finger domains. Vertebrates possess six GATA paralogs that are divided into two subfamilies based on initial discovery of their spatial and temporal expression patterns: GATA1/2/3 for the hematopoietic system and GATA4/5/6 for the cardiac system. Consistent with their role as key developmental regulators, alterations in GATA TFs lead to a variety of human diseases. For instance, GATA4 and GATA6 genes are amplified in gastric cancers and are associated with misregulation of genes involved in processes such as cell movement, death, and survival [19,20].

A variety of molecular mechanisms describe the modes of GATA TF regulation [21]. Of importance to this dissertation is the ability for GATA TFs to serve as both activators and repressors of target gene expression depending on genomic context. For example, in the mouse small intestine GATA4 activates genes that promote jejunal identity while simultaneously inhibiting an ileal identity. GATA4 promotes ileal identity by activating genes associated with transcription-related processes and repressing genes associated with cell death, signal transduction, the cytoskeleton, and lipid metabolism [22]. A suite of 14 sequence motifs were associated with GATA4 activated target genes, while one motif corresponding to nuclear receptor transcription factor Nr1d1 was associated with GATA4 repressed target genes. Additionally, GATA1 has a similar dual activator and repressor role as an essential transcription factor for red blood cell development [23,24]. Genome-wide studies performed in mouse models demonstrated that the mode of regulation for GATA1 target genes is dependent on the integration of regulatory control on a locus-specific basis similar to GATA4. For instance, GATA1 activated genes are associated with co-binding of GATA1 with TAL1 to gene promoters that contain a composite DNA motif for both TFs, known as a GATA/E-box motif [25–27]. Alternatively, repressed GATA1 gene promoters are associated with co-binding of GATA1 and LDB1, lack TAL1, and are marked with PRC2 catalyzed histone post-translational modification H2K27me3 [28,29]. These examples demonstrate the dual role of GATA1 in facilitating simultaneous activation and repression gene regulation to drive red blood cell development. Overall, the mode of regulation for GATA TF target genes appear to be dependent on a combination of local DNA regulatory elements and the suite of available transcription factors within the cell.

While much is known about GATA TFs, a major challenge remains in understanding the totality of GATA TF regulation in the complex GRNs they participate in [18,21]. For instance, only a subset of the regulatory connections in GATA TF controlled GRNs have been characterized in detail. The continued use of systems biology approaches will further clarify how GATA TFs

regulate their GRNs and provide valuable insight in their molecular role in developmental diseases.

1.2.2 C. elegans as a model organism

C. elegans is useful for studying GRNs in developmental biology for several reasons. First, it is a multicellular organism composed of just five major organs: epidermis, muscle, digestive system, nervous system and reproductive system [30]. Next, *C. elegans* has as an invariant cell lineage, where cell divisions that lead to organ development are predictable [31]. Finally, 60-80% of human genes have an ortholog in the *C. elegans* genome which makes most discoveries in *C. elegans* relevant to the study of human health and disease [32]. It has been proposed that genetically and biologically facile model organisms such as *C. elegans* can be used to "infer the function of human genes and to place these genes in the context of their informational pathway" [9].

C. elegans was established as a model organism by Sydney Brenner in 1963 for the fields of developmental biology and neurobiology [33]. Additional features make *C. elegans* a powerful model organism. It is a free-living soil-dwelling nematode that is simple to culture in the lab and has a rapid 3 to 4 day lifecycle from egg to egg-laying adult. *C. elegans* is a self-fertilizing hermaphroditic species with a low frequency male population allowing for genetic crosses and has facile genetic manipulation techniques. Finally, *C. elegans* was the first multicellular organism to have its whole genome sequenced and there is robust scientific community dedicated to studying it [34,35].

1.2.3 The C. elegans intestine gene regulatory network

In addition to being the site of many biological functions ranging from digestion to the immune response and home of the gut microbiome the *C. elegans* intestine provides an attractive model for organogenesis [36,37]. Additionally, the developmental progression that gives

rise to the *C. elegans* intestine is well established (Figure 1.1). The intestine arises clonally from a single cell, the E cell, and ultimately produces a 20-cell-long tube that extends from the pharynx to the rectum. At the 4-cell stage of embryogenesis, the endomesoderm is specified when a non-canonical Wnt-signaling pathway responds to positional cues from the posterior P cell. This alleviates endoderm repression orchestrated by SKN-1 [38–40]. Beginning in the E-cell (8-cell stage of embryogenesis), transient pulses of GATA TF pairs, first MED-1/MED-2 then END-1/END-3, lead to the eventual expression of the final pair ELT-2 and ELT-7 [41–46].

ELT-2 and ELT-7 are partially redundant TFs expressed through embryonic, larval, and adult stages, only declining in aged, post-fertile worms [45,47–49]. Evolutionary analysis has demonstrated that *C. elegans* intestine GATA TFs are most closely related to the human GATA4/5/6 subfamily [50]. After initial ELT-2 expression, the intestine continues to divide and grow, undergoing a further 3-4 rounds of cell division. During these divisions, the E lineage ingresses, aligns into two rows of cells down the embryo center, fuses, and finally creates a lumen to form the alimentary canal [36]. The expression of ELT-2 and its homolog ELT-7 are then sustained throughout the organism's life [45,51]. Sustained intestine expression of ELT-2 and ELT-7 is a result of auto- and cross-regulatory feedback between these two key transcription factors [45]. Additionally, ELT-2 contributes to diverse intestinal processes such as digestion, immunity, detoxification, and aging [49,52–58].

1.3 Rationale and Hypotheses

1.3.1 Tissue-specific transcriptional dynamics of the developing C. elegans intestine

A major effort in biology involves determining the complete atlas of transcription factors, their expression patterns, their regulatory networks, and how their downstream effects interplay to generate a functional organ [59–61]. While the gene regulatory network in early *C. elegans* intestine development is well characterized through the 4-cell intestine stage, the GRN components after early intestine development are currently unknown[42]. This leads us to our goal

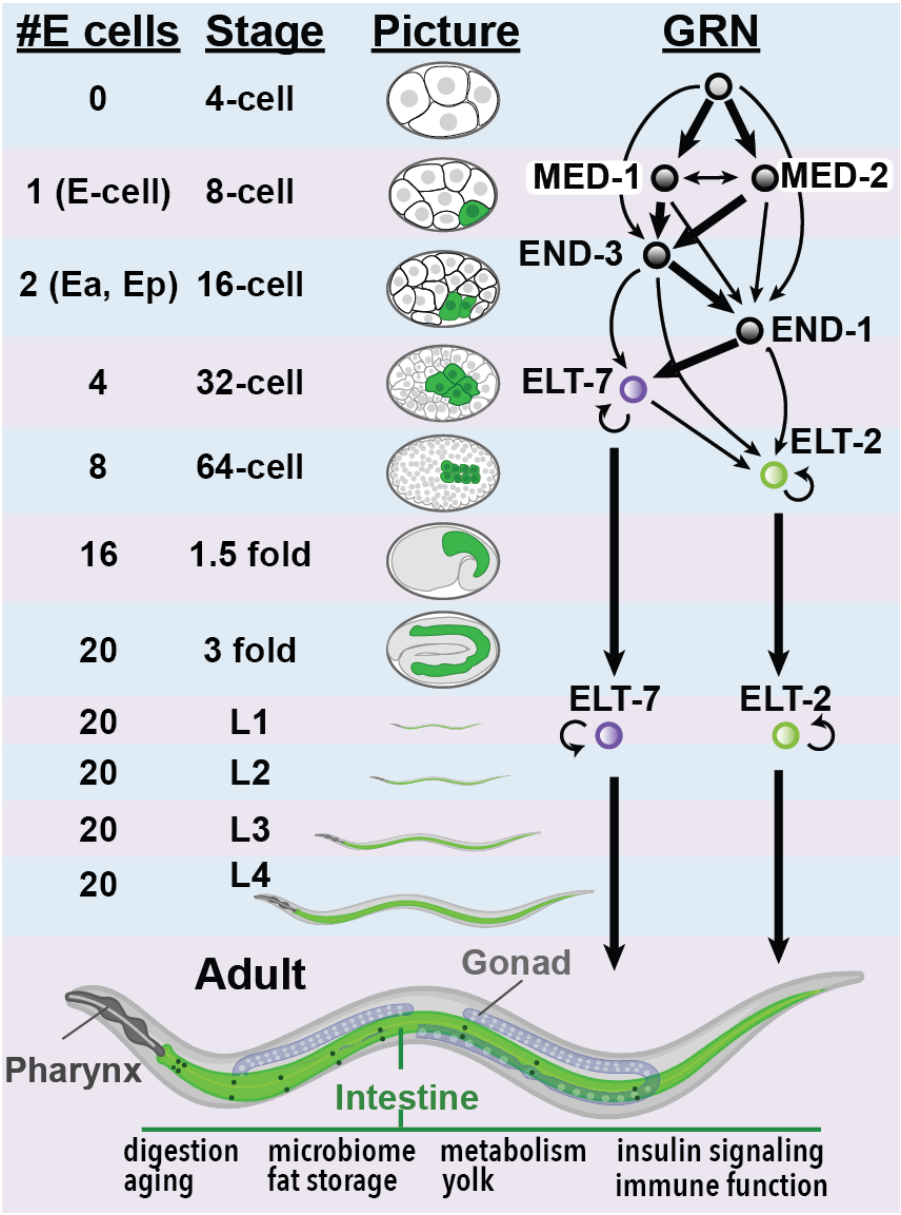

Figure 1.1: A summary of the known intestine GRN components and biological functions that the intestine participates in.

of determining the intestine specific transcriptome over developmental time. Measuring the intestine transcriptome will provide an atlas of intestine gene expression and aid in understanding the gene regulatory components necessary for intestine organogenesis.

Previous work characterizing the intestine specific transcriptome has been successful. These include single cell dissection of embryonic cells, RNA immunoprecipitation techniques, and intestine purification through Fluorescence Activated Cell Sorting (FACS), primarily in the embryo, L1 and L2 stages[5–7]. However, differences in tissue isolation methods and RNA abundance measurements (RNA-seq vs microarray) introduce technical variation that creates barriers to integrating previous results. Major differences in sensitivity hinder the biological relevance of meta-analyses integrating data from microarray and RNA-seq experiments [62–66]. Overall, the field currently lacks a comprehensive view of *C. elegans* intestine gene expression through development.

Our goal was to circumvent these issues by standardizing the isolation and transcriptome measurement practices between developmental stages. In this dissertation I have developed Fluorescence Activated Cell Sorting (FACS) methodology to isolate intestine tissue from three stages of *C. elegans* intestine development and measured the intestine transcriptome with either bulk RNA-seq or single-cell RNA-seq. This dataset provides an atlas of intestine expressed genes which will further the understanding of molecular factors involved in driving intestine organogenesis. One major goal of this dissertation was to develop the approaches and methodologies necessary to isolate intestine cells in embryonic and post-embryonic *C. elegans* through Fluorescence Activated Cell Sorting (FACS). By allowing access to biochemical and molecular biology techniques in a tissue specific manner, this methodology allow dissection of the intestine GRN with additional genome-wide tools providing new avenues for hypothesis generation and experimental perturbation.

1.3.2 Genome-wide characterization of the regulatory role of the C. elegans GATA transcription factor ELT-2

ELT-2 promotes intestine-specific gene expression, its loss leads to a larval lethal intestinal phenotype with collapsed intestine lumen, and its ectopic expression induces intestinal features marked by *ges-1* expression in non-intestine cells, all characteristics underscoring the central role ELT-2 plays in intestine development [44,48,67–71]. However, ELT-2 is induced by and co-expressed with the other members of the intestinal GRN that together promote cellular specification, commitment, differentiation, and downstream biological processes [39,42]. Because ELT-2 DNA binding sites are over-represented in the promoters of genes expressed in both embryonic and larval stage intestines, ELT-2 was initially suggested to activate expression of all intestinal genes [48]. However, ChIP-seq assays of ELT-2 have revealed that ELT-2 binding is detected at a small fraction of potential binding sites, echoing a general theme in the TF field that sequence alone does not dictate TF occupancy [38,48,49,72–75]. Thus to date it remains unclear whether ELT2 binding is necessary and sufficient for intestinal expression. Additionally, previous work has demonstrated that ELT-2 regulates its own expression, and the presence of ELT-2 protein at its own gene promoter has been extensively characterized [38,76]. The purpose of this regulation was hypothesized to be to sustain ELT-2 expression through autoactivation until it declines in old age [49]. However, the ELT-2 positive autoregulation hypothesis had yet to be evaluated by removing ELT-2 from the biological system.

We have performed a systems biology investigation in the regulatroy role of ELT-2 in the developing intestine to determine how ELT-2 regulates the intestine GRN. Through integration of three key datasets we set out to evaluate three prevailing hypotheses regarding ELT-2 function: 1) ELT-2 binding is associated with expression of all intestine genes, and 2) ELT-2 functions solely as a transcriptional activator, and 3) ELT-2 engages in positive autoregulation. We found that only 33% of embryo stage intestine enriched genes had observable ELT-2 occupancy in ChIP-seq assays, but that this number increased to 75% in the L3 stage,

illustrating that intestinal gene expression in the embryo is either indirectly associated with ELT-2 expression or independent from it. Further, our analysis suggests that ELT-2 both positively and negatively regulates subsets of direct target genes. This finding led us to reevaluate the hypothesis that ELT-2 performs positive autoregulation and demonstrated that ELT-2 negatively regulates target genes as well as its own promoter. By using systems biology approaches we have discovered additional roles of ELT-2 transcriptional regulation and refined the role ELT-2 plays in regulating the intestine GRN. The approaches we undertook here may serve as a roadmap to study other TFs that contribute to diverse *C. elegans* GRNs.

CHAPTER 2

GENOME-WIDE CHARACTERIZATION OF THE *CAENORHABDITIS ELEGANS* INTESTINE GATA TRANSCRIPTION FACTOR ELT-2[1]

2.1 Introduction

Transcription factors (TFs) work together to form gene regulatory networks (GRNs) that direct expression of target genes [6,77,78]. These GRNs form positive, negative, feed-forward, and auto-inhibitory connections to orchestrate downstream transcriptional responses critical for development and health [79–82]. The *Caenorhabditis elegans* intestine is a model for understanding how GRNs contribute to organogenesis [36,39,42]. In embryos, the intestinal GRN initiates through the combined action of maternally loaded transcription factors and inductive cues. The embryonic intestinal GRN is comprised of successive pairs of GATA transcription factors that culminate in the expression of the GATA transcription factor ELT-2 [36,43]. ELT-2 promotes intestine-specific gene expression, its loss leads to a larval lethal intestinal phenotype, and its misexpression ectopically induces intestinal features, all characteristics underscoring ELT-2's central role [44,48,67–71]. However, ELT-2 is induced by and co-expressed with the other members of the intestinal GRN that together promote cellular specification, commitment, differentiation, and downstream biological processes [39,42]. A missing aspect of the *C. elegans* intestine GRN model is how it changes over developmental time, especially beyond embryonic stages. Many TFs within the described embryonic network are transiently expressed, though ELT-2 and its partner ELT-7 persist for the duration of the worm's lifespan, prompting questions as to how the intestine transcriptional landscape is controlled over developmental time.

[1] This chapter is a modified version of a manuscript under review

The developmental progression that gives rise to the *C. elegans* intestine is well established. The intestine arises clonally from a single cell, the E cell, and ultimately produces a 20-cell-long tube that extends from the animal's pharynx to its rectum. At the 4-cell stage, specification of the endomesoderm occurs when a non-canonical Wnt-signaling pathway responds to positional cues to alleviate endoderm repression orchestrated by SKN-1 [38–40]. Beginning in the E-cell (8-cell stage of embryogenesis), transient pulses of GATA TF pairs, first MED-1/MED-2 then END-1/END-3, lead to the eventual expression of the final pair ELT-2 and ELT-7 [41–46]. ELT-2 and ELT-7 are partially redundant TFs whose expression is sustained through embryonic, larval, and adult stages, only declining in aged, post-fertile worms [45,47–49]. After initial ELT-2 expression, the intestine continues to divide and grow, undergoing a further 3-4 rounds of cell division. During these divisions, the E lineage ingresses, aligns into two rows of cells down the embryo center, fuses, and finally creates a lumen to form the alimentary canal [36]. During larval and adult stages, ELT-2 continues to mark intestinal identity and contributes to diverse intestinal processes such as digestion, immunity, detoxification, and aging [49,52–58].

Previous analysis identified that the intestine transcriptome is developmentally dynamic as only 20% of genes are shared between embryonic and larval stages as initially assessed by SAGE analysis, a forerunner of microarrays [48,68]. These changes illustrate the multiple functions the intestine undertakes upon hatching. Because ELT-2 DNA binding sites (TGATAA) are over-represented in the promoters of genes expressed in both embryonic and larval stages, ELT-2 was initially suggested to activate all intestinal genes. However, ChIP-seq assays of ELT-2 have revealed that ELT-2 binds only a small fraction of potential binding sites, echoing a general theme in the TF field that sequence alone does not dictate TF occupancy [38,48,49,72–75]. Chromatin, flanking sequences, long-range interactions, and combinatorial binding also contribute to the final set of genomic loci a TF inhabits. Indeed, recent evidence suggests a quantitative relationship between ELT-2 *cis*-regulatory site binding affinity and transcriptional strength [83]. Moreover, TFs differ in the degree to which their occupancy at a genomic locus leads to

functional transcriptional activation [84]. Altogether, differences in intestinal transcriptomics over time could be governed by a host of different molecular changes at the chromatin, DNA, or transcription factor levels.

To date, a systems biology investigation of ELT-2's regulatroy role in the developing intestine utilizing whole genome approaches has yet to be performed as the field has lacked a comprehensive transcriptome profile of the intestine over developmental time. To amend this issue, we characterized intestinal transcriptomes in embryo, L1 and L3 stages by Fluorescence Activated Cell Sorting (FACS) followed by bulk RNA-seq. We then combined that dataset with publically available datasates including ELT-2 DNA binding maps (ChIP-seq) in embryo, L1 and L3 stages, and the whole genome transcriptional response to *elt-2* depletion (RNA-seq on whole L1 worms with or without *elt-2* function) [47,72].

Through integration of the newly generated set of intestine enriched genes with ELT-2 binding maps and the transcriptional response to *elt-2* deletion, we set out to evaluate three prevailing hypotheses regarding ELT-2 function: 1) ELT-2 binding is associated with expression of all intestine genes, 2) ELT-2 functions soley as a transcriptional activator, and 3) ELT-2 performs positive autoregulation. We found that only 33% of embryo stage intestine enriched genes had observable ELT-2 occupancy in ChIP-seq assays, but that this number increased to 75% in the L3 stage, illustrating that intestinal gene expression in the embryo is either indirectly associated with ELT-2 expression or independent from it. Further, our analysis suggests that ELT-2 both positively and negatively regulates subsets of direct target genes. This finding led us to reevaluate the hypothesis that ELT-2 performs positive autoregulation and demonstrated that ELT-2 negatively regulates target genes as well as its own promoter. By using systems biology approaches we have discovered additional roles of ELT-2 transcriptional regulation. As more ChIP-seq datasets become available through the modERN Resource (model organism Encyclopedia of Regulatory Networks) [72], the approaches we undertook here may serve as a roadmap to study other TFs that contribute to diverse *C. elegans* GRNs.

2.2 **Results**

2.2.1 C. elegans intestine transcriptional profiling by FACS isolation in embryonic and post-embryonic stages

To evaluate the role of ELT2 in the gene regulatory network of the developing intestine, we measured the intestine-specific transcriptome across three stages: embryo, L1 and L3. Previous work characterizing the intestine-specific transcriptome at individual or mixed stages has been successful. These include dissection of embryos into single cells, RNA immunoprecipitation, and intestine cell purification through Fluorescence Activated Cell Sorting (FACS), primarily in the embryo, L1 and L2 stages [85–90]. However, those studies had been conducted on mixed-stage worms or at only a single developmental stage with differences in tissue isolation and assay methods (RNA-seq vs microarray) preventing integration of these results across stages. Therefore, a comprehensive view of intestine gene expression through developmental time was missing.

Here, we have taken a FACS-based approach to separate intestine cells from non-intestine cells in embryo, L1 and L3 developmental stages for RNA-seq analysis (Figure 2.1A). Intestine cells were labeled with transgenic reporter *elt-2p::H2B::GFP* which produces histone fused GFP localizing to intestine nuclei (Figure 2.1B). Worm dissociation protocols and FACS gating strategies were optimized for intestine cell isolation in each stage (see Methods, Appendix B, Figure 2.2A-C). A flow cytometry profile of dissociated cells detected ~2% of cells as GFP-labeled which corresponds to the number of intestine cells in the worm (20/~1000 = 2%). GFP labeled intestine cells were then isolated by FACS with a purity of 87% (Figure 2.1C) and verified visually (Figure 2.1D).

RNA-seq libraries were generated from intestine cell and non-intestine cell populations. To determine the purity of isolated intestine cells, we compared the RNA-seq enrichment score

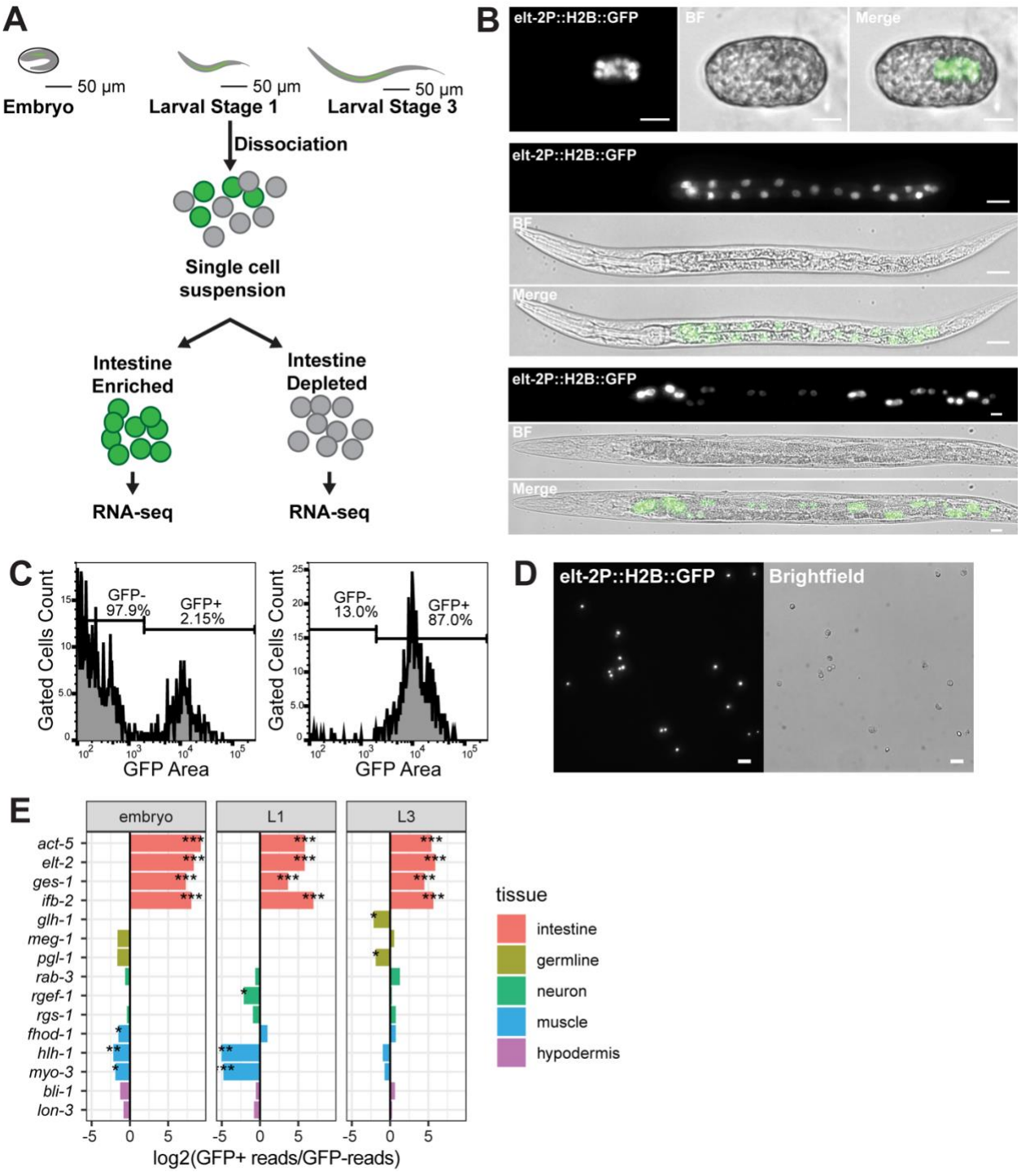

Figure 2.1. *C. elegans* intestine transcriptional profiling by FACS isolation in embryonic and post-embryonic stages. (A) Schematic depicting the procedure for measuring the C. elegans intestine in embryo, L1 and L3 stages investigated in this study. Worms were dissociated into single cell suspension with chitinase treatment and mechanical disruption (embryos) or SDS-DTT treatment and mechanical disruption (L1, L3). Cells were separated into intestine and non-intestine samples with FACS. RNA-seq was then performed on FACS isolated

cells. **(B)** Visualization of the intestine reporter transgene (*caIs71[elt-2p::GFP::HIS-2B]* allele) utilized in this study at embryo (top), L1 (middle) and L3 (bottom) stages (60x magnification, scale bars = 20 microns). **(C)** Histogram of dissociated C. elegans cells before (left) and after (right) FACS purification. GFP signal gates were set using wildtype strain N2. Percentage of GFP+ and GFP- cells was quantitated (n > 5,000). Analysis and FACS was performed on strain JM149 (elt-2P::H2B::GFP). **(D)** Visualization of GFP+ intestine cells (20x magnification, scale bars = 20 microns). **(E)** GFP+ (intestine) vs GFP- (non-intestine) log2 transformed RNA-seq read fold change for tissue specific genes (*P-value < 0.01, **P-value < 1x10−5, ***P-value < 1x10−10).

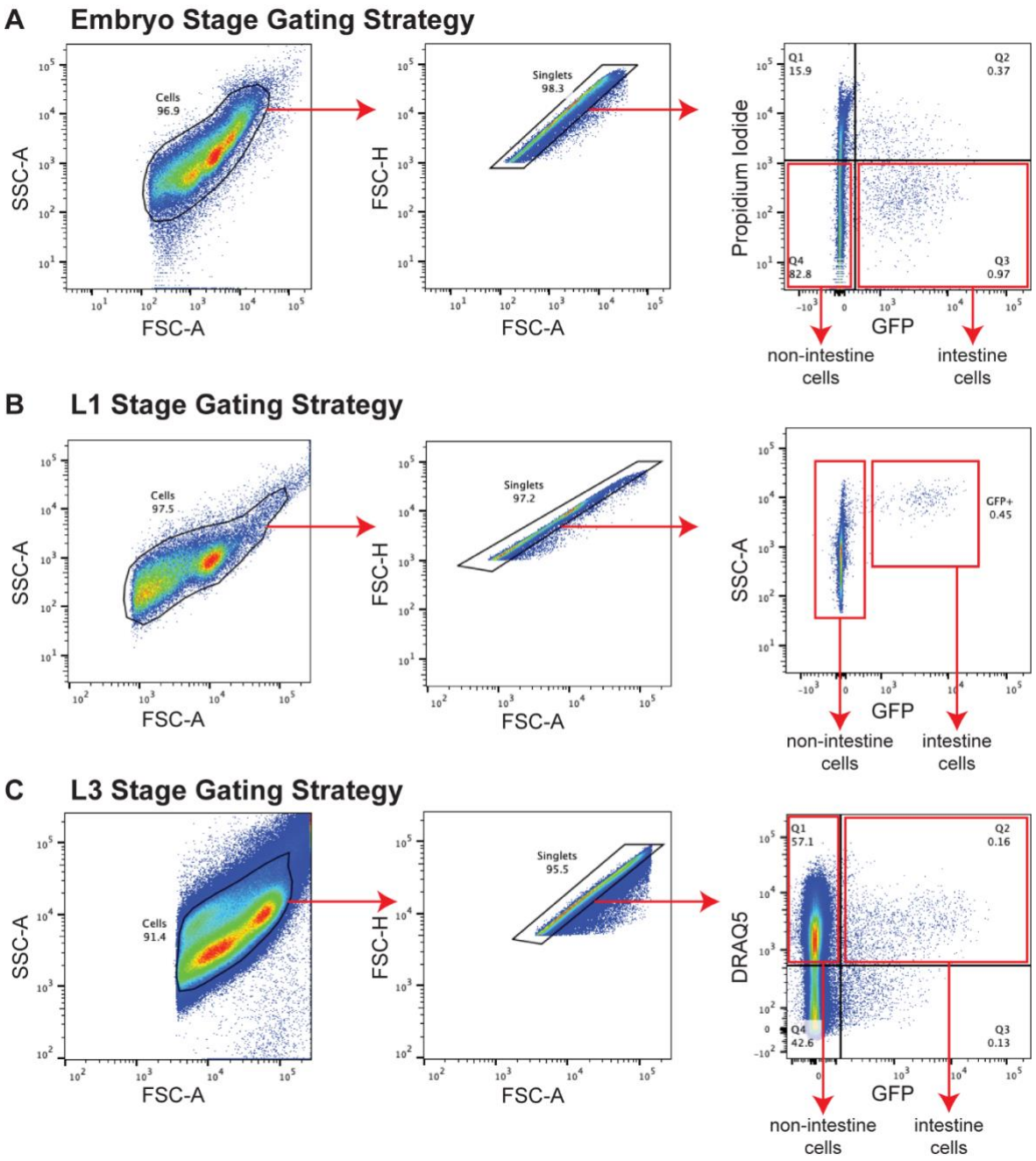

Figure 2.2: FACS gating strategy for isolation of intestine cells. Diagram of gating strategy used for embryo (A), L1 (B), and L3 (C) stage FACS intestine isolation.

of tissue-specific genes in GFP+ intestine cell and GFP- non-intestine cell populations corresponding to intestine (*act-5*, *elt-2*, *ges-1*, *ifb-2*), germline (*glh-1*, *meg-1*, *pgl-1*), neuron (*rab-3*, *rgef-1*, *rgs-1*), muscle (*fhod-1*, *hlh-1*, *myo-3*), and hypodermis tissues (*bli-1*, *lon-3*) (Figure 2.1E). In all assayed developmental stages, only intestine genes scored a significant fold-change score for enrichment in the isolated GFP+ intestine cells. Of the 11 non-intestine genes investigated, only 6 scored as significantly depleted in intestine. This result suggests high confidence in the identification of intestine-enriched genes but an inability to accurately detect genes depleted in the intestine. Thus, further analysis in this study will focus on the set of genes enriched in the intestine.

To determine on a genome-wide scale the similarity and differences between the sequenced libraries, we performed sample-to-sample correlation analysis (Figure 2.3). This analysis identified that samples clustered within each stage and were further subclustered based on intestine vs. non-intestine samples. We also identified that L1 samples from replicate 2clustered within embryo stage samples and were subsequently removed from downstream analysis.

We were concerned that the cell dissociation process may introduce biases to our intestine transcriptome data. To detect dissociation-induced transcriptome bias, we performed differential expression analysis on whole worms and freshly dissociated cells (Figure 2.4A). We found that embryo stage samples had the least difference in transcript abundance compared to post-embryonic stages. We identified 502 and 440 genes significantly increased in dissociated cell populations compared to whole worms in L1 and L3 stages (Figure 2.4B). In contrast, we identified few genes that are significantly reduced in dissociated cell populations compared to whole worms in all developmental stages (Embryo = 1, L1 = 14, L3 = 14). Dissociation-induced genes were subsequently removed from downstream analysis.

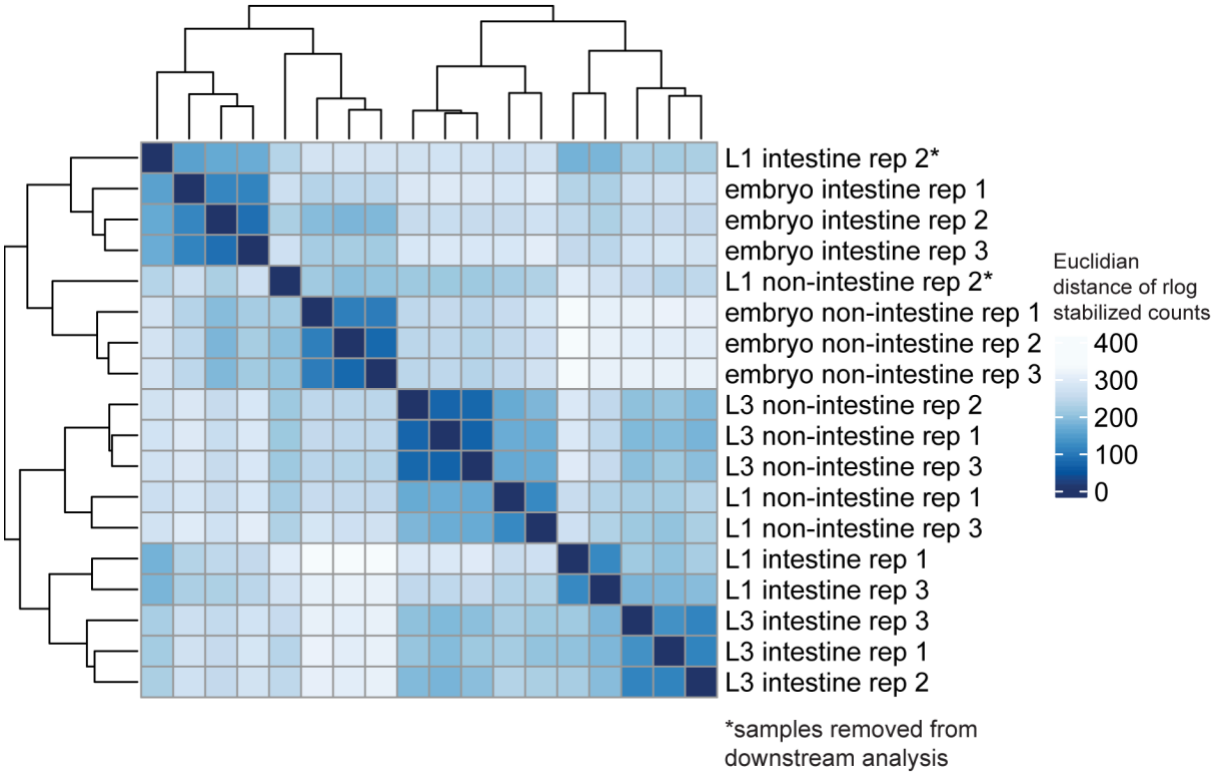

Figure 2.3: Sample-to-sample Euclidean distances of embryo and L1 stage RNA-seq samples. Pairwise measurements for Euclidian distance of regularized logarithm transformed RNA-seq read counts between intestine and non-intestine RNA-seq libraries for assayed developmental stages. L1 stage samples from replicate 2 clustered with embryo stage samples and were removed from downstream RNA-seq analysis.

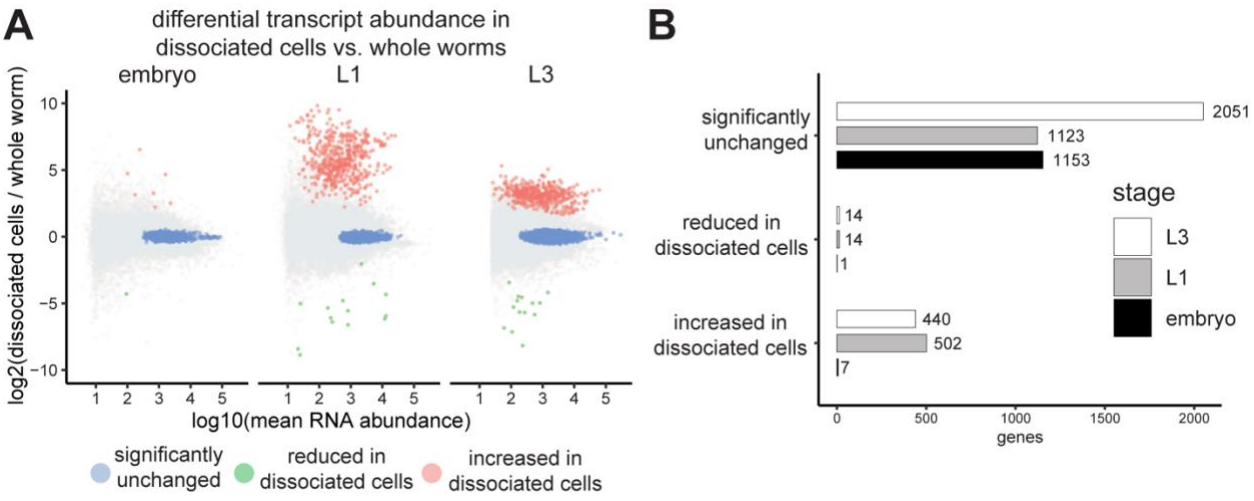

Figure 2.4: Identification of dissociation-induced transcript abundance difference in C. elegans single cell suspension compared to whole worms. (A) MA plots visualizing differential transcript abundance is dissociated cells compared to whole worms. Points are colored based on the category: significantly unchanged (blue), reduced in dissociated cells (green), increased in dissociated cells (red). **(B)** Bar plot displaying the number of transcripts in each category.

2.2.2 Differential RNA-seq analysis detects intestine enriched transcripts

To identify intestine expressed transcripts in each of the embryo, L1, and L3 stages, differential expression analysis was performed between GFP+ intestine and GFP- non-intestine RNA-seq samples (Figure 2.5A-C). We performed Gene Ontology (GO) analysis for intestine enriched transcripts in each developmental stage (Figure 2.5D-E). The associated terms across all stages (*innate immune response*, *peroxisome, membrane raft, carbohydrate binding* and *glucuronosyltransferase activity*) were consistent with known intestinal functions. GO terms associated specifically with the embryo stage included *DNA-binding transcription factor activity* indicating an overrepresentation of transcriptional regulators. Additionally, embryo intestine GO terms included *iron ion binding* and *organic acid metabolic process* suggesting that some aspects of metabolism and homeostasis gain prominence before hatching. GO terms associated specifically with the L1 stage include *proton-transporting V-type ATPase complex*, and terms associated specifically with the L3 stage include *defense response to Gram-negative bacterium,* suggesting the importance of defense at this stage.

To understand how the intestine transcriptome changes over developmental time, we categorized intestine-enriched genes that are unique to one stage or shared between developmental stages. We then performed Gene Ontology (GO) analysis on genes within each category to identify the functional role of intestine enriched genes distinct or shared between stages (Figure 2.5G). We identified 883 genes shared between all three developmental stages, consisting of the largest set intestine enriched genes, and associated with the GO term *response to external stimulus* (p-value < 1E-10). The second largest set of 691 genes were distinct to the embryo stage and were associated with the GO term *small molecule metabolic process* (p-value < 1E-15). Overall, these results demonstrate that the functional role of the intestine is developmentally dynamic particularly between embryonic and post-embryonic stages.

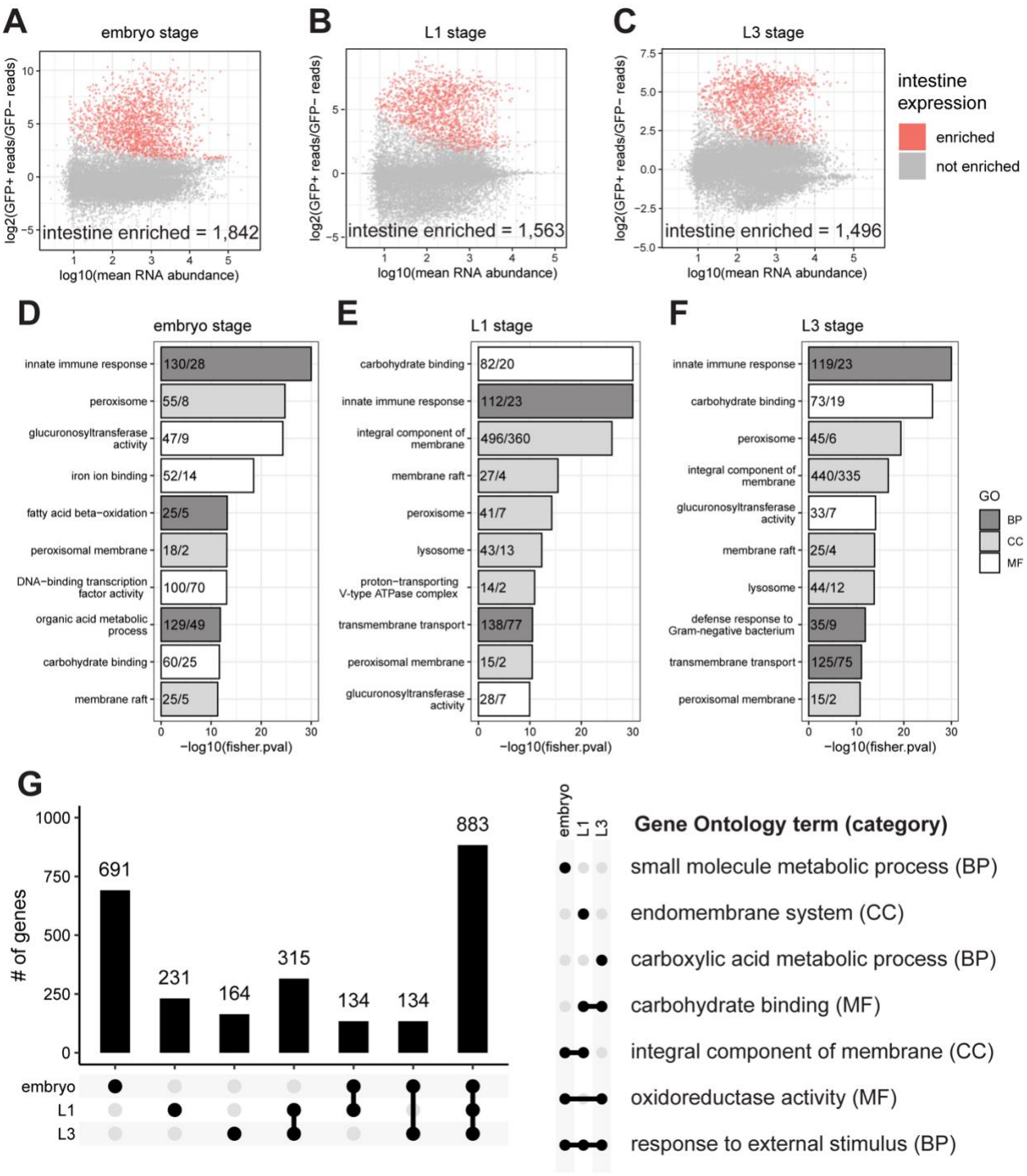

Figure 2.5: Differential expression analysis identifies intestine specific genes in the embryo and L1 stage. (A-C) Scatterplot of pairwise differential expression in FACS isolated GFP+ intestine GFP- non-intestine cells. The log2-fold change of gene expression (y-axis) is plotted against mean normalized read count per gene (x-axis). Significantly enriched intestine genes are highlighted in red. This analysis identified 1,842 embryo enriched genes (A), 1,563 L1

intestine enriched genes and 1,496 L3 intestine enriched genes. **(D-F)** Gene Ontology (GO) analysis of intestine enriched genes identified in each developmental stage. The top 10 GO terms for all three categories are displayed (BP, biological process; CC, cellular component; MF, molecular function). The x-axis corresponds to the -log10 transformed p-value, and the y-axis corresponds to the identified GO term. Numbers within each bar represent the number of "observed vs expected" genes in the input set that correspond to the given GO term. **(G)** (left) Gene set overlap of significantly intestine enriched genes distinct or shared between developmental stages. Dots below the x-axis indicate gene set identification. Single dots are specific to one stage, while dots connected by lines indicate genes shared between stages. (right) the most significant GO term associated with each stage shared or distinct intestine enriched gene set. GO category is indicated in parenthesis (BP, biological process; CC, cellular component; MF, molecular function).

To validate our intestine-enriched RNA-seq dataset, we utilized smiFISH (single molecule inexpensive fluorescence in-situ hybridization) [91,92]. We were interested in determining if our dataset could accurately detect previously uncharacterized intestine transcripts by microscopic visualization. The transcripts for *acy-4*, *bre-2* and *C02D5.4* were detected as intestine expressed across embryo, L1 and L3 stages by RNA-seq (Figure 2.6A). To determine whether these were indeed intestine-enriched transcripts, we assessed their mRNA abundance and localization in embryos while co-staining for the intestine marker transcript *ges-1* [93]. *acy-4* and *bre-2* mRNA were exclusive to intestine tissue (Figure 2.6B, C). In contrast, *C02D5.4* mRNA localized to both intestine and epithelial cells (Figure 2.6D). Together these data serve to support that intestine-specific transcriptome profiling is able to identify new intestine-enriched genes.

2.2.3 Integration of genome-wide datasets

Previous reports have proposed that ELT-2 participates in the transcription of every gene expressed in the differentiating and mature intestine [48,68]. This assertion was based on the high prevalence of ELT-2 binding site sequences (TGATAA) in intestine promoters. However, ChIP-seq data across many fields illustrates that only a fraction of possible binding sites are typically occupied by TFs, usually due to either chromatin context, a requirement for combinatorial binding with other TFs, or other causes [84]. To understand the scope of influence ELT-2 has in the developing intestine GRN, we integrated 1) previously published ELT-2 binding maps in embryo, L1 and L3 stages (ELT-2 ChIP-seq) [72], 2) our intestine-specific RNA-seq profiles (intestine RNA-seq (GFP+/GFP-), and 3) whole transcriptome responses to *elt-2* deletion in the L1 stage (RNA-seq) [47] (Figure 2.7). We plotted all three datasets as aligned heatmaps, filtering for all protein coding genes in the genome. We divided the data into subsets based on ELT-2 occupancy (ELT-2 bound) and expression status within the intestine (intestine-enriched). “ELT-2 bound” genes contained ELT-2 ChIP-seq peak within 1 kb upstream and 200

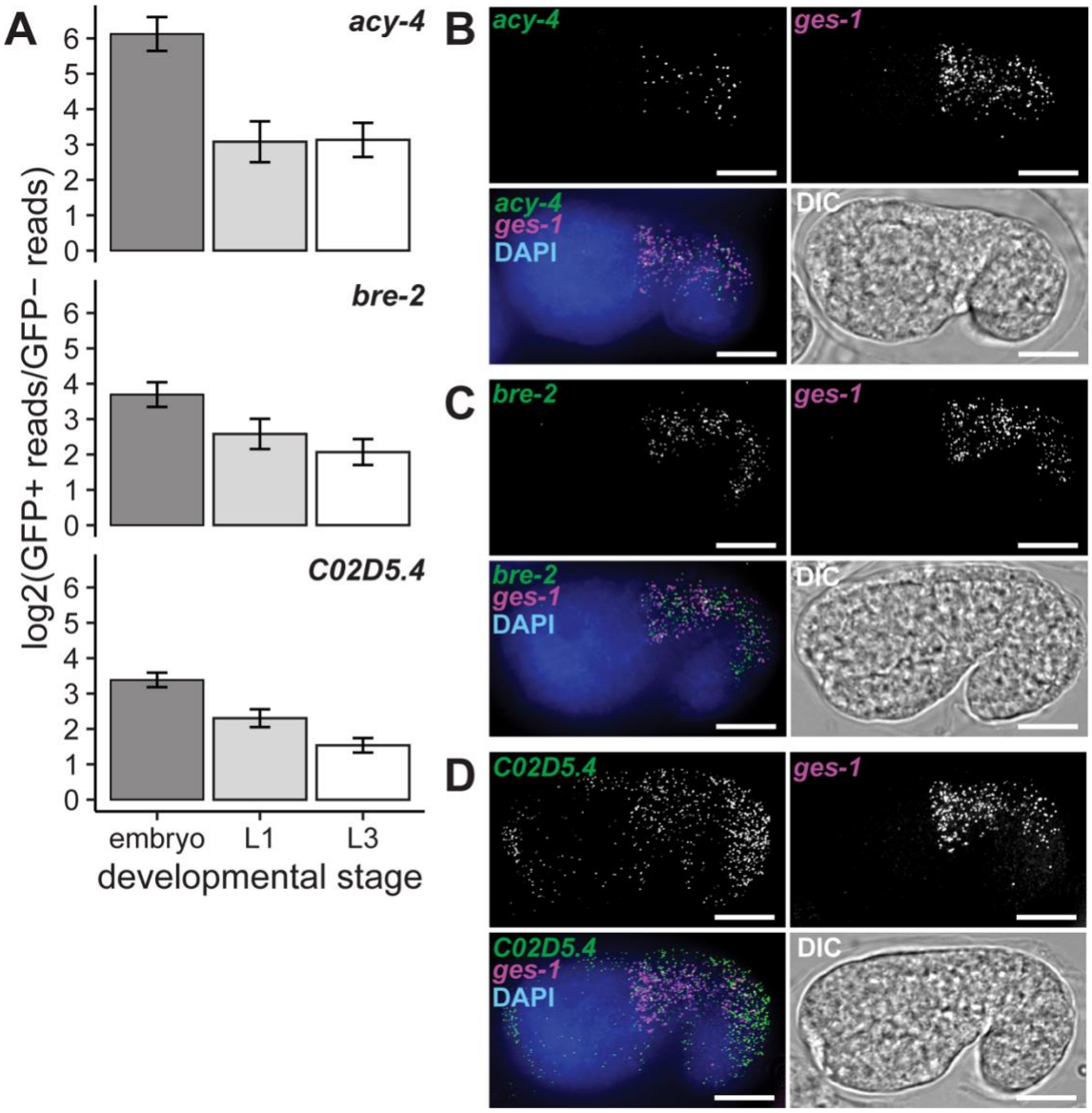

Figure 2.6: Visualization of novel intestine genes detected in intestine FACS dataset. (A) GFP+ (intestine) vs GFP- (non-intestine) log2 transformed RNA-seq read fold change for genes detected enriched in the FACS isolated intestine sample in embryo, L1, and L3 stages (*acy-4*, *bre-2*, *C02D5.4*). All genes had a log2(fold change) > 1 with Benjamini-Hochberg adjusted p-value < 0.01. Bar plots represent the mean and standard deviation of replicate log2(fold change) values (embryo and L3 N = 3, L1 N = 2). **(B-D)** Transcript visualization using smiFISH in representative comma-stage embryos of *acy-4* (A), *bre-2* (B), and *C02D5.4* (C). Transcript of interest (green), intestine marker *ges-1* (magenta), DNA (blue), and DIC images (grey) are shown. Colocalization of *ges-1* intestine marker with transcripts of interest were observed. Representative images from three biological replicates collecting at total of at least 30 images. Scale bars = 10 microns.

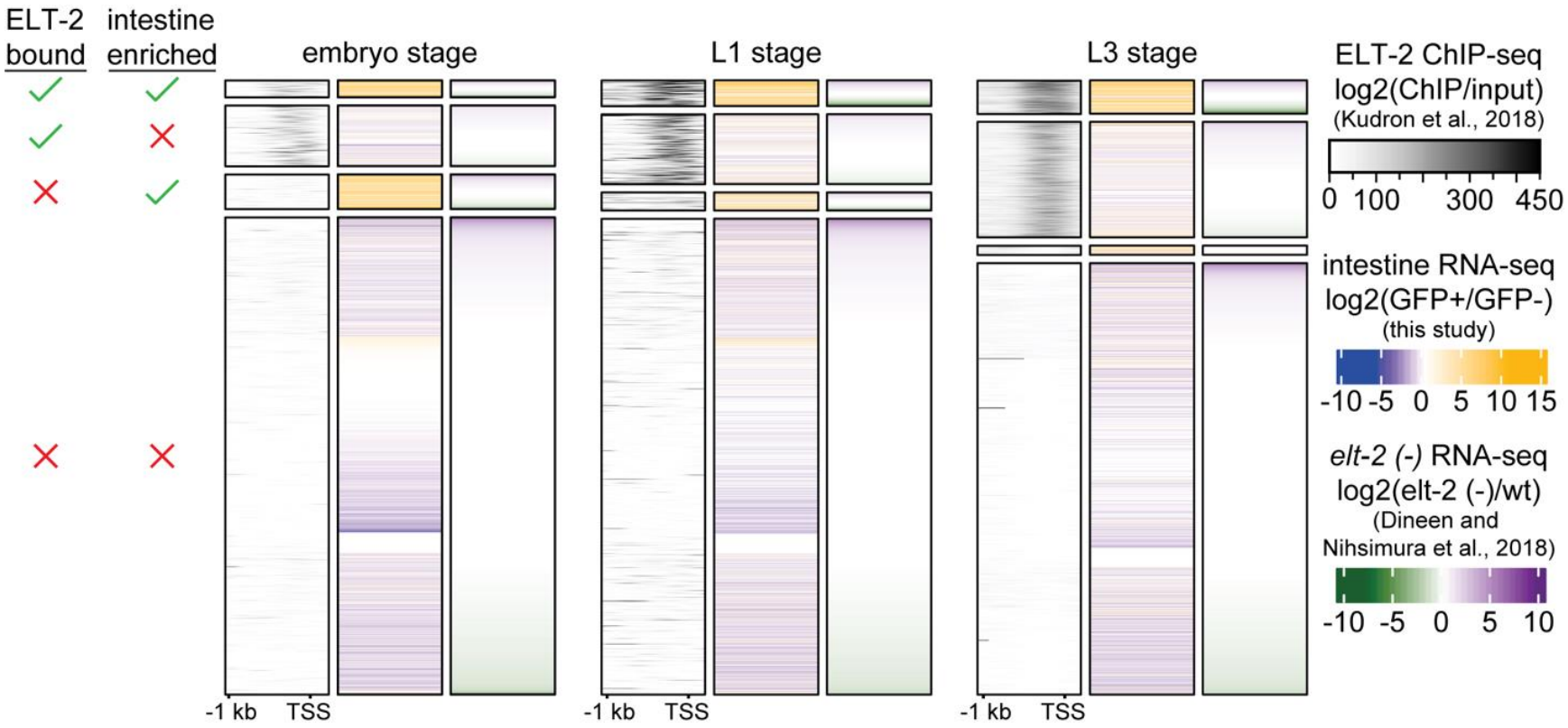

Figure 2.7: Integration of ELT-2 and intestine genome-wide datasets. Heatmap visualization ELT-2 binding (left, black), log2 transformed intestine enrichment (middle, blue to yellow), and log2 transformed *elt-2 (-)* transcription response (right, green to purple). ELT-2 binding (ChIP-seq [72]) and intestine enrichment (RNA-seq, this study) was measured in embryo, L1 and L3 stages. *elt-2 (-)* transcription (RNA-seq [47]) response was measured in the L1 stage. Genes are separated based on presence or absence of promoter localized ELT-2 binding, and significant intestine enrichment. Genes within each category are ranked from high to low based on *elt-2 (-)* transcriptional response.

bp downstream of their transcription start site. Genes were “intestine enriched” if they scored an intestine enrichment score greater than 1 and p-value < 0.01 (see Methods). In addition, we plotted the fold-change ratio of transcriptional response to *elt-2* loss and sorted the order of genes within the heatmap based on this response. Genes dependent on ELT-2 for their expression (ELT-2 activated) are shown in green and genes that become over abundant with ELT-2 loss (ELT-2 repressed) shown in purple.

There are important caveats to note regarding the use of this data relating to the fact that the transcriptome dataset measuring response to *elt-2* loss was collected by whole animal flow cytometry sorting of L1 worms that had either lost or retained an unstable ELT-2::GFP rescue array [47]. Because of this, three main caveats are: 1) the data was collected only in the L1 stage and is therefore not truly representative of genes regulated in embryo or L3 stages, 2) the data was collected in whole worms thereby capturing expression events occurring outside of the intestine, and 3) the data represents both direct and indirect responses to *elt-2* loss. However, this dataset still provides useful information regarding the role of ELT-2 in the intestine GRN.

2.2.4 Quantification of ELT-2 target genes in the intestine GRN

Previous work has hypothesized that ELT-2 serves to regulate the expression of intestine genes. We set to evaluate this hypothesis with the available genome-wide resources. If ELT-2 occupancy is a major predictor of intestine enriched expression, we would expect to see a large category of genes in which “ELT-2 bound” and “intestine enriched” co-occur. This would suggest that ELT-2 occupancy is predictive of intestine enrichment. In contrast, if ELT-2 occupancy is not a predictor of intestine enriched expression, we would expect a large fraction of “intestine enriched” genes to have “ELT-2 absent” promoters.

To evaluate if ELT-2 binding is associated with expression of all intestine genes, we measured the percent of intestine enriched genes that are ELT-2 bound (Figure 2.8A). In the embryo stage only 31% (572/1745 genes) of intestine enriched genes are bound by ELT-2.

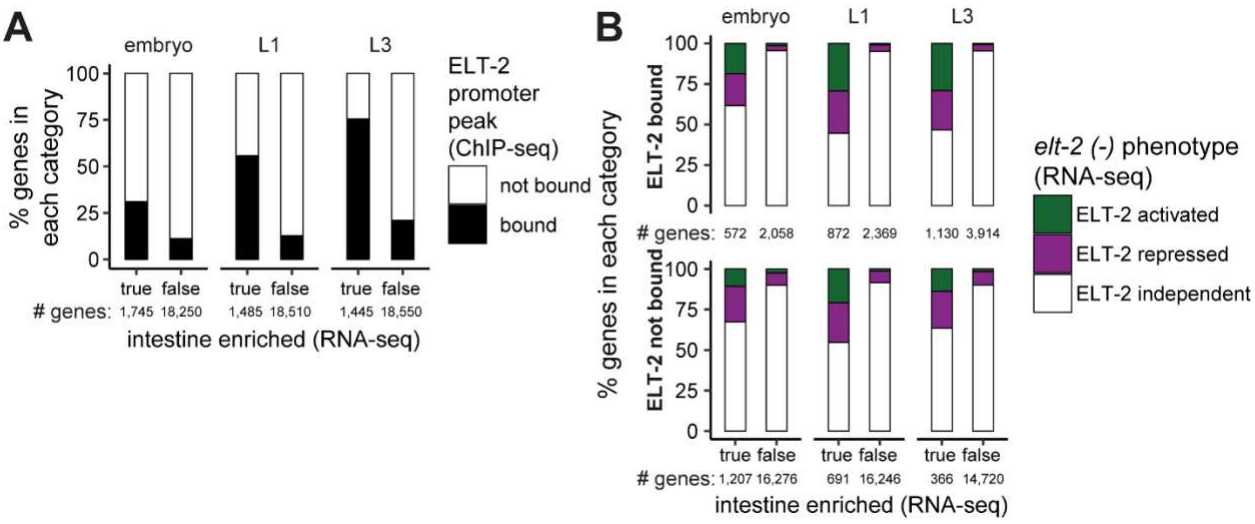

Figure 2.8: ELT-2 regulated genes are both activated and repressed. (A) Percent of intestine enriched (true) or not intestine enriched (false) genes with a significant ELT-2 ChIP-seq peak in a gene promoter (black, ELT-2 promoter peak present; white, ELT-2 promoter peak absent). The total number of genes in each set are indicated below each bar. **(B)** Percent of intestine enriched (true) or not intestine enriched (false) genes with transcriptional response to *elt-2 (-)* (green, ELT-2 activated; purple, ELT-2 repressed; white, ELT-2 independent). Percentages were measured for direct targets with ELT-2 promoter peak (top, ELT-2 bound) and indirect targets without an ELT-2 promoter peak (bottom, ELT-2 not bound). The total number of genes in each set are indicated below each bar.

However, this percentage increased in L1 (55%, 872/1485 genes) and L3 (75%, 1130/1445 genes) stages. In all stages, ELT-2 bound genes were more likely to be intestine enriched than genes not bound by ELT-2 (chi-square p-value < 1E-100). These results suggest that while ELT-2 binding is associated with intestine enrichment, not all intestine enriched genes are bound by ELT-2, particularly in the embryo stage.

2.2.5 ELT-2 regulates target genes through activation and repression

Previous reports have suggested that ELT-2 functions exclusively as an activator of the intestine gene regulatory network [38,68,83]. To evaluate if ELT-2 serves to primarily activate target gene transcription we integrated RNA-seq data measuring differentially expressed genes in *elt-2 (-)* versus wildtype conditions (Figure 2.8B). Approximately half of intestine enriched genes directly bound by ELT-2 had no requirement for ELT-2 for their transcriptional activity (embryo 61.6%, L1 44%, L3 46%). We observed approximately a fifth to a third of intestine enriched genes bound by ELT-2 are positively regulated by ELT-2 (embryo 18.6%, L1 29.3%, L3 29.0%). The remaining genes were negatively regulated by ELT-2 (embryo 19.6%, L1 26.0%, L3 24%). We also observed similar proportions to genes either activated or repressed by ELT-2 that were not bound by ELT-2. This suggests that ELT-2 regulates genes through both activation and repression in a direct and indirect fashion.

Indirect ELT-2 regulated genes measured lower percentage of activated genes and a higher percentage of independent genes. In all cases, we identified a statistically significant relationship between intestine enrichment and the distribution of ELT-2 regulated gene states (chi-square p-value < 1E-35). These results suggest that the bulk of direct ELT-2 targets are not fully dependent on ELT-2 for their transcriptional activity. It is therefore likely that other factors or contexts may be required for ELT-2 transcriptional regulation. Additionally, these results demonstrate that ELT-2 functions not only as an activator but also repress distinct sets of target genes.

2.2.6 Intestine-enriched gene expression as a function of ELT-2 ChIP signal

A simple mechanistic model describing the relationship between transcription factor regulation and target gene expression asserts that transcription factor binding quantitatively influences transcriptional output. If this were the case, genes with high ELT-2 occupancy would be highly intestine expressed while genes with low ELT-2 occupancy would have lower expression in the intestine. In contrast, ELT-2 occupancy may be bimodal (on/off) and not affect the overall level of intestine expression for target genes.

To distinguish between these scenarios, we plotted intestine RNA-seq counts (regularized and log-transformed) for ELT-2 bound versus genes not bound by ELT-2 (defined as -1kb and +0.2kb centered on the TSS) (Figure 2.9). We found that for all developmental stages, ELT-2 bound genes had significantly higher intestine RNA-seq counts than genes without ELT-2 binding. However, the linear correlation between intestine RNA-seq counts against ELT-2 ChIP-seq ranged from R^2 values of 0.04 – 0.14, indicating low predictive value and suggesting against a scenario where the degree of ELT-2 occupancy informs gene expression level (Figure 2.10A). Stratifying based on ELT-2 regulatory status (Figure 2.10B) failed to improve the correlation coefficients, which ranged from R^2 values from 0.03 – 0.16, again indicating a low predictive value.

These results suggest that although ELT-2 target genes show an overall increase in intestine expression compared to those not targeted by ELT-2, the effect size was small. Furthermore, our findings do not support a model that ELT-2 occupancy levels dictate target gene transcriptional output.

We were interested in determining whether ELT-2 occupancy levels were higher in genes whose expression was enriched in the intestine, suggesting that the degree of ELT-2 binding could influence transcriptional output. To evaluate this, we plotted ELT-2 ChIP-seq signal averaged across promoter regions for different expression categories (Figure 2.11). ELT-2 occupancy was greater for intestine enriched genes than non-intestine genes in L1 and L3

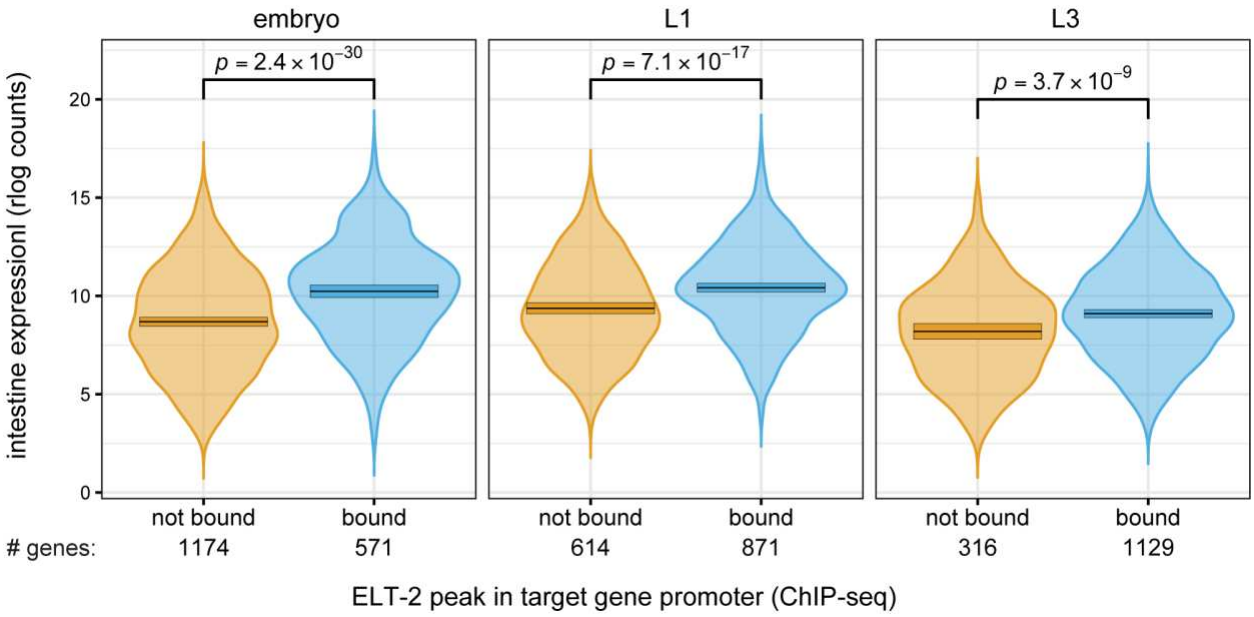

Figure 2.9: ELT-2 binding status is associated with higher mean expression of intestine-enriched genes in every stage. Violin plot of genes intestine expression for genes categorized based on presence or absence of ELT-2 peak in target gene promoter. The number of genes in each set are indicated below the each histogram. Student's t-test was used to measure statistical significance. Bars within the violin plot represent the 99% confidence interval.

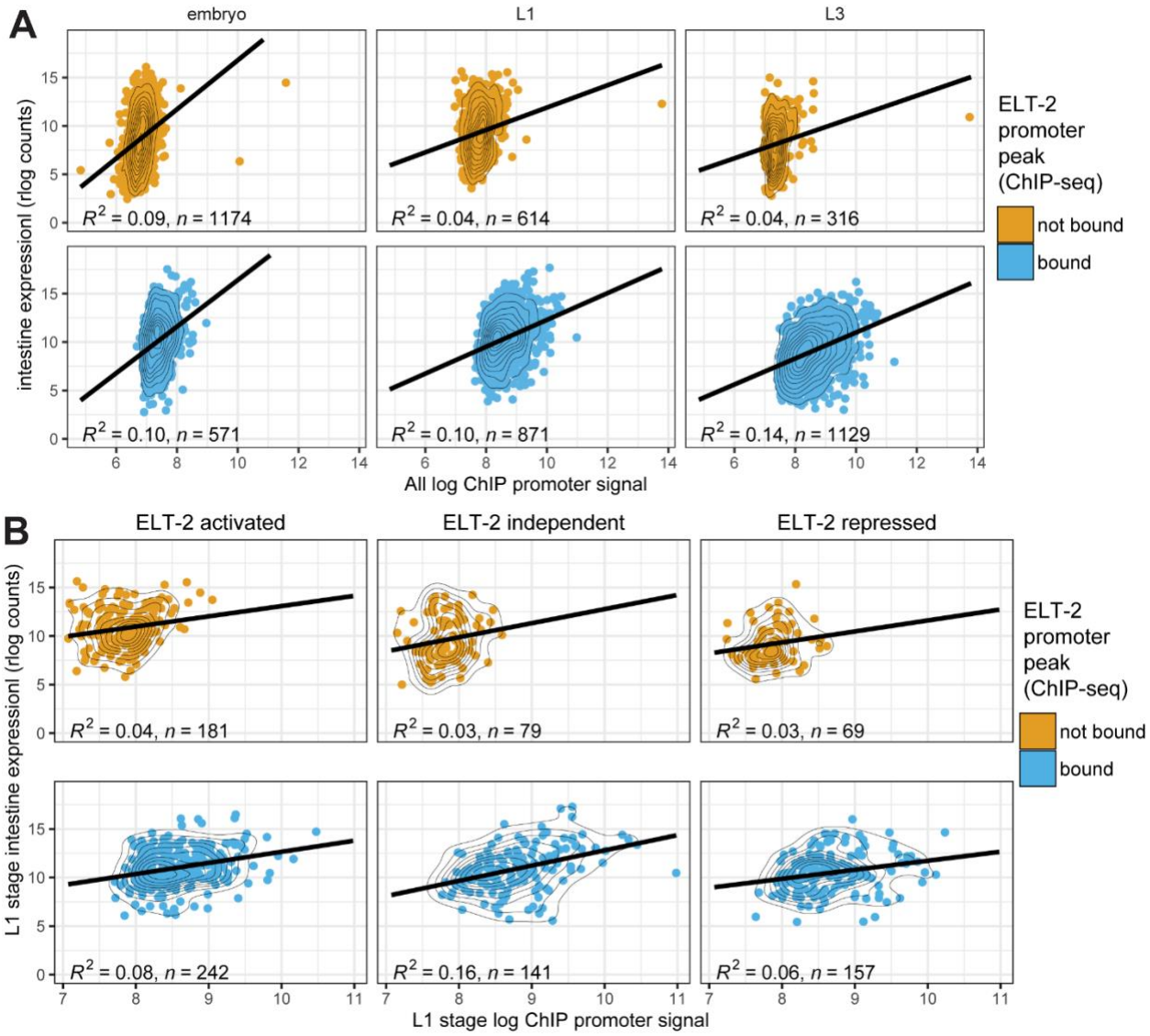

Figure 2.10: Intestine expression is not correlated to ELT-2 occupancy of target gene promoters. (A) As a quantitative predictor, the relationship between ELT-2 occupancy signal and gene expression is not higher for genes associated with a promoter that is reproducibly bound by ELT-2 (slopes are the same or lower in magnitude between the two binding statuses, for each stage independently). The greater R^2 values for ELT-2 bound data indicate that those data fit their respective models better, but still reflect a relatively low variance explained (.10 - .14). **(B)** Slopes did not improve for when genes were further subset based on dependence for ELT-2 transcriptional regulation. Thin black lines represent contour lines for point density.

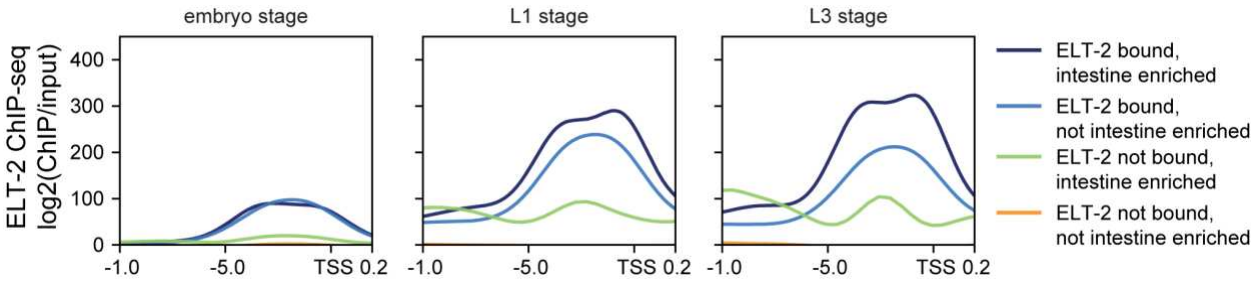

Figure 2.11: Visualization of ELT-2 ChIP-seq read density averaged across promoter regions centered on the gene transcription start site. Promoters were defined as -1kb upstream and 0.2kb downstream of a gene transcription start site (dark blue, ELT-2 bound and intestine enriched; light blue, ELT-2 bound and not intestine enriched; ELT-2 not bound, intestine enriched; ELT-2 not bound, not intestine enriched).

stages but not in embryos. This corroborates our previous findings that the relative contribution of ELT-2 to the intestinal transcriptome increases over developmental time (Figure 2.5). Other transcription factors are likely participating in shaping the embryonic transcriptome in ways we have not yet characterized.

2.2.7 ELT-2 represses defense response genes

Since ELT-2 is intestine expressed throughout development, we were interested to determine if the function of ELT-2 changed over development. That is, we wondered whether repression and activation by ELT-2 was specific to one or more stages. We observed 14 ELT-2 repressed genes and 2 ELT-2 activated genes unique to the embryonic stage (Figure 2.12A). Conversely, we observed 50 ELT-2 repressed genes and 63 ELT-2 repressed genes shared between all assayed stages. This suggests that there is not a heavy bias in developmental stage for either ELT-2 repression or activation activity in the intestine GRN. Furthermore, we identified 69 ELT-2 repressed genes and 89 ELT-2 activated genes shared between L1 and L3 stages. Together, these results suggest that ELT-2 has a developmentally dynamic role in the intestine GRN, but the unknown factors that dictate ELT-2 regulatory dynamics are not stage specific.

To better understand the type of genes ELT-2 is regulating, we characterized the tissues that ELT-2 dependent target genes are associated with. Previous reports of human homolog GATA factors have demonstrated that they serve to promote one fate while repressing an alternative fate [94]. We were interested in determining if ELT-2 functions this way. Alternatively, since the *elt-2 (-)* differential expression data was collected on whole worms, we may be observing transcriptional responses that may not correspond to gene regulatory events within the intestine.

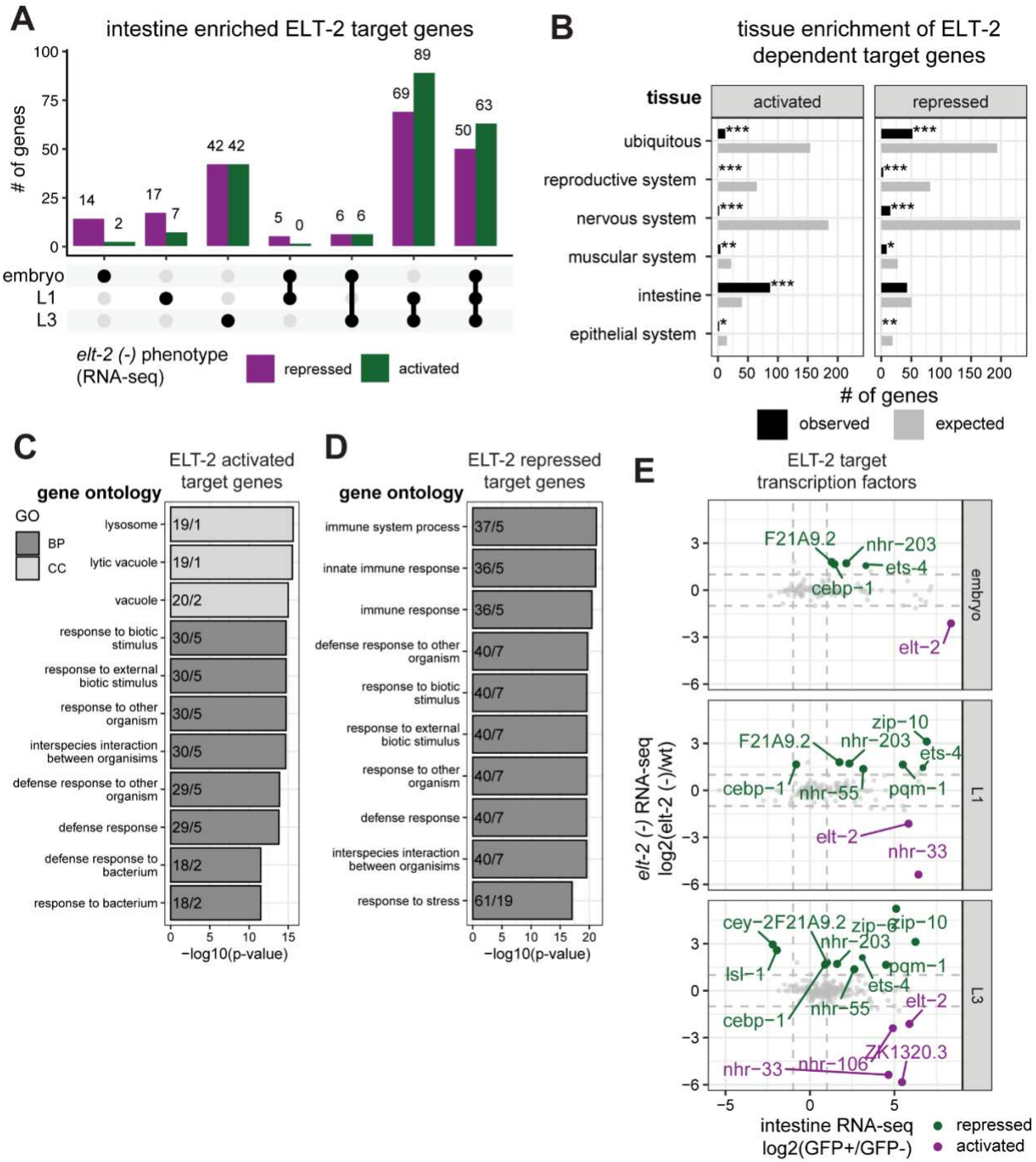

Figure 2.12: ELT-2 repressed target genes are defense response genes. (A) Gene set overlap of ELT-2 dependent intestine enriched genes distinct or shared between developmental stages. Dots below the x-axis indicate gene set identification. Single dots are specific to one stage, while dots connected by lines indicate genes shared between stages. **(B)** Enrichment of ELT-2 activated and repressed direct target genes from all developmental stages for genes associated with ubiquitous or organ system specific expression. The hypergeometric statistical test was used to test enrichment and depletion of gene set numbers for each category (black,

observed number of genes; grey, expected number of genes; *P-value < 0.01, **P-value < 1x10−5, ***P-value < 1x10−10). **(C-D)** Gene Ontology (GO) analysis of direct ELT-2 activated (C) or repressed (D) genes identified in all developmental stages. The top 10 GO terms for all three categories are displayed (BP, biological process; CC, cellular component). The x-axis corresponds to the -log10 transformed p-value, and the y-axis corresponds to the identified GO term. Numbers within each bar represent the number of “observed vs expected” genes in the input set that correspond to the given GO term. **(E)** Scatterplot visualizing intestine enrichment and *elt-2(-)* transcription response for ELT-2 direct target transcription factors (ChIP-seq). Dotted lines correspond the log2 fold change threshold of 1 used for differential expression significance testing. Genes with significant *elt-2 (-)* transcriptional response are indicated with genes names (purple, repressed; green, activated, Benjamini-Hochberg adjusted p-value < 0.01). Grey dots indicate transcription factors without significant response to *elt-2 (-)* (Benjamini-Hochberg adjusted p-value > 0.01).

To evaluate these possibilities, we generated non-overlapping sets of anatomy ontology terms for major organ systems including *reproductive system, nervous system, muscular system, intestine,* and *epithelial system*, and a separate set of ubiquitously expressed genes [95]. We used hypergeometric statistics to evaluate the enrichment or depletion of terms for all ELT-2 target genes from all stages stratified based on ELT-2 activation or repression (Figure 2.12B).

We observed that ELT-2 activated target genes were significantly enriched in genes associated with *intestine* and significantly depleted in genes associated with both non-intestine and ubiquitous genes. We identified that ELT-2 repressed target genes were similarly depleted in non-intestine and ubiquitous genes but were not significantly enriched in *intestine* annotated genes. This result suggests that ELT-2 activated genes are strongly associated with intestine genes, but that ELT-2 repressed genes are not associated any one tissue. It is possible that ELT-2 repressed genes are intestine expressed but may not be reliably detected in normal growth conditions and may represent genes that are repressed by ELT-2 for inducible expression.

We were curious whether the genes that ELT-2 represses shared similar functions suggesting that repression by ELT-2 could prevent specific biological characteristic. To evaluate the biological processes associated with ELT-2 activation versus repression, we performed GO analysis on ELT-2 activated or repressed direct target genes in all three developmental stages (Figure 2.12C, D). ELT-2 activated target genes were associated with *defense response to bacterium* and related *lysosome* terms. ELT-2 repressed target genes were associated with *response to stress* and related *innate immune response* terms. Both ELT-2 activated and repressed target genes were associated with several related *defense response* terms and *interspecies interaction between organisms*. Overall, these results suggest that both ELT-2 activation and repression activity is central in regulating responses to biotic stimulus. Our results suggest that ELT-2 turns on lysosome and a set of immune-related genes, whereas ELT-2 repression is central to regulating innate immunity genes. These findings are consistent with

previous reports which have demonstrated ELT-2 regulation of lysosome and immunity pathways [96,97].

Our results suggest that ELT-2 is not alone in driving intestine enriched gene expression in embryonic stages but that other TFs contribute independently or downstream of ELT-2 action. To determine which other TFs are downstream targets of ELT-2, we investigated ELT-2 dependent target TFs to expand the ELT-2 gene regulatory network. We subset ELT-2 regulated target genes for known and predicted TFs (Figure 2.12E) [98]. We identified ELT-2 activated target transcription factors *nhr-33*, *nhr-106*, and *ZK1320.3*, and ELT-2 repressed target transcription factors *cebp-1*, *ets-4*, *nhr-203*, *F21A9.2*, *nhr-55*, *zip-10*, *pqm-1*, *cey-1*, and *lsl-1*. *elt-2* was detected as an activated target due to the nature of the *elt-2* deletion allele used in this study. We observed that the number of ELT-2 regulated target transcription factors increases over developmental time. ELT-2 regulation of these transcription factors may explain the observed ELT-2 transcriptional dependence for genes without evidence of ELT-2 promoter binding (Figure 2.8B). These results further emphasize that ELT-2 regulates the intestine gene regulatory network though both activation and repression of target genes.

2.2.8 ELT-2 negatively regulates expression of transcription factors CEBP-1 and ETS-4 in the intestine

Our analysis suggests that ELT-2 regulates the intestine gene regulatory network through activation and repression. Previous work has focused on ELT-2 as a transcriptional activator but has not demonstrated ELT-2 acts as a direct repressor [47]. ELT-2 may either downregulate gene expression through direct repression or through indirect regulatory feedback loops. Additionally, it is possible that some genes that become over-expressed upon ELT-2 depletion are the result of over-expression in other tissues. To better assess ELT-2 negative regulatory behavior, we selected three genes that are the direct targets of ELT-2 binding that were up-regulated upon *elt-2* deletion: *cebp-1* (CCAAT/enhancer-binding protein), *ets-4* (ETS

class transcription factor) and *pqm-1* (ParaQuat Methylviologen responsive) (Figure 2.13A-C). We assessed whether GFP translation fusion constructs for these TFs recapitulated the over-expression phenotype observed upon *elt-2* deletion in RNA-seq data. We simultaneously evaluated whether over-expression was intestine-specific. We depleted ELT-2 by RNA interference (RNAi), and ELT-2 knockdown was validated by observing a significant 2-fold reduction in ELT-2::GFP signal in intestinal nuclei compared to control worms at matched developmental stages (p-value = 1.2E-18) (Figure 2.14A, E).

cebp-1 encodes a bZIP (basic leucine zipper) class transcription factor with annotated expression in intestine, hypodermis, pharynx, muscle, and neuronal cells. *cebp-1* function has been demonstrated as essential for neuronal axon regeneration, stress response, and intestinal immune response [99–102]. In native ELT-2 conditions, CEBP-1::GFP fluorescence was observed in pharangeal nuclei. Depletion of ELT-2 led to a significant 1.4-fold increase of CEBP-1::GFP reporter activity specifically in intestine nuclei (p-value = 3.3E-6) (Figure 2.14B, E).

ets-4 encodes an ETS (E26 transformation-specific) class transcription factor with a winged helix-turn-helix DNA binding motif [103]. *ets-4* has annotated expression in the intestine, germline, and adult hypodermis and seam cells and functions in lifespan regulation and axon regeneration [104–108]. In native ELT-2 conditions, we observed ETS-4::GFP reporter activity localized to the nuclei of the pharyngeal-intestinal valve (vpi) cells, anterior intestine nuclei, and rectal gland cells (Figure 2.14C). Upon ELT-2 depleiton by RNAi, we observed a significant 1.7-fold increase of ETS-4::GFP reporter activity in intestine nuclei (p-value = 1.3E-8) (Figure 2.14C, E).

pqm-1 is a C2H2-type zinc finger transcription factor that is primarily expressed in the intestine and neuronal cells [98,109–111]. *pqm-1* functions in several processes including stress response, defense response, lipid metabolism, and lifespan regulation [112–116]. We observed PQM-1::GFP reporter signal localized to intestine nuclei (Figure 2.14D). In contrast to ETS-4::GFP and CEBP-1::GFP, ELT-2 depletion had an opposite effect on PQM-1::GFP reporter

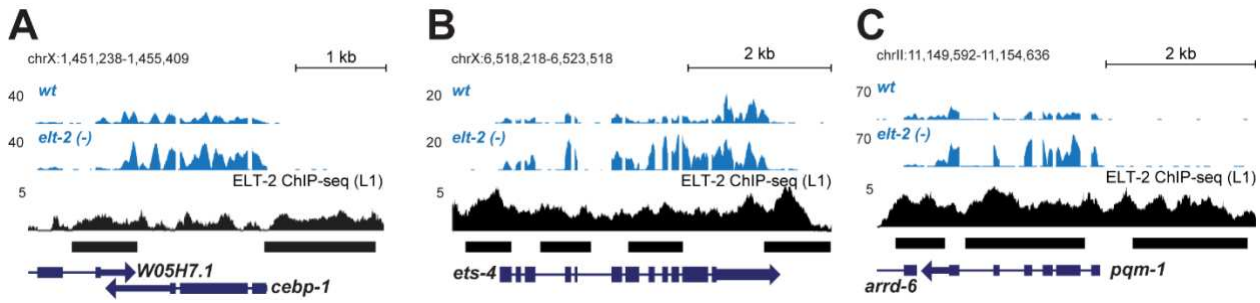

Figure 2.13: Visualization of repressed ELT-2 target gene locus. (A-C) Genome browser tracks of the *cebp-1* (A), *ets-4* (B), and *pqm-1* (C) genomic locus. RNA-seq tracks (blue) display increased transcript abundance in *elt-2(-)* genetic background compared to wildtype (WT, N2). ELT-2 ChIP-seq tracks (black) display ELT-2 binding sites in the gene promoters (black bars).

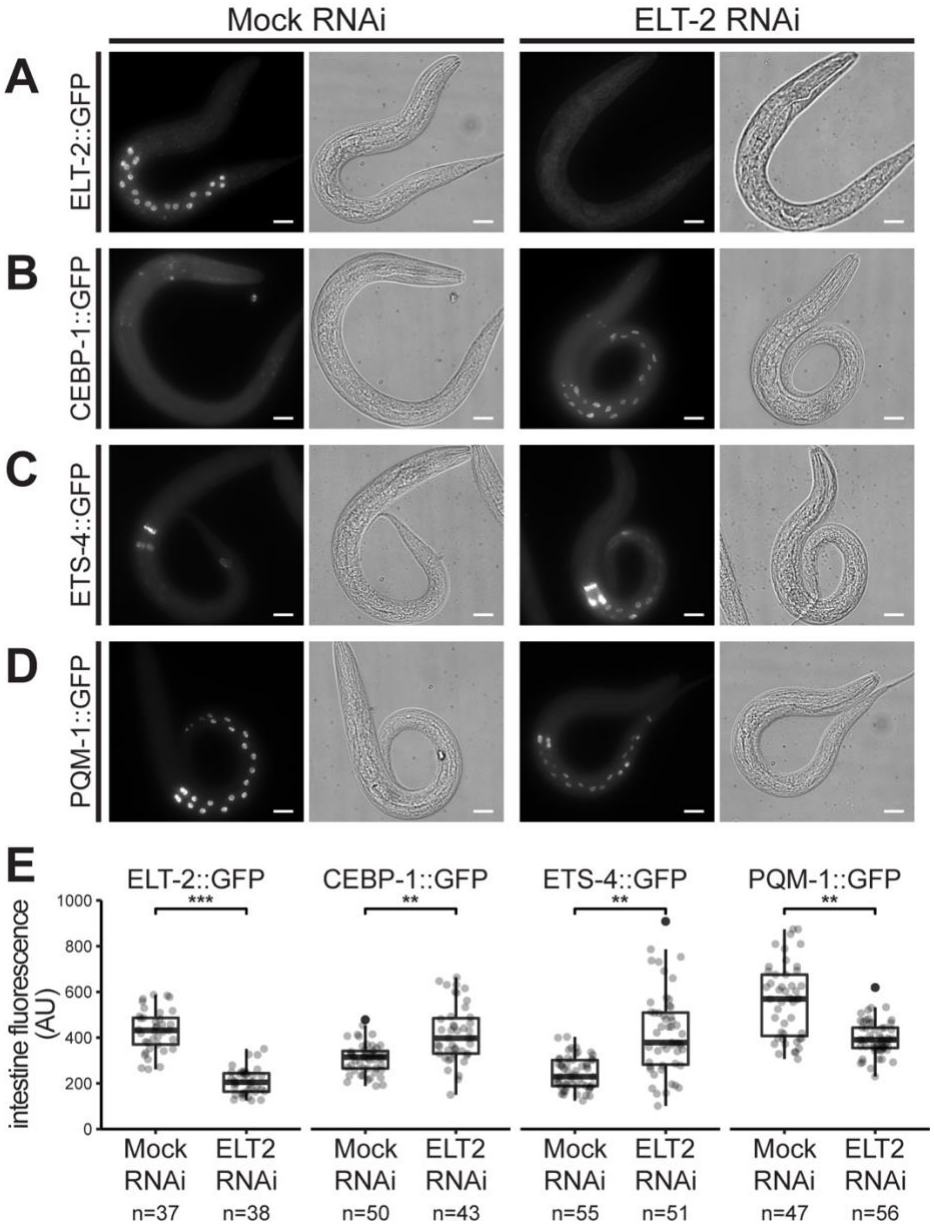

Figure 2.14: ELT-2 negatively regulates expression of transcription factor CEBP-1 and ETS-4 in the intestine. (A-D) Representative images of GFP translation reporter signal (gray) in immobilized L1 stage animals during mock RNAi (L4440, left) or ELT-2 RNAi (right) treatments. The strains image include: (A) ELT-2::GFP (*wgIs56 [elt-2::GFP]* allele), (B) CEBP-1::GFP (*wgIs563 [cebp-1::EGFP]* allele), (C) ETS-4::GFP (*wgIs509 [ets-4::EGFP]* allele), (D) PQM-1::GFP (*wgIs201 [pqm-1::EGFP]* allele). To remove birefringent gut granule autofluorescence, animals were fixed and permeabilized through liquid nitrogen freeze-crack and subsequent washes with methanol, acetone, and PBST. Differential interference contrast microscopy (DIC) is also shown. Worms were imaged at 60x and multi-panel images were stitched as necessary. Scale bars = 10 microns. **(E)** Quantification of GFP signal in *C. elegans* intestines in GFP translation reporter strains (A-D). GFP measurements were collected from background subtracted maximum Z projections from three biological replicates. Intensity measurements were normalized for intestine area. Box and whisker plots represent the

distribution of data points. Each point represents data measured from a single worm; filled points represent outliers (black). Total number of worms imaged are listed below the column label. Student's t-test was calculated to measure statistical significance in GFP intestine fluorescence between mock and ELT-2 RNAi treatments (*P-value < 0.01, **P-value < 1x10−5, ***P-value < 1x10−10).

activity than expected. Intestine localized PQM-1::GFP reporter activity was significantly reduced 1.5-fold when ELT-2 was depleted by RNAi (p-value = 1.4E-8) (Figure 2.14D, E).

The increase in intestine localized GFP for CEBP-1::GFP and ETS-4::GFP upon ELT-2 depletion is consistent with the model that ELT-2 functions to directly repress these direct target genes. We selected these transcripts for hypothesis validation because we detected an ELT-2 peak in their promoters (Figure 2.14D, Figure 2.13A, B). Our reporter assay results illustrate that both CEBP-1 and ETS-4 depend on direct ELT-2 regulation to remain lowly expressed in the intestine. It is still formally possible that in addition to ELT-2's direct action at these promoters that indirect negative feedback loops are also involved. In contrast, the evidence that PQM-1::GFP protein abundance decreases upon ELT-2 depletion is incongrous with the evidence that *pqm-1* transcript abundance increases upon ELT-2 depletion (Figure 2.12E, Figure 2.14C). This result does not support our initial hypothesis for the regulatory connection between ELT-2 and PQM-1 but does demonstrate that PQM-1 is dependent on ELT-2 genetic control. This suggests that PQM-1 is not controlled by a simple ELT-2 transcriptional repression model and that PQM-1 may be under more complex genetic control than anticipated possibly with transcriptional and translational regulatory components. Additionally, this result emphasizes previous findings that mRNA transcript abundance is often insufficient to predict protein abundance [117,118].

Together, our finding that *cebp-1* and *ets-4* are upregualted in the absence of *elt-2* support our hypothesis that *elt-2* plays a repressive role in their transcriptional regulation. The evidence of promoter localized ELT-2 binding and changes to transcriptional abundance of target genes suggests that ELT-2 participates in the repression of these target genes, but the molecular mechanisms regulating this process requires futher experimentation.

2.2.9 ELT-2 negatively regulates its own promoter

ELT-2 regulates its own expression, and the presence of ELT-2 protein at its own gene promoter has been extensively characterized [38,76]. The reported purpose of this regulation is to sustain ELT-2 expression throughout the worm's lifetime through autoactivation until it declines in old age [49]. Previous transcriptomics studies were unable to discern how *elt-2* promoter output responded to *elt-2* depletion, as the studies were conducted in *elt-2* deletion backgrounds [47]. Having identified ELT-2 repressive activity, we thought it pertinent to reevaluate this hypothesis. By using an *elt-2* promoter reporter transgene driving histone H2B fused to GFP, we could study *elt-2* promoter activity divorced from production of ELT-2 protein (Figure 2.15A). If ELT-2 positively regulates its own promoter's expression, we expect to observe reduced *elt-2* promoter activity upon ELT-2 depletion. Suprisingly, depletion of ELT-2 protein in the L1 stage resulted in a ~5-fold over-expression of GFP, suggesting that ELT-2 *negatively* regulates *elt-2* promoter activity (Figure 2.15B-D). We observed this negative autoregulation in the L1 and L3 and adult stages suggesting that this activity is not restricted to one developmental stage (Figure 2.16).

The observation of upregulation in the absence of *elt-2* resembled the previously identified "*elt-7* over-compensated" gene set [47]. These genes were demonstrated to have increased transcriptional output in the absence of *elt-2* that was abolished when *elt-7* was also removed. This work was unable to determine how ELT-2 responds to *elt-2* deletion as the *elt-2 (ca15)* allele previously used to perform transcriptomics was a complete coding sequence deletion. Use of this allele obscured any transcriptional response of *elt-2* to either *elt-2* or *elt-7* mutations. To determine if *elt-2* is also an "*elt-7* over-compensated" gene, we measured *elt-2* promoter activity in an *elt-7* genetic deletion background. We observed that the *elt-2* promoter is no longer over expressed when both ELT-2 and ELT-7 are depleted (Figure 2.15B-D), thereby demonstrating that ELT-2 is also "*elt-7* over-compensated". We were interested to determine if our results demonstrating ELT-2 negative autoregulation was specific to the reporter construct or if this phenomena could be observed outside of a reporter gene context. We visualized the

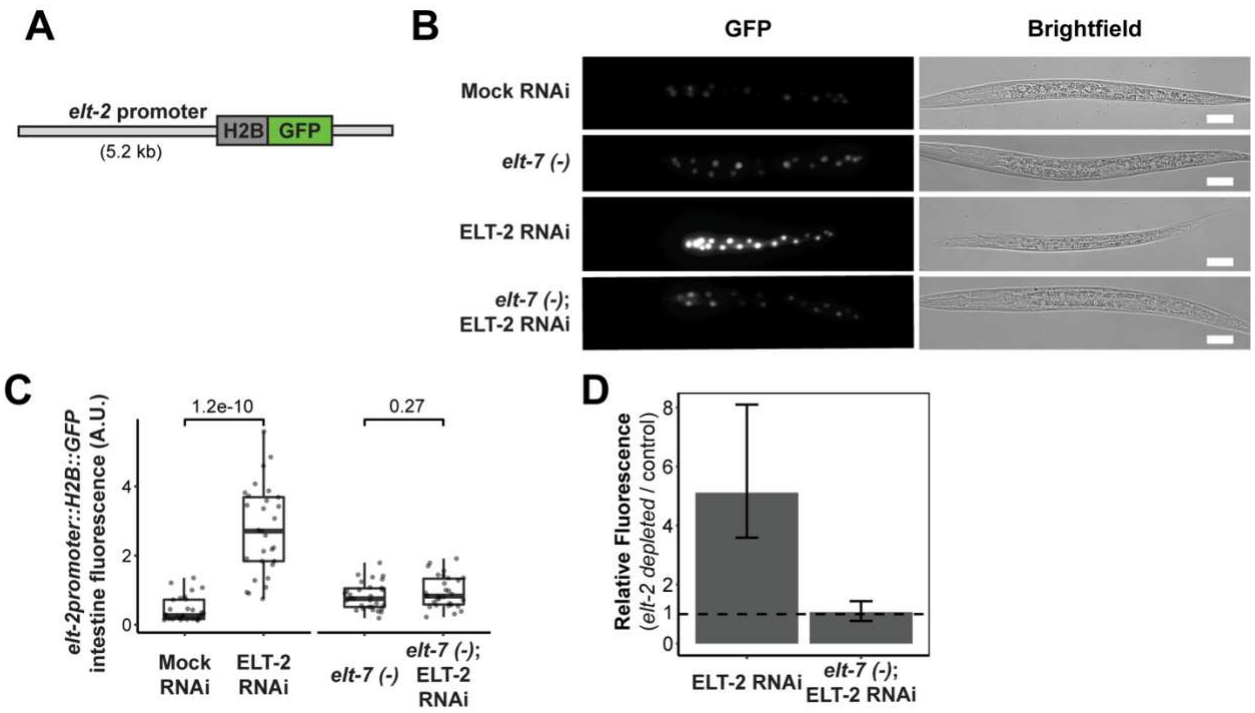

Figure 2.15: ELT-2 negatively regulates its own promoter. (A) Diagram of *elt-2* promoter reporter (*caIS71[elt-2p::H2B::GFP]* allele) used in this analysis. **(B)** Representative images of *elt-2* promoter activity in immobilized L1 stage animals. *elt-2* promoter reporter fluorescence was measured in *elt-7* wildtype or *elt-7 (tm840)* loss-of-function genetic background. Mock RNAi (L4440) or ELT-2 RNAi depletion were performed. Consistent imaging exposure was used between treatments and genetic backgrounds. Brightfield images are also shown. Animals were imaged at 40x. Scale bars = 20 microns. **(C)** Quantification of the *elt-2* promoter reporter's fluorescence intensity. This analysis included intestine GFP signal quantification, background subtraction, and area normalization using ImageJ. Student's t-test was used to determine statistical significance of measured fluorescence signal. Data represents 30 worms per treatment for three biological replicates. Box and whisker plots display the data point spread. Individual measurements are overlaid as points. **(D)** Comparison of *elt-2* promoter reporter activity for *elt-2* RNAi conditions with or without *elt-7* activity. Horizontal dotted line indicates relative fluorescence of 1. Error bars represent the t-test 95% confidence interval for the ratio of means between *elt-2* RNAi and control RNAi measurements.

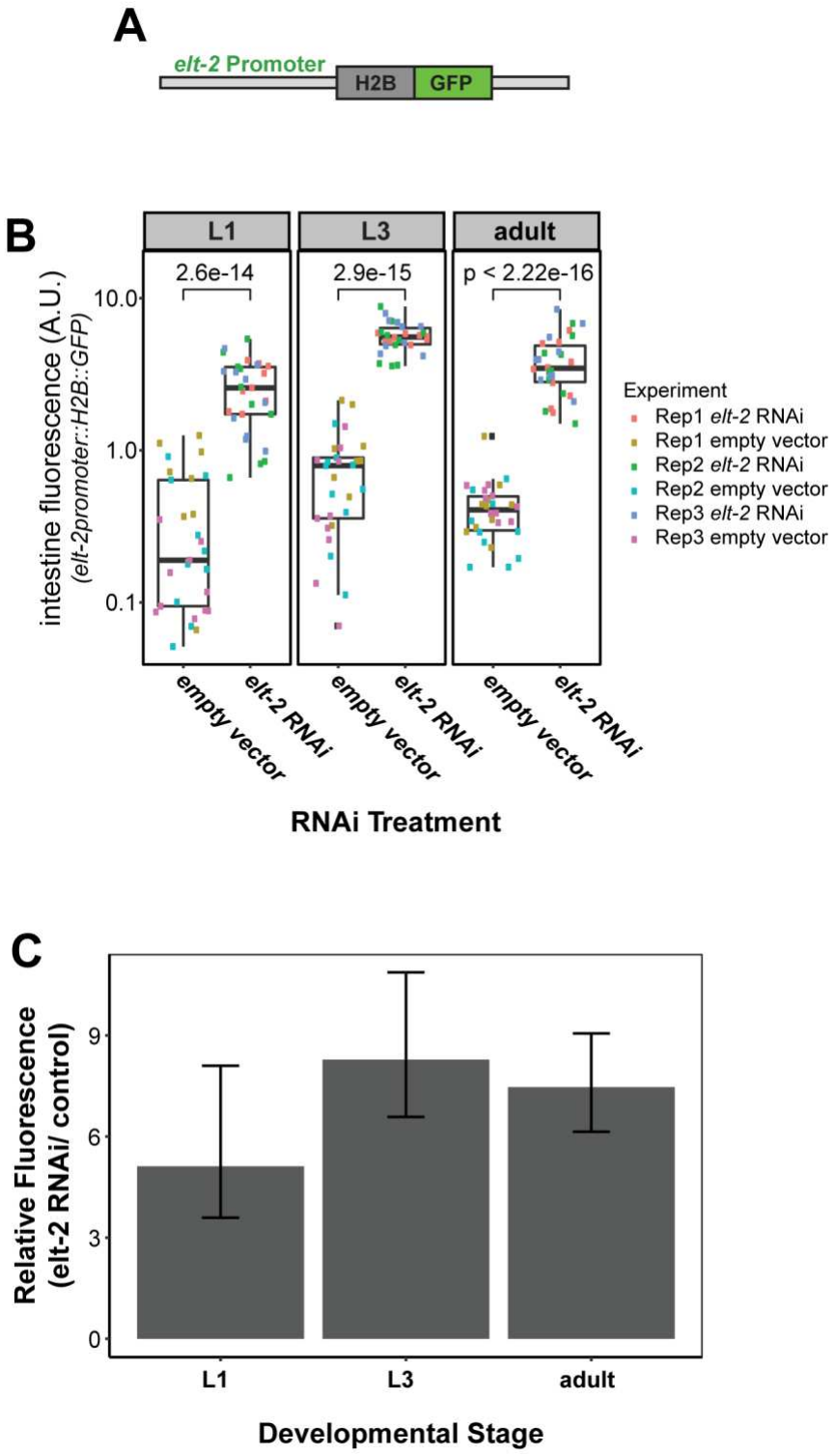

Figure 2.16. Negative regulation of the *elt-2* promoter is observed in L1, L3 and adult stages. (A) Block diagram of the *elt-2* promoter reporter allele used in this study. **(B)** Quantification of elt-2 promoter intestine fluorescence in either control RNAi (L4440) or *elt-2* RNAi conditions. Measurements were performed in L1, L3 and adult stages. GFP measurements were collected from background subtracted maximum Z projections from three

biological replicates and normalized for intestine area (n = 30 worms for all stages and RNAi treatments). Points represent a single worm and are colored based on biological replicate. Data distribution is represented by box and whisker plots. Student's t-test was used for significance testing. Plot y-axis is log-scale. **(C)** Relative fluorescence of *elt-2* promoter reporter activity measured in *elt-2* RNAi over control RNAi in L1, L3 and adult developmental stages. Error bars represent the t-test 95% confidence interval for the ratio of means between *elt-2* RNAi and control RNAi measurements.

elt-2 genomic locus for RNA-seq reads in *elt-2(-)* mutants with or without *elt-7* (Figure 2.17A). We measured increased reads aligning to the *elt-2* promoter in *elt-2(-)* mutants compared to wildtype (Figure 2.17B). Additionally, the increase in *elt-2* promoter-localized RNA-seq reads were reduced in *elt-2(-);elt-7(-)* double mutants. These results serve to support the biological significance of our findings in a reporter construct context.

Together, these results suggest that ELT-2 protein negatively regulates its own promoter in an ELT-7 dependent manner. We suggest that ELT-7 is responsible for over-activating the *elt-2* promoter when relieved of ELT-2 repression. Previous transcriptomics work has shown that *elt-7* is not over-expressed upon loss of *elt-2* at the mRNA level [47], suggesting that the higher activity of ELT-7 on the *elt-2* promoter is post-transcriptional, potentially occuring through an increased ability of ELT-7 to stimulate transcriptional output. Intriguingly, *elt-2* is still expressed at wildtype levels when both ELT-2 and ELT-7 are absent, implying that an additional factor may contribute to *elt-2* transcription (Figure 2.15C, D). These findings demonstrate that the wildtype level of *elt-2* mRNA production is affected by both positive and negative inputs. This example emphasizes combinatorial control by both ELT-2 and ELT-7 in regulating the intestine gene regulatory network, and illustrates that an additional unknown factor (or factors) participates in the intestine gene regulatory network.

2.3 Discussion

ELT-2 is the major transcription factor that marks intestinal tissue. ELT-2 production commences in embryos where its role is well studied, yet it remains present during larval and adult stages where its impact is more nebulous. *C. elegans* larval and adult stage intestines perform digestion and metabolism characteristic of most animal intestines. However, they also perform additional functions that in other animals are overseen by separate organs such as the

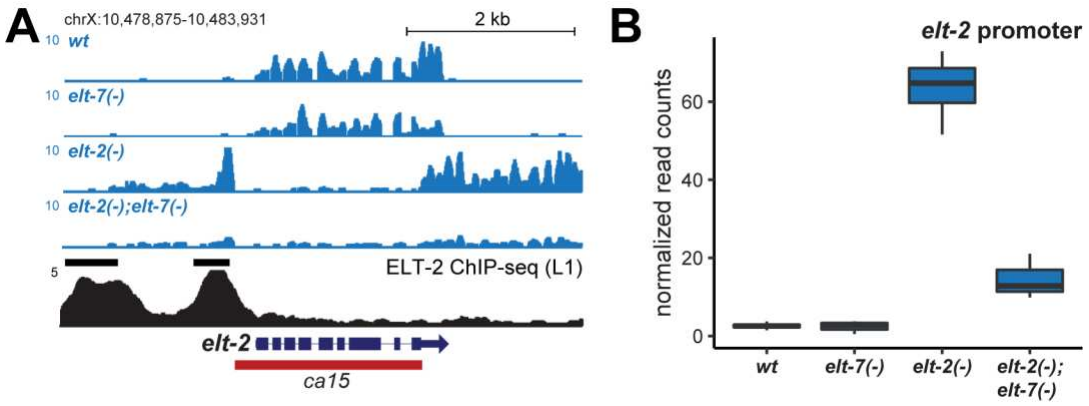

Figure 2.17: Visualization of ELT-2 negative autoregulation in RNA-seq data. (A) Genome browser tracks of the *elt-2* locus. RNA-seq tracks (blue) display increased read density aligning to the *elt-2* promoter in *elt-2(-)* genetic background compared to wildtype (WT, N2). The increase in *elt-2* promoter read density is reduced in *elt-2(-);elt-7(-)* genetic background. **(B)** Quantification of RNA-seq reads aligning to the elt-2 promoter. The elt-2 promoter here is defined as the region spanning the TSS to -1kb upstream of the TSS. RNA-seq reads were normalized with the DESeq2 package.

functions of yolk production, insulin signaling, regulation of developmental aging, immune response, stress response, and detoxification [49,53,55,96,97,119–125].

ELT-2 is required for optimal performance of many of these sub-functions, but an integrative genome-wide analysis of the role of ELT-2 in the intestine gene regulatory network has yet to link ELT-2 to these sub-functions. In this study we used a systems biology approach to place ELT-2 within the larger gene regulatory network and were confronted with findings that countered previous assumptions that: 1) ELT-2 binding is associated with expression of all intestine genes, 2) ELT-2 functions solely as a transcriptional activator, and 3) ELT-2 performs positive autoregulation.

Previous work suggests that ELT-2 is predominantly responsible for intestine gene activation with high prevalence of TGATAA sequences in intestine gene promoters indicating ELT-2 as an activator of all intestine genes [45,48,67,68,126–129]. Here, we combined ELT-2 binding maps (ChIP-seq), transcriptome response to elt-2 (-) (RNA-seq), and our newly generated intestine transcriptome data to evaluate this hypothesis. We identified that 31% of embryo stage intestine enriched genes had observable ELT-2 binding in their promoters, but the percent of intestine enriched genes bound by ELT-2 increased over developmental time to 75% in the L3 stage. This suggests that not all intestine enriched genes are regulated by ELT-2, but that the role of ELT-2 appears to expand in later developmental stages. Further analysis with the intestine transcriptome data generated in this study will be of benefit for identifying additional intestine GRN components.

ELT-2 has largely been assumed to function predominantly as a transcriptional activator [48,68]. By integrating publicly available data we evaluated how ELT-2 target genes are dependent on ELT-2 for their expression. Roughly half of ELT-2 target gene transcription is unchanged in *elt-2* loss of function mutants. It is possible that ELT-2 target genes independent of ELT-2 regulation may be redundantly regulated by other intestine GATA transcription factors, such as ELT-7. Surprisingly, we identified approximately equal shares of ELT-2 direct target genes are

either activated or repressed. This result is contrary to the prevailing hypothesis that ELT-2 functions solely as an activator [48,68]. As expected, GO analysis demonstrated that ELT-2 activated target genes are associated with known intestine genes with functions related to lysosome and defense response (Figure 2.18A). However, ELT-2 repressed target genes were not strongly associated with ubiquitously expressed genes or genes expressed exclusively in any tissue. ELT-2 repressed target genes were strongly associated with innate immunity and defense response genes. We hypothesize that ELT-2 establishes a baseline expression of defense genes while also repressing other defense response genes, keeping them poised until ELT-2 repression is relieved (Figure 2.18B). It is likely that ELT-2 repression is relieved upon pathogen exposure, but molecular details regulating this process will need to be experimentally determined [113].

Of the repressed genes, we verified that ELT-2 target gene *cebp-1* and *ets-4* are negatively regulated within the intestine. *cebp-1* and *ets-4* are transcription factors that function in pathogen response with additional roles in neuronal axon regeneration [99,101,102,130–133]. This result is consistent with our hypothesis that ELT-2 repression likely serves to keep defense response genes off until a stimulus-triggered relief of ELT-2 repression, consistent with previous reports [97,134]. It is possible that negative regulation by ELT-2 could occur through negative feedback loops within the intestinal GRN or by recruiting repressive machinery. Further studies will be required to differentiate between these options.

Having identified that ELT-2 functions as both an activator and repressor, we were interested in reevaluating the prevailing hypothesis that ELT-2 performs positive autoregulation. Previous work established that ELT-2 binds its own promoter through microscopy, ChIP-seq, and *in vitro* binding assays [38,135]. The role of this binding was posited to maintain ELT-2 levels throughout the worm lifespan. Indeed, heat shock-induced ectopic over-expression of ELT-2 expression leads to ectopic *elt-2* promoter reporter activity [38,45,67]. However, the ELT-2 positive autoregulation hypothesis had yet to be evaluated by removing ELT-2 from the biological

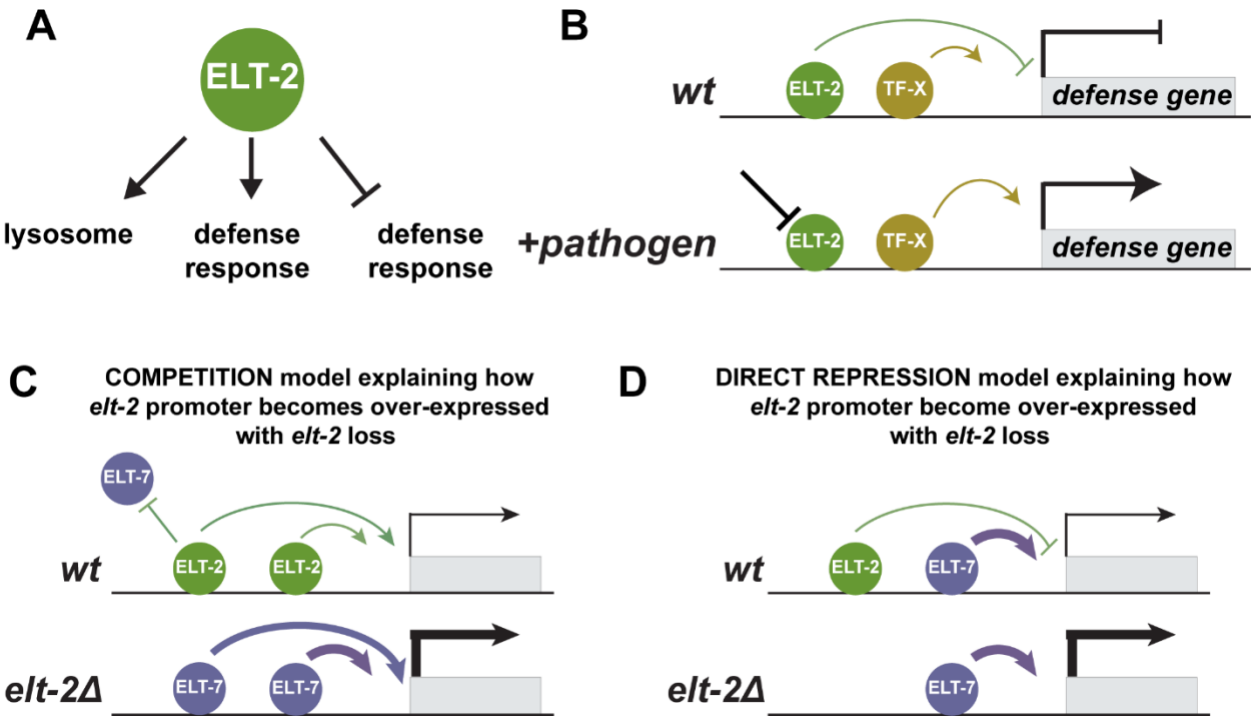

Figure 2.18: Model summarizing ELT-2 regulatory activity. (A) Through integration of genome-wide datasets, we found that ELT-2 serves to promote lysosome and defense response genes while simultaneously repressing a distinct set of defense response genes. **(B)** We hypothesize that ELT-2 repression of defense-response genes are inducible in situations such as pathogen exposure where relief of ELT-2 repression is required for defense response gene expression. Additionally, our data suggests that ELT-2 performs negative autoregulation with apparent ELT-7 overcompensation. We hypothesize two scenarios explaining this result. Either **(C)** ELT-2 and ELT-7 are directly competing for binding to the *elt-2* promoter and ELT-7 serves as a stronger activator than ELT-2, or **(D)** ELT-2 is directly repressing the *elt-2* promoter and the ELT-7 is a simultaneously activating.

system. We found that loss of ELT-2 led to an upregulation of *elt-2* promoter activity in the intestine. Additionally, we identified that simultaneous depletion of ELT-2 and ELT-7 reduced *elt-2* promoter activity back to wildtype levels (Figure 2.15). This illustrates that ELT-7 is required for the over-expression of the *elt-2* promoter [45]. We propose two mechanisms for how ELT-2 and ELT-7 may regulate the *elt-2* promoter. ELT-2 and ELT-7 may both serve to activate *elt-2* and perform competition for the elt-2 promoter, with ELT-7 serving as a stronger activator (Figure 2.18C). Alternatively, ELT-2 may perform direct repression with the wildtype level of *elt-2* mRNA production a result of both positive and negative inputs (Figure 2.18D). Finally, the observation that *elt-2* promoter activity is still expressed even in the absence of both ELT-2 and ELT-7 protein suggests an unknown TF also contributes to its activation.

In both yeast and mammals, GATA factors act as dual activators and repressors with negative regulation of target genes occurring through either competition and direct repression [134,136–138]. Over-expression of human GATA factors GATA-4, GATA-5, and GATA-6 can inhibit tetracycline-catalyzed induction by out-competing higher activity tetracycline activators due to naturally occurring GATA sites within the tetO promoter [139], illustrating a competitive mechanism of repression. In contrast, numerous examples of direct repression have been documented [94]. In humans, GATA factors catalyze cell differentiation, act as either tumor suppressors or oncogenes, and serve as critical markers of cancer onset and progression [136–138,140,141]. Human GATA-1 positively regulates subsets of direct targets through the recruitment of histone acetyltransferase complexes (CBP/p300) [142] but represses others through recruitment of direct repressive factors such as PU.1, FOG-1, or the NuRD co-repressor complex [143,144]. In yeast, GATA factors function as metabolic switches responding to changing levels of amino acid availability. Remarkably, the yeast GATA factor Gaf1 responds to amino acid depletion by activating amino acid biosynthetic genes through RNA Pol II activation while simultaneously repressing tRNA transcription through RNA Pol III repression [145]. While there is ample evidence

for both competitive relationships and direct repression by GATA factors in other systems, further experimentation will be required to determine which applies to ELT-2 in *C. elegans*.

2.4 Materials and methods

C. elegans strains and culture conditions

All worm strains were maintained as described [146] and cultured at 20°C on NGM plates seeded with OP50 E. coli unless otherwise stated. The wild-type strain N2 (Bristol) was used. Transgenic strains used in this study are below:

OP56 *unc-119(ed3) III; wgIs56 [elt-2::TY1::EGFP::3xFLAG(92C12) + unc-119(+)]*

OP563 *unc-119(tm4063) III; wgIs563 [cebp-1::TY1::EGFP::3xFLAG + unc-119(+)]*

OP509 *unc-119(tm4063) III; wgIs509 [ets-4::TY1::EGFP::3xFLAG + unc-119(+)]*

OP201 *unc-119(tm4063) III; wgIs201 [pqm-1::TY1::EGFP::3xFLAG(92C12) + unc-119(+)]*

JM149 *caIs71[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]*

JM259 *elt-7(tm840) V; caIs71[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]*

Some strains were provided by the Caenorhabditis Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

C. elegans culture and preparation of dissociated *C. elegans* cells

A detailed protocol for C. elegans culture is available in Appendix B and on the protocols.io platform: dx.doi.org/10.17504/protocols.io.8epv59zjng1b/v2. In brief, C. elegans were cultured on agar-based NGM plates to reduce any confounding effects that may be introduced by large scale liquid culture. To produce synchronized cultures sufficient for FACS, two rounds of mixed stage culture growth were performed followed by two rounds of synchronized growth. *C. elegans* were fed E. coli strain OP50 and maintained at 20°C. Synchronized worm populations were achieved through hypochlorite treatment.

For embryo stage experiments, 100,000 synchronized embryos were seeded onto 20 total 150 mm NGM plates at a density of 5,000 embryos per plate. Worms were cultured for 72 hours until gravid and harvested for mixed stage embryos through hypochlorite treatment. Embryo stage worms were dissociated through Chitinase and Pronase E treatment and mechanical disruption with a 21-gauge syringe needle. A detailed protocol for embryo dissociation is available in Appendix B and on the protocols.io platform: https://dx.doi.org/10.17504/protocols.io.dm6gpbw9plzp/v1 .

For L1 stage experiments, mixed stage embryos were synchronized to the L1 stage by incubating in M9 overnight rotating in a 20°C incubator for 24 hours. L1 worms were then fed OP50 E. coli on peptone enriched NGM plates for six hours before cell dissociation to reduce any observable starvation-induced responses in the measured transcriptional data. L1 stage worms were dissociated through SDS-DTT and Pronase E treatment and mechanical disruption with a Dounce homogenizer. A detailed protocol for L1 dissociation is available in Appendix B and on the protocols.io platform: https://dx.doi.org/10.17504/protocols.io.rm7vzy365lx1/v1 . For L3 stage experiments, synchronized L1 worms were cultured for an additional 48 hours on peptone enriched NGM plates for 48 hours at 20°C. L3 stage worms were dissociated similarly to L1, where the final centrifugation speed for cell harvest was reduced from 100 rcf to 20 rcf. A detailed protocol for L3 dissociation is available in Appendix B and on the protocols.io platform: https://dx.doi.org/10.17504/protocols.io.14egn7zwmv5d/v1 .

FACS isolation of intestine cells

GFP+ intestine cells were isolated from GFP- non-intestine cells by FACS using a BD FACSAria III cell sorter. Additional dyes were used depending on developmental stage. For embryo stage experiments, viability dye propidium iodide was used to separate live cells (PI-) from dead cells (PI+). For L1 stage experiments, viability dyes were not used because preliminary experiments identified that intestine cells preferentially take up viability dyes

confounded sorting (data not shown). For L3 stage experiments, viability dyes are also not used. We observed that L3 stage cell preps have a high degree of debris, so a cell permeable nucleic acid dye such as DRAQ5 was used to distinguish cells (DRAQ5+) from debris (DRAQ5-). FACS plots and post-hoc analysis were performed with FlowJo (v 10) (Figure 1, S1). Cell staining protocols and gating strategies (Figure S1) are available in Appendix B and on the protocols.io platform: https://dx.doi.org/10.17504/protocols.io.j8nlkk43dl5r/v1 .

Approximately 300,000 GFP+ intestine cells and 1,000,000 GFP- non-intestine cells were collected for RNA isolation for each developmental stage. Three biological replicates for each developmental stage were performed. FACS isolated cells were pelleted by centrifuging at 10,000 rcf for 5 mins in a 4°C cooled centrifuge. The supernatant was removed from the cell pellet and resuspended in 1 ml Qiazol for RNA extraction. Cells in Qiazol were stored at -80°C until RNA extraction was performed.

RNA extraction

Total RNA extractions were performed with miRNeasy Micro kit (Qiagen 217084). RNA quantity was measured using a Qubit 3.0 fluorometer (Thermo Fisher Q33216) and Qubit High Sensitivity RNA assay kit (Thermo Fisher Q32852). RNA was of a consistently high quality (RIN > 8.0) as determined on an Agilent 2200 TapeStation using high sensitivity RNA reagents (Agilent 5067-5579). Total RNA samples were stored at -80°C until RNA-seq library preparation was performed.

RNA-seq library preparation and analysis

To measure the intestine specific transcriptome, RNA-seq library preparation was performed with the NEBNext Ultra II RNA Library Prep Kit and sequenced with the Illumina platform (NEB E7770S). Ribosomal RNAs (rRNAs) were depleted through the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB E7490S). All samples were sequenced on the Illumina

NovaSeq 6000 instrument except for replicates 1 and 2 for L1 stage samples, which were sequenced on an Illumina HiSeq X-ten instrument. Sequencing reads were processed using a custom pipeline on the Rocky Mountain Advanced Computing Consortium Summit Supercomputer. Processing pipeline included the following steps: 1) sequencing reads were filtered for low quality or adapter sequence containing reads with *fastp* (v 0.20.0) [147], 2) aligned to the ce11 *C. elegans* genome (c_elegans.PRJNA13758.WS263.canonical_geneset.gtf) with *hisat2* (v 2.1.0) [148], 3) tabulated for number of mapped reads aligning to the WS263 genome assembly using *featureCounts* (v 1.6.4) [149]. Sequencing data collected is available through NCBI Gene Expression Omnibus (GEO) database under accession number GSE214581.

Differential expression analysis and plotting was performed with the *DESeq2* package (v 1.28.1) in the R Environment (v 4.0.3, BiocManager, v 1.30.10, tidyverse v 1.3.1). [150–153]. RNA-seq count data were filtered for detected genes (> 10 counts per million across all samples). To visualize genome-wide similarities between samples, sample-to-sample Euclidian distances were calculated using the Complete cluster method on normalized, variance-stabilized, log transformed count data (Figure S. In all pairwise differential expression analyses, genes were identified as differentially expressed if they had a normalized log2-transformed fold change value greater than 1 with a Benjamini-Hochberg corrected p-value less than 0.01 (Figure 2, Figure S3). UpSet plots were generated with the UpSet function of *ComplexHeatmap* (v 2.8.0) [154].

ELT-2 ChIP-seq data and analysis

ELT-2 ChIP-seq data was downloaded from the ModERN Resource for embryo, L1, and L3 stages (https://epic.gs.washington.edu/modERN/) [72]. Downloaded files included aligned reads and optimal IDR thresholded peaks. Peaks were assigned to genes based on presence of ≥1 IDR passing peak in a window of -1kb and +0.2kb centered on a gene TSS. ELT-2 ChIP-seq heatmaps were generated with *deeptools* (v 3.5.0) [155].

elt-2(-) RNA-seq data and analysis

RNA-seq data measuring the transcriptional response to *elt-2* deletion was downloaded from the publication supplemental material [47]. Files contained per gene aligned RNA-seq read count data. Differential expression analysis and plotting was performed with the *DESeq2* package (v 1.28.1) in the R Environment (v 4.0.3, BiocManager, v 1.30.10, tidyverse v 1.3.1). [150–153]. Genes were identified as differentially expressed if they had a normalized log2-transformed fold change value greater than 1 with a Benjamini-Hochberg corrected p-value less than 0.01.

Gene set definitions

Tissue-specific genes were generated by downloading gene lists associated with the following anatomy ontology terms from WormBase:

WBbt:0005730 (epithelial system)

WBbt:0005735 (nervous system)

WBbt:0005736 (excretory system)

WBbt:0005737 (muscular system)

WBbt:0005747 (reproductive system)

WBbt:0005749 (coelomic system)

WBbt:0005772 (intestine)

The list of ubiquitous genes was previously defined [95]. Tissue-specific or ubiquitous gene set definitions were made by retaining genes from the input lists that were unique to one category. The hypergeometric test was utilized to evaluate if a given set of genes corresponded to any gene set listed above. The list of known and putative *C. elegans* transcription factors was provided in the wTF3.0 TF list [98].

Computational Analysis

RNA-seq data and ELT-2 target genes were subset for protein coding genes in all analyses. All analysis and plots were performed in the R Environment unless otherwise stated (v 4.0.3, BiocManager, v 1.30.10, tidyverse v 1.3.1) [150–153]. All scripts for analysis are available in the following GitHub repository: https://github.com/meekrob/ELT-2-ChIP-revision .

Gene ontology analysis

Gene Ontology (GO) terms for gene sets were evaluated using R package topGO (v 2.44.0) [157]. GO analysis was performed using the "weight01" algorithm and the "fisher" statistic. When measuring ontology terms for intestine enriched FACS samples (Figure 2), background gene sets included all genes in the genome. When measuring ontology terms for ELT-2 regulated targets (Figure 6), background gene sets included all genes with an ELT-2 bound promoter.

RNAi feeding

To visualize the effect of ELT-2 depletion, we used *E. coli* feeding strains engineered to produce double-stranded RNA (dsRNA) cognate to *elt-2* transcript. RNAi feeding strains were obtained from the Ahringer RNAi feeding library [157]. Freshly starved worms were chunked to large NGM/OP50 plates to grow for 48 hours. Synchronized embryos were isolated from mixed stage populations through hypochlorite treatment [146]. To visualize the effect of ELT-2 depletion on target gene expression (Figure 7), we grew synchronized embryos to L3 stage by incubation at 20°C for 48 hours before RNAi exposure. L3 stage worms were transferred to NGM plates seeded with RNAi-inducing *E. coli* and incubated for an additional 48 hours until gravid. Synchronized L1s were collected by transferring RNAi treated gravid adults to 50 ml drops of M9 and culturing for 24 hours. To image the response of *elt-2* promoter across developmental

stages to ELT-2 depletion, we modified the above procedure by exposing worms to RNAi feeding strains 24 hours before the queried developmental stage (Figure 8, Figure S8). Negative control experiments utilized feeding strains containing L4440 empty vector. To confirm RNAi efficiency, we utilized POP-1 targeted RNAi as a positive control. RNAi experiments were considered successful when POP-1 RNAi resulted in 100% embryonic lethality. All RNAi plasmids were sequence verified before experimental use. Three biological replicates were performed for each GFP translation reporter strain.

Fluorescence Microscopy

Response of GFP translation reporter gene fluorescence to ELT-2 depletion was visualized using fluorescence microscopy. RNAi-treated worms were imaged on a DeltaVision Elite inverted microscope equipped with an Olympus PLAN APO 60x (1.42 NA, PLAPON60XOSC2) objective, an Insight SSI 7-Color Solid State Light Engine and SoftWorx software (Applied Precision) using 1 µm z-stacks. To remove birefringent gut granule autofluorescence, animals were fixed and permeabilized through liquid nitrogen freeze-crack and subsequent washes with methanol, acetone, and PBST.

For *elt-2* promoter experiments, images were collected using a Keyence BZ-X800 fluorescence microscope equipped with a Keyence PLAN APO 40x (0.95 NA, BZ-PA40) objective using 1 µm z-stacks. Worms were paralyzed with 25 mM sodium azide (Sigma Aldrich S2002) and mounted on microscope slides prepared with 5% agar pads and sealed with Valap. Imaging settings were kept constant between RNAi treatments and developmental stages per strain. Multi-panel images were processed and stitched, maximum Z projections were generated, and intestine fluorescence reporter was measured in ImageJ distribution Fiji [158,159].

Signal quantification was restricted to the intestine by utilizing a DIC reference image. Analysis and plots were generated using custom R scripts. Where ratios of reporter intensity are

shown against controls with a p-value and confidence interval, a *t*-test of the ratio of means was performed using the *ttestratio* function in R package *mratios* (v 1.4.2), with the null hypothesis of a ratio of 1. Confidence intervals are 95% standard error around the ratio of means [160].

Single molecule inexpensive fluorescence *in situ* hybridization (smiFISH)

smiFISH was used to confirm intestine specificity of bioinformatically identified intestine genes. Transcripts were probed in embryo stage using smiFISH as previously described [91,92]. Probes were designed with FLAPY extensions for *acy-4*, *bre-2*, and *C02D5.4* and FLAPX for *ges-1*. For each gene 12 primary probes were designed with the R script Oligostan (commit d87b4b2). Probe sequences are included in Appendix C. Secondary FLAP probes were ordered from Stellaris LGC with dual 5′ and 3′ fluorophore labeling. FLAPY probes were labeled with Cal Fluor 610 and FLAPX probers were labeled with Quasar 670 (Biosearch Technologies, BNS-5082 and FC-1065, respectively). All samples were imaged on a DeltaVision Elite inverted microscope described above. Images were processed in the ImageJ distribution Fiji [158].

CHAPTER 3

FUTURE DIRECTIONS AND DISCUSSION

3.1 Future directions for FACS intestine isolation and intestine transcriptome atlas

In this dissertation, we developed methodology for isolating intestine cells in embryonic and larval *C. elegans* through Fluorescence Activated Cell Sorting (FACS, Appendix B). By allowing tissue specific application of biochemical and molecular approaches, this methodology facilitated dissection of the intestine gene regulatory network. We utilized FACS intestine isolation to profile the intestine transcriptome at three distinct developmental stages in bulk cell samples (Chapter 2). We found that the intestine is developmentally dynamic between the embryo and L1 developmental stages, suggesting a rapid shift in GRN components between these stages. Other tools such as measuring intestine chromatin accessibility, intestine specific transcription factor profiling (through ChIP-seq, CUT&TAG, or CUT&RUN), or histone modifications will further link intestines-specific chromatin changes to intestine-specific gene expression [161–164]. Data generated through these applications will provide a clearer perspective of the molecular mechanisms governing the intestine GRN in development. A clear future direction could be distinguishing the regulatory components common and distinct to each stage to determine which factors may be driving the observed differences between the embryonic and post-embryonic intestine transcriptome

We also utilized this technique to perform single-cell RNA-seq (scRNA-seq) analysis on purified intestine cell populations (Appendix A). As demonstrated in our bulk intestine transcriptome data, it appears that the embryo transcriptome is distinct from larval stages. It is likely that the intestine undergoes rapid transcriptional changes during embryonic development when cells are rapidly dividing, differentiating, and undergoing organogenesis. The application of scRNA-seq provides an opportunity to assay the complete transcriptome of thousands of

individual cells providing an opportunity to resolve distinct cell subtypes within a population of cells [14,15]. Previous work profiling single cell transcriptomes for the *C. elegans* embryo do not currently provide a complete picture of intestine development [86,165]. Through scRNA-seq analysis of FACS isolated embryo stage intestine cells, we identified 6 clusters of intestine cell populations with distinct transcriptome profiles, demonstrating additional complexities in the embryonic intestine GRN. Through computational analysis, we identified intestine TFs and key intestine marker genes distinct to each cluster, thereby providing an opportunity to investigate key TFs for the embryonic intestine development. Whether these clusters represent different stages of intestine development or different sub-regions of intestine tissue remains unclear. Together, these datasets will serve as rich atlases for future work dissecting the *C. elegans* intestine GRN.

3.2 Future directions for the genome-wide characterization of GATA TF ELT-2

Through integration of genome-wide resources we performed a systematic investigation of the regulatory role of GATA TF ELT-2 in the developing intestine (Chapter 2). Surprisingly, we found that approximately 50% of ELT-2 targets are not transcriptionally dependent on ELT-2 regulation alone. Of those genes that were dependent on ELT-2, equal shares exhibited activation or repression. By analyzing the sets of activated and repressed ELT-2 targets, we found that ELT-2 activated genes are involved in lysosome and defense response processes, while ELT-2 repressed genes are involved in defense response alone. Through fluorescence microscopy we found evidence that ELT-2 negatively impacts direct target TFs *cebp-1* and *ets-4* which are known to participate in defense response, axon regeneration, and life span regulation [102,130,132,133,166]. These findings emphasize recent studies demonstrating a key role for ELT-2 in defense response [53,55–57,96,97,113,134]. We hypothesize that ELT-2 establishes a baseline expression of defense genes while also repressing other defense response genes, keeping them poised until ELT-2 repression is relieved. Furthermore, our results suggest that ELT-2

influences a larger proportion of the intestine transcriptome in larval stages compared to the embryonic stage. It is possible that other TFs work together with ELT-2 in embryo stages in intestine differentiation that are distinct from larval stages. Together, these results demonstrate that the role of ELT-2 is dynamic over developmental time.

The *C. elegans* intestine immune response has been demonstrated to be regulated by the conserved p38-MAPK pathway [167]. In this pathway, p38-driven phosphorylation cascades lead to NF-kB (nuclear factor kappa B) activation. NF-kB is a key transcription factor that serves to regulate the induced infection response. However, studies have shown that *C. elegans* lacks a clear homolog to NF-kB [168]. This raises the question of whether ELT-2 serves as the replacement for NF-kB in *C. elegans* to regulate induced infection responses. In support of this hypothesis, it has been demonstrated that ELT-2 functions downstream of the p38 pathway and participates with key TFs ATF-7 and SKN-1 [97,113,134]. It has yet to be determined if ELT-2 is a target of p38 phosphorylation, the demonstration of which may further support this hypothesis. Future studies could investigate if ELT-2 is regulated by phosphorylation during immune response through SDS-PAGE, Phos-tag gel, or mass spectrometry approaches [169–171].

Our findings lead to a suite of new questions surrounding the molecular basis for ELT-2 regulation. What sequence features differentiate ELT-2 activated and repressed genes. Through DNA motif analysis with the currently available data, we may be able to determine DNA sequences or TF binding site motifs that are distinct between activated and repressed ELT-2 targets. Furthermore, through computational analysis we may be able to determine the identity of TFs that work in combination with ELT-2 by utilizing TF binding maps provided by the modENCODE and modERN projects [72,172]. Alternatively, we may be able to utilize a proximity labeling approach to identify ELT-2 binding partners that are distinct between activated and repressed ELT-2 target in a loci-specific manner in isolated intestine tissue [173]. Furthermore, it is possible that ELT-2 regulates the intestine GRN by reorganizing chromatin structure to promote recruitment of additional transcriptional TFs, as GATA TFs have been implicated as pioneer

factors which remodel chromatin [94,174,175]. By perturbing the intestine GRN through ELT-2 depletion, we may measure differences in chromatin accessibility upon ELT-2 depletion revealing how activated and repressed ELT-2 targets are differentially regulated. For whichever mechanism holds true, it is likely that ELT-2 recruits transcriptional machinery differentially to activated and repressed target genes.

The finding that ELT-2 performs negative autoregulation is an interesting story. This result was initially found by accident when I was performing early RNAi experiments. I was unsure if RNAi would work in my hands, so I set up an experiment with transgenic strains I had available with the expected outcomes based on what I read in the literature. The expectation I had when performing ELT-2 RNAi was that I should visualize a reduction in both translation reporter and promoter reporter fluorescence. In strains treated with ELT-2 RNAi, I was relieved to observe the ELT-2 translation reporter signal disappear but was confused and surprised to visualize brighter fluorescence than mock RNAi in the *elt-2* promoter reporter strain. I was convinced I did something wrong but asked my talented undergraduate Izabella to investigate this further. After three biological replicates of this experiment, I was convinced that this result wasn't a mistake and that decades of ELT-2 literature were now in question with one simple experiment. After discussion with Jim McGhee (the father of ELT-2 research), he suggested we evaluate if *elt-7* is required for the observed upregulation in *elt-2* promoter fluorescence and provided the key strain to ask this question before his retirement. We found that upregulation of the *elt-2* promoter in the absence ELT-2 protein does not occur when *elt-7* is absent. Together, these results suggest that ELT-2 protein negatively regulates its own promoter, and that ELT-7 is responsible for over-activating the *elt-2* promoter when relieved of ELT-2 repression

These results provide a unique opportunity to refine a previously established genetic network and has led to many interesting questions for the future. Previous work has shown that ectopic expression of ELT-2 is able to ectopically induce *elt-2* promoter activity [176]. If our hypothesis regarding ELT-2 and ELT-7 regulating the *elt-2* promoter is correct, then in this same

experimental setting we would not observe ectopic *elt-2* promoter activity when ELT-2 is ectopically expressed in the absence of *elt-7*. Additionally, we proposed two mechanisms for how ELT-2 and ELT-7 may regulate the *elt-2* promoter. In the competition model, ELT-2 and ELT-7 may both serve to activate *elt-2* and perform competition for the *elt-2* promoter, with ELT-7 serving as a stronger activator. Alternatively, in the direct repression model ELT-2 may perform direct repression where the wildtype level of *elt-2* mRNA production is a result of both positive and negative inputs. Future studies can distinguish between these two models by measuring TF occupancy in the *elt-2* promoter when both factors are reciprocally removed from the system. While the purpose of elt-2 negative autoregulation is not immediately clear, modeling experiments in yeast and E. coli suggest that this genetic circuit leads to faster and linear protein production with a stabilized amount of produced protein [177–179].

Initially, ELT-2 was named the intestinal master regulator, but, in fact, ELT-2 is supported by its homolog ELT-7 [45,48,67]. That is, loss of *elt-2* and *elt-7* in combination produces a more severe lethal phenotype than the loss of elt-2 alone in which the intestine lumen is discontinuous, though both are lethal [45,47]. In contrast, loss of elt-7 alone produces worms that are superficially wild type with only a several minutes delay in hatching compared to wildtype [47]. This relationship implies that ELT-2 can direct intestine development in the absence of *elt-7*, but the reverse is not true. ELT-2 and ELT-7 are therefore said to be partially redundant. Though their relationship has been studied, several questions remain. For example, ELT-2 and ELT-7 bind identical DNA sequences in vitro, are expressed in the same tissues at roughly the same levels and have 64% identical amino acid identity in the zinc finger protein domain [38,47]. Intriguingly, when placed under genetic control of the *elt-2* promoter, ELT-7 can replace ELT-2 to specify the intestine lineage [47]. Furthermore, elt-7 under control of both *elt-2* and *end-1* promoters can completely replace the other GATA TFs in the intestine GRN emphasizing the redundant relationship these factors share (*end-1*, *end-3* and *elt-2* deletion). These complexities

demonstrate that it is still unclear why ELT-7 is unable to perform the same role as ELT-2 and what makes ELT-2 unique.

3.3 Overall discussion

A major effort in biology involves determining the complete atlas of transcription factors, their expression patterns, their regulatory networks, and how their downstream effects interplay to generate a functional organ. This dissertation has contributed to this effort by providing an atlas of *C. elegans* intestine transcriptomes over developmental time. Additionally, we utilized systems biology approaches to better understand the regulatory role of a key TF in intestine development. The tools and methodologies established in this dissertation can be extended to additional factors in the intestine GRN thereby leading to a complete understanding of the molecular mechanisms for organogenesis. These findings can then be translated to other organisms, providing a clearer picture for organismal development.

REFERENCES

1. Abbott, S. & Fairbanks, D. J. Experiments on Plant Hybrids by Gregor Mendel. Genetics 204, 407–422 (2016).

2. Hershey, A. D. & Chase, M. INDEPENDENT FUNCTIONS OF VIRAL PROTEIN AND NUCLEIC ACID IN GROWTH OF BACTERIOPHAGE. J Gen Physiology 36, 39–56 (1952).

3. CRICK, F. Central Dogma of Molecular Biology. Nature 227, 561–563 (1970).

4. Oudelaar, A. M. & Higgs, D. R. The relationship between genome structure and function. Nat Rev Genet 22, 154–168 (2021).

5. Trainor, P. A. Chapter Thirty Developmental Biology We Are All Walking Mutants. Curr Top Dev Biol 117, 523–538 (2016).

6. Emmert-Streib, F., Dehmer, M. & Haibe-Kains, B. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. Frontiers Cell Dev Biology 2, 38 (2014).

7. Davidson, E. H. & Peter, I. S. Genomic Control Process. 41–77 (2015). doi:10.1016/b978-0-12-404729-7.00002-2

8. Peter, I. S. & Davidson, E. H. Chapter Thirteen Implications of Developmental Gene Regulatory Networks Inside and Outside Developmental Biology. Curr Top Dev Biol 117, 237–251 (2016).

9. Ideker, T., Galitski, T. & Hood, L. A NEW APPROACH TO DECODING LIFE: Systems Biology. Annu Rev Genom Hum G 2, 343–372 (2001).

10. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. Genome Biol 18, 83 (2017).

11. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. Science 316, 1497–1502 (2007).

12. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10, 57–63 (2009).

13. Chu, Y. & Corey, D. R. RNA Sequencing: Platform Selection, Experimental Design, and Data Interpretation. Nucleic Acid Ther 22, 271–274 (2012).

14. Saliba, A.-E., Westermann, A. J., Gorski, S. A. & Vogel, J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42, 8845–8860 (2014).

15. Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Medicine 50, 1–14 (2018).

16. Angelini, C. & Costa, V. Understanding gene regulatory mechanisms by integrating ChIP-seq and RNA-seq data: statistical solutions to biological problems. Frontiers Cell Dev Biology 2, 51 (2014).

17. Muhammad, I. I., Kong, S. L., Abdullah, S. N. A. & Munusamy, U. RNA-seq and ChIP-seq as Complementary Approaches for Comprehension of Plant Transcriptional Regulatory Mechanism. Int J Mol Sci 21, 167 (2019).

18. Tremblay, M., Sanchez-Ferras, O. & Bouchard, M. GATA transcription factors in development and disease. Development 145, dev164384 (2018).

19. Deng, N., Goh, L. K., Wang, H., Das, K., Tao, J., Tan, I. B., Zhang, S., Lee, M., Wu, J., Lim, K. H., Lei, Z., Goh, G., Lim, Q.-Y., Tan, A. L.-K., Poh, D. Y. S., Riahi, S., Bell, S., Shi, M. M., Linnartz, R., Zhu, F., Yeoh, K. G., Toh, H. C., Yong, W. P., Cheong, H. C., Rha, S. Y., Boussioutas, A., Grabsch, H., Rozen, S. & Tan, P. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. Gut 61, 673 (2012).

20. Chia, N.-Y., Deng, N., Das, K., Huang, D., Hu, L., Zhu, Y., Lim, K. H., Lee, M.-H., Wu, J., Sam, X. X., Tan, G. S., Wan, W. K., Yu, W., Gan, A., Tan, A. L. K., Tay, S.-T., Soo, K. C., Wong, W. K., Dominguez, L. T. M., Ng, H.-H., Rozen, S., Goh, L.-K., Teh, B.-T. & Tan, P. Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. Gut 64, 707 (2015).

21. Romano, O. & Miccio, A. GATA factor transcriptional activity: Insights from genome-wide binding profiles. Iubmb Life 72, 10–26 (2020).

22. Aronson, B. E., Aronson, S. R., Berkhout, R. P., Chavoushi, S. F., He, A., Pu, W. T., Verzi, M. P. & Krasinski, S. D. GATA4 represses an ileal program of gene expression in the proximal small intestine by inhibiting the acetylation of histone H3, lysine 27. Biochimica Et Biophysica Acta Bba - Gene Regul Mech 1839, 1273–1282 (2014).

23. Takahashi, S., Onodera, K., Motohashi, H., Suwabe, N., Hayashi, N., Yanai, N., Nabesima, Y. & Yamamoto, M. Arrest in Primitive Erythroid Cell Development Caused by Promoter-specific Disruption of the GATA-1 Gene*. J Biol Chem 272, 12611–12615 (1997).

24. Pevny, L., Lin, C. S., D'Agati, V., Simon, M. C., Orkin, S. H. & Costantini, F. Development of hematopoietic cells lacking transcription factor GATA-1. Development 121, 163–172 (1995).

25. Cheng, Y., Wu, W., Kumar, S. A., Yu, D., Deng, W., Tripic, T., King, D. C., Chen, K.-B., Zhang, Y., Drautz, D., Giardine, B., Schuster, S. C., Miller, W., Chiaromonte, F., Zhang, Y., Blobel, G. A., Weiss, M. J. & Hardison, R. C. Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. Genome Res 19, 2172–2184 (2009).

26. Kassouf, M. T., Hughes, J. R., Taylor, S., McGowan, S. J., Soneji, S., Green, A. L., Vyas, P. & Porcher, C. Genome-wide identification of TAL1's functional targets: Insights into its mechanisms of action in primary erythroid cells. Genome Res 20, 1064–1083 (2010).

27. Tripic, T., Deng, W., Cheng, Y., Zhang, Y., Vakoc, C. R., Gregory, G. D., Hardison, R. C. & Blobel, G. A. SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. Blood 113, 2191–2201 (2009).

28. Yu, M., Riva, L., Xie, H., Schindler, Y., Moran, T. B., Cheng, Y., Yu, D., Hardison, R., Weiss, M. J., Orkin, S. H., Bernstein, B. E., Fraenkel, E. & Cantor, A. B. Insights into GATA-1-Mediated Gene Activation versus Repression via Genome-wide Chromatin Occupancy Analysis. Mol Cell 36, 682–695 (2009).

29. Soler, E., Andrieu-Soler, C., Boer, E. de, Bryne, J. C., Thongjuea, S., Stadhouders, R., Palstra, R.-J., Stevens, M., Kockx, C., IJcken, W. van, Hou, J., Steinhoff, C., Rijkers, E., Lenhard, B. & Grosveld, F. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Gene Dev 24, 277–289 (2010).

30. Corsi, A. K., Wightman, B. & Chalfie, M. A Transparent Window into Biology: A Primer on Caenorhabditis elegans. Genetics 200, 387–407 (2015).

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

32. Kaletta, T. & Hengartner, M. O. Finding function in novel targets: C. elegans as a model organism. Nat Rev Drug Discov 5, 387–399 (2006).

33. Brenner, S. In the Beginning Was the Worm …. Genetics 182, 413–415 (2009).

34. Bieri, T., Blasiar, D., Ozersky, P., Antoshechkin, I., Bastiani, C., Canaran, P., Chan, J., Chen, N., Chen, W. J., Davis, P., Fiedler, T. J., Girard, L., Han, M., Harris, T. W., Kishore, R., Lee, R., McKay, S., Müller, H.-M., Nakamura, C., Petcherski, A., Rangarajan, A., Rogers, A., Schindelman, G., Schwarz, E. M., Spooner, W., Tuli, M. A., Auken, K. V., Wang, D., Wang, X., Williams, G., Durbin, R., Stein, L. D., Sternberg, P. W. & Spieth, J. WormBase: new content and better access. Nucleic Acids Res 35, D506–D510 (2007).

35. Consortium*, T. C. elegans S. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 282, 2012–2018 (1998).

36. McGhee, J. The C. elegans intestine. Wormbook 1–36 (2007). doi:10.1895/wormbook.1.133.1

37. Zhang, F., Berg, M., Dierking, K., Félix, M.-A., Shapira, M., Samuel, B. S. & Schulenburg, H. Caenorhabditis elegans as a Model for Microbiome Research. Front Microbiol 8, 485 (2017).

38. Wiesenfahrt, T., Berg, J. Y., Nishimura, E. O., Robinson, A. G., Goszczynski, B., Lieb, J. D. & McGhee, J. D. The function and regulation of the GATA factor ELT-2 in the C. elegans endoderm. Dev Camb Engl 143, 483–91 (2015).

39. Maduro, M. F. & Rothman, J. H. Making Worm Guts: The Gene Regulatory Network of the Caenorhabditis elegans Endoderm. Dev Biol 246, 68–85 (2002).

40. Maduro, M. F., Kasmir, J. J., Zhu, J. & Rothman, J. H. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Dev Biol 285, 510–523 (2005).

41. Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R. & Rothman, J. H. Genetic redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol 284, 509–522 (2005).

42. Dimov, I. & Maduro, M. F. The C. elegans intestine: organogenesis, digestion, and physiology. Cell Tissue Res 377, 383–396 (2019).

43. Maduro, M. F. Evolutionary Dynamics of the SKN-1 → MED → END-1,3 Regulatory Gene Cascade in Caenorhabditis Endoderm Specification. G3 Genes Genomes Genetics 10, g3.400724.2019 (2019).

44. Choi, H., Broitman-Maduro, G. & Maduro, M. F. Partially compromised specification causes stochastic effects on gut development in C. elegans. Dev Biol 427, 49–60 (2017).

45. Sommermann, E. M., Strohmaier, K. R., Maduro, M. F. & Rothman, J. H. Endoderm development in Caenorhabditis elegans: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. Dev Biol 347, 154–166 (2010).

46. Ewe, C. K., Sommermann, E. M., Kenchel, J., Flowers, S. E., Maduro, M. F., Joshi, P. M. & Rothman, J. H. Feedforward regulatory logic controls the specification-to-differentiation transition and terminal cell fate during Caenorhabditis elegans endoderm development. Development 149, (2022).

47. Dineen, A., Nishimura, E. O., Goszczynski, B., Rothman, J. H. & McGhee, J. D. Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the C. elegans endoderm. Dev Biol 435, 150–161 (2018).

48. 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. ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. Dev Biol 327, 551–565 (2009).

49. Mann, F. G., Nostrand, E. L. V., Friedland, A. E., Liu, X. & Kim, S. K. Deactivation of the GATA Transcription Factor ELT-2 Is a Major Driver of Normal Aging in C. elegans. Plos Genet 12, e1005956 (2016).

50. Eurmsirilerd, E. & Maduro, M. F. Evolution of Developmental GATA Factors in Nematodes. J Dev Biology 8, 27 (2020).

51. Dineen, A., Nishimura, E. O., Goszczynski, B., Rothman, J. H. & McGhee, J. D. Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the C. elegans endoderm. Dev Biol 435, 150–161 (2018).

52. Elliott, S. L., Sturgeon, C. R., Travers, D. M. & Montgomery, M. C. Mode of bacterial pathogenesis determines phenotype in elt-2 and elt-7 RNAi Caenorhabditis elegans. Dev Comp Immunol 35, 521–4 (2010).

53. Head, B. & Aballay, A. Recovery from an Acute Infection in C. elegans Requires the GATA Transcription Factor ELT-2. Plos Genet 10, e1004609 (2014).

54. Roh, H. C., Dimitrov, I., Deshmukh, K., Zhao, G., Warnhoff, K., Cabrera, D., Tsai, W. & Kornfeld, K. A modular system of DNA enhancer elements mediates tissue-specific activation of transcription by high dietary zinc in C. elegans. Nucleic Acids Res 43, 803–16 (2014).

55. Zárate-Potes, A., Yang, W., Pees, B., Schalkowski, R., Segler, P., Andresen, B., Haase, D., Nakad, R., Rosenstiel, P., Tetreau, G., Colletier, J.-P., Schulenburg, H. & Dierking, K. The C. elegans GATA transcription factor elt-2 mediates distinct transcriptional responses and opposite infection outcomes towards different Bacillus thuringiensis strains. Plos Pathog 16, e1008826 (2020).

56. Yang, W., Dierking, K., Rosenstiel, P. C. & Schulenburg, H. GATA transcription factor as a likely key regulator of the Caenorhabditis elegans innate immune response against gut pathogens. Zoology 119, 244–253 (2016).

57. Head, B. P., Olaitan, A. O. & Aballay, A. Role of GATA transcription factor ELT-2 and p38 MAPK PMK-1 in recovery from acute P. aeruginosa infection in C. elegans. Virulence 8, 261–274 (2016).

58. Su, L., Zhao, T., Li, H., Li, H., Su, X., Ba, X., Zhang, Y., Huang, B., Lu, J. & Li, X. ELT-2 promotes O-GlcNAc transferase OGT-1 expression to modulate Caenorhabditis elegans lifespan. J Cell Biochem 121, 4898–4907 (2020).

59. Murray, J. I. Systems biology of embryonic development: Prospects for a complete understanding of the Caenorhabditis elegans embryo. Wiley Interdiscip Rev Dev Biology 7, e314 (2018).

60. Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., Hacohen, N., Haniffa, M., Hemberg, M., Kim, S., Klenerman, P., Kriegstein, A., Lein, E., Linnarsson, S., Lundberg, E., Lundeberg, J., Majumder, P., Marioni, J. C., Merad, M., Mhlanga, M., Nawijn, M., Netea, M., Nolan, G., Pe'er, D., Phillipakis, A., Ponting, C. P., Quake, S., Reik, W., Rozenblatt-Rosen, O., Sanes, J., Satija, R., Schumacher, T. N., Shalek, A., Shapiro, E., Sharma, P., Shin, J. W., Stegle, O., Stratton, M., Stubbington, M. J. T., Theis, F. J., Uhlen, M., Oudenaarden, A. van, Wagner, A., Watt, F., Weissman, J., Wold, B., Xavier, R., Yosef, N. & Participants, H. C. A. M. The Human Cell Atlas. Elife 6, e27041 (2017).

61. Ko, M. S. H. Embryogenomics: developmental biology meets genomics. Trends Biotechnol 19, 511–518 (2001).

62. Rau, A., Marot, G. & Jaffrézic, F. Differential meta-analysis of RNA-seq data from multiple studies. Bmc Bioinformatics 15, 91 (2014).

63. McIntyre, L. M., Lopiano, K. K., Morse, A. M., Amin, V., Oberg, A. L., Young, L. J. & Nuzhdin, S. V. RNA-seq: technical variability and sampling. Bmc Genomics 12, 293 (2011).

64. Kloet, F. M. van der, Buurmans, J., Jonker, M. J., Smilde, A. K. & Westerhuis, J. A. Increased comparability between RNA-Seq and microarray data by utilization of gene sets. Plos Comput Biol 16, e1008295 (2020).

65. Thompson, J. A., Tan, J. & Greene, C. S. Cross-platform normalization of microarray and RNA-seq data for machine learning applications. Peerj 4, e1621 (2016).

66. Fu, X., Fu, N., Guo, S., Yan, Z., Xu, Y., Hu, H., Menzel, C., Chen, W., Li, Y., Zeng, R. & Khaitovich, P. Estimating accuracy of RNA-Seq and microarrays with proteomics. Bmc Genomics 10, 161 (2009).

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

68. 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. The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. Dev Biol 302, 627–645 (2007).

69. Riddle, M. R., Weintraub, A., Nguyen, K. C. Q., Hall, D. H. & Rothman, J. H. Transdifferentiation and remodeling of post-embryonic C. elegans cells by a single transcription factor. Dev Camb Engl 140, 4844–9 (2013).

70. Du, L., Tracy, S. & Rifkin, S. A. Mutagenesis of GATA motifs controlling the endoderm regulator elt-2 reveals distinct dominant and secondary cis-regulatory elements. Dev Biol 412, 160–70 (2016).

71. Riddle, M. R., Spickard, E. A., Jevince, A., Nguyen, K. C. Q., Hall, D. H., Joshi, P. M. & Rothman, J. H. Transorganogenesis and transdifferentiation in C. elegans are dependent on differentiated cell identity. Dev Biol 420, 136–147 (2016).

72. Kudron, M. M., Victorsen, A., Gevirtzman, L., Hillier, L. W., Fisher, W. W., Vafeados, D., Kirkey, M., Hammonds, A. S., Gersch, J., Ammouri, H., Wall, M. L., Moran, J., Steffen, D., Szynkarek, M., Seabrook-Sturgis, S., Jameel, N., Kadaba, M., Patton, J., Terrell, R., Corson, M., Durham, T. J., Park, S., Samanta, S., Han, M., Xu, J., Yan, K.-K., Celniker, S. E., White, K. P., Ma, L., Gerstein, M., Reinke, V. & Waterston, R. The modERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis elegans Transcription Factors. Genetics 208, genetics.300657.2017 (2017).

73. Carr, A. & Biggin, M. D. A comparison of in vivo and in vitro DNA-binding specificities suggests a new model for homeoprotein DNA binding in Drosophila embryos. Embo J 18, 1598–1608 (1999).

74. Liu, X., Lee, C.-K., Granek, J. A., Clarke, N. D. & Lieb, J. D. Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. Genome Res 16, 1517–1528 (2006).

75. Yang, A., Zhu, Z., Kapranov, P., McKeon, F., Church, G. M., Gingeras, T. R. & Struhl, K. Relationships between p63 Binding, DNA Sequence, Transcription Activity, and Biological Function in Human Cells. Mol Cell 24, 593–602 (2006).

76. Fukushige, T., Hendzel, M. J., Bazett-Jones, D. P. & McGhee, J. D. Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. Proc National Acad Sci 96, 11883–11888 (1999).

77. Davidson, E. H., McClay, D. R. & Hood, L. Regulatory gene networks and the properties of the developmental process. Proc National Acad Sci 100, 1475–1480 (2003).

78. Levine, M. & Davidson, E. H. Gene regulatory networks for development. P Natl Acad Sci Usa 102, 4936–4942 (2005).

79. Alon, U. Network motifs: theory and experimental approaches. Nat Rev Genet 8, 450–461 (2007).

80. Delás, M. J. & Briscoe, J. Repressive interactions in gene regulatory networks: When you have no other choice. Curr Top Dev Biol 139, 239–266 (2020).

81. Murray, J. I., Boyle, T. J., Preston, E., Vafeados, D., Mericle, B., Weisdepp, P., Zhao, Z., Bao, Z., Boeck, M. & Waterston, R. H. Multidimensional regulation of gene expression in the C. elegans embryo. Genome Res 22, 1282–1294 (2012).

82. Oestreich, K. J. & Weinmann, A. S. Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. Nat Rev Immunol 12, 799–804 (2012).

83. Lancaster, B. R. & McGhee, J. D. How affinity of the ELT-2 GATA factor binding to cis-acting regulatory sites controls C. elegans intestinal gene transcription. Dev Camb Engl dev.190330 (2020). doi:10.1242/dev.190330

84. Ucar, D., Beyer, A., Parthasarathy, S. & Workman, C. T. Predicting functionality of protein–DNA interactions by integrating diverse evidence. Bioinformatics 25, i137–i144 (2009).

85. Kaletsky, R., Yao, V., Williams, A., Runnels, A. M., Tadych, A., Zhou, S., Troyanskaya, O. G. & Murphy, C. T. Transcriptome analysis of adult Caenorhabditis elegans cells reveals tissue-specific gene and isoform expression. Plos Genet 14, e1007559 (2018).

86. Tintori, S. C., Osborne Nishimura, E., Golden, P., Lieb, J. D. & Goldstein, B. A Transcriptional Lineage of the Early C. elegans Embryo. Dev Cell 38, 430–444 (2016).

87. Hashimshony, T., Feder, M., Levin, M., Hall, B. K. & Yanai, I. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature 519, 219–222 (2015).

88. Spencer, W. C., Zeller, G., Watson, J. D., Henz, S. R., Watkins, K. L., McWhirter, R. D., Petersen, S., Sreedharan, V. T., Widmer, C., Jo, J., Reinke, V., Petrella, L., Strome, S., Stetina, S. E. V., Katz, M., Shaham, S., Rätsch, G. & Miller, D. M. A spatial and temporal map of C. elegans gene expression. Genome Res 21, 325–341 (2011).

89. Blazie, S. M., Babb, C., Wilky, H., Rawls, A., Park, J. G. & Mangone, M. Comparative RNA-Seq analysis reveals pervasive tissue-specific alternative polyadenylation in Caenorhabditis elegans intestine and muscles. Bmc Biol 13, 4 (2015).

90. Haenni, S., Ji, Z., Hoque, M., Rust, N., Sharpe, H., Eberhard, R., Browne, C., Hengartner, M. O., Mellor, J., Tian, B. & Furger, A. Analysis of C. elegans intestinal gene expression and polyadenylation by fluorescence-activated nuclei sorting and 3′-end-seq. Nucleic Acids Res 40, 6304–6318 (2012).

91. Parker, D. M., Winkenbach, L. P., Parker, A., Boyson, S. & Nishimura, E. O. Improved Methods for Single-Molecule Fluorescence In Situ Hybridization and Immunofluorescence in Caenorhabditis elegans Embryos. Curr Protoc 1, e299 (2021).

92. Tsanov, N., Samacoits, A., Chouaib, R., Traboulsi, A.-M., Gostan, T., Weber, C., Zimmer, C., Zibara, K., Walter, T., Peter, M., Bertrand, E. & Mueller, F. smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. Nucleic Acids Res 44, e165–e165 (2016).

93. Kennedy, B. P., Aamodt, E. J., Allen, F. L., Chung, M. A., Heschl, M. F. P. & McGhee, J. D. The Gut Esterase Gene (ges-1) From the Nematodes Caenmorhabditis elegans and Caenorhabditis briggsae. J Mol Biol 229, 890–908 (1993).

94. Tremblay, M., Sanchez-Ferras, O. & Bouchard, M. GATA transcription factors in development and disease. Development 145, dev164384 (2018).

95. Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T. M., Egelhofer, T. A., Wang, W., Kimura, H., Lieb, J. D. & Strome, S. The Histone H3K36 Methyltransferase MES-4 Acts Epigenetically to Transmit the Memory of Germline Gene Expression to Progeny. Plos Genet 6, e1001091 (2010).

96. Keith, S. A., Maddux, S. K., Zhong, Y., Chinchankar, M. N., Ferguson, A. A., Ghazi, A. & Fisher, A. L. Graded Proteasome Dysfunction in Caenorhabditis elegans Activates an Adaptive Response Involving the Conserved SKN-1 and ELT-2 Transcription Factors and the Autophagy-Lysosome Pathway. Plos Genet 12, e1005823 (2016).

97. Block, D. H. S., Twumasi-Boateng, K., Kang, H. S., Carlisle, J. A., Hanganu, A., Lai, T. Y.-J. & Shapira, M. The Developmental Intestinal Regulator ELT-2 Controls p38-Dependent Immune Responses in Adult C. elegans. Plos Genet 11, e1005265 (2015).

98. Bass, J. I. F., Pons, C., Kozlowski, L., Reece-Hoyes, J. S., Shrestha, S., Holdorf, A. D., Mori, A., Myers, C. L. & Walhout, A. J. A gene-centered C. elegans protein–DNA interaction network provides a framework for functional predictions. Mol Syst Biol 12, 884 (2016).

99. Yan, D., Wu, Z., Chisholm, A. D. & Jin, Y. The DLK-1 Kinase Promotes mRNA Stability and Local Translation in C. elegans Synapses and Axon Regeneration. Cell 138, 1005–1018 (2009).

100. Yang, Y., Liu, L., Naik, I., Braunstein, Z., Zhong, J. & Ren, B. Transcription Factor C/EBP Homologous Protein in Health and Diseases. Front Immunol 8, 1612 (2017).

101. McEwan, D. L., Feinbaum, R. L., Stroustrup, N., Haas, W., Conery, A. L., Anselmo, A., Sadreyev, R. & Ausubel, F. M. Tribbles ortholog NIPI-3 and bZIP transcription factor CEBP-1 regulate a Caenorhabditis elegans intestinal immune surveillance pathway. Bmc Biol 14, 105 (2016).

102. Wu, C., Karakuzu, O. & Garsin, D. A. Tribbles pseudokinase NIPI-3 regulates intestinal immunity in Caenorhabditis elegans by controlling SKN-1/Nrf activity. Cell Reports 36, 109529 (2021).

103. Sharrocks, A. D. The ETS-domain transcription factor family. Nat Rev Mol Cell Bio 2, 827–837 (2001).

104. Hart, A. H., Reventar, R. & Bernstein, A. Genetic analysis of ETS genes in C. elegans. Oncogene 19, 6400–6408 (2000).

105. Kaymak, E., Farley, B. M., Hay, S. A., Li, C., Ho, S., Hartman, D. J. & Ryder, S. P. Efficient generation of transgenic reporter strains and analysis of expression patterns in Caenorhabditis elegans using library MosSCI. Dev Dynam 245, 925–936 (2016).

106. Thyagarajan, B., Blaszczak, A. G., Chandler, K. J., Watts, J. L., Johnson, W. E. & Graves, B. J. ETS-4 Is a Transcriptional Regulator of Life Span in Caenorhabditis elegans. Plos Genet 6, e1001125 (2010).

107. Li, C., Hisamoto, N. & Matsumoto, K. Axon Regeneration Is Regulated by Ets–C/EBP Transcription Complexes Generated by Activation of the cAMP/Ca2+ Signaling Pathways. Plos Genet 11, e1005603 (2015).

108. Sakai, Y., Hanafusa, H., Pastuhov, S. I., Shimizu, T., Li, C., Hisamoto, N. & Matsumoto, K. TDP2 negatively regulates axon regeneration by inducing SUMOylation of an Ets transcription factor. Embo Rep 20, e47517 (2019).

109. Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A. M., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J. S., Hope, I. A., Tissenbaum, H. A., Mango, S. E. & Walhout, A. J. M. A Gene-Centered C. elegans Protein-DNA Interaction Network. Cell 125, 1193–1205 (2006).

110. O'Brien, D., Jones, L. M., Good, S., Miles, J., Vijayabaskar, M. S., Aston, R., Smith, C. E., Westhead, D. R. & Oosten-Hawle, P. van. A PQM-1-Mediated Response Triggers Transcellular Chaperone Signaling and Regulates Organismal Proteostasis. Cell Reports 23, 3905–3919 (2018).

111. MacNeil, L. T., Pons, C., Arda, H. E., Giese, G. E., Myers, C. L. & Walhout, A. J. M. Transcription Factor Activity Mapping of a Tissue-Specific In Vivo Gene Regulatory Network. Cell Syst 1, 152–162 (2015).

112. O'Brien, D., Jones, L. M., Good, S., Miles, J., Vijayabaskar, M. S., Aston, R., Smith, C. E., Westhead, D. R. & Oosten-Hawle, P. van. A PQM-1-Mediated Response Triggers Transcellular Chaperone Signaling and Regulates Organismal Proteostasis. Cell Reports 23, 3905–3919 (2018).

113. Shapira, M., Hamlin, B. J., Rong, J., Chen, K., Ronen, M. & Tan, M.-W. A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. Proc National Acad Sci 103, 14086–14091 (2006).

114. Heimbucher, T., Hog, J., Gupta, P. & Murphy, C. T. PQM-1 controls hypoxic survival via regulation of lipid metabolism. Nat Commun 11, 4627 (2020).

115. Tepper, R. G., Ashraf, J., Kaletsky, R., Kleemann, G., Murphy, C. T. & Bussemaker, H. J. PQM-1 Complements DAF-16 as a Key Transcriptional Regulator of DAF-2-Mediated Development and Longevity. Cell 154, 676–690 (2013).

116. Dowen, R. H. CEH-60/PBX and UNC-62/MEIS Coordinate a Metabolic Switch that Supports Reproduction in C. elegans. Dev Cell 49, 235-250.e7 (2019).

117. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between Protein and mRNA Abundance in Yeast. Mol Cell Biol 19, 1720–1730 (1999).

118. Bauernfeind, A. L. & Babbitt, C. C. The predictive nature of transcript expression levels on protein expression in adult human brain. Bmc Genomics 18, 322 (2017).

119. Kimble, J. & Sharrock, W. J. Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans. Dev Biol 96, 189–196 (1983).

120. An, J. H. & Blackwell, T. K. SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. Gene Dev 17, 1882–1893 (2003).

121. Libina, N., Berman, J. R. & Kenyon, C. Tissue-Specific Activities of C. elegans DAF-16 in the Regulation of Lifespan. Cell 115, 489–502 (2003).

122. Martinez-Finley, E. J. & Aschner, M. Revelations from the Nematode Caenorhabditis elegans on the Complex Interplay of Metal Toxicological Mechanisms. J Toxicol 2011, 895236 (2011).

123. Ludewig, A. H. & Schroeder, F. C. Ascaroside signaling in C. elegans. Wormbook 1–22 (2013). doi:10.1895/wormbook.1.155.1

124. Chun, H., Sharma, A. K., Lee, J., Chan, J., Jia, S. & Kim, B.-E. The Intestinal Copper Exporter CUA-1 Is Required for Systemic Copper Homeostasis in Caenorhabditis elegans * ♦. J Biol Chem 292, 1–14 (2017).

125. Lee, K. & Mylonakis, E. An Intestine-Derived Neuropeptide Controls Avoidance Behavior in Caenorhabditis elegans. Cell Reports 20, 2501–2512 (2017).

126. Kalb, J. M., Beaster-Jones, L., Fernandez, A. P., Okkema, P. G., Goszczynski, B. & McGhee, J. D. Interference Between the PHA-4 and PEB-1 Transcription Factors in Formation of the Caenorhabditis elegans Pharynx. J Mol Biol 320, 697–704 (2002).

127. Fukushige, T., Goszczynski, B., Tian, H. & McGhee, J. D. The Evolutionary Duplication and Probable Demise of an Endodermal GATA Factor in Caenorhabditis elegans. Genetics 165, 575–588 (2003).

128. Fukushige, T., Goszczynski, B., Yan, J. & McGhee, J. D. Transcriptional control and patterning of the pho-1 gene, an essential acid phosphatase expressed in the C. elegans intestine. Dev Biol 279, 446–461 (2005).

129. Oskouian, B., Mendel, J., Shocron, E., Lee, M. A., Fyrst, H. & Saba, J. D. Regulation of Sphingosine-1-phosphate Lyase Gene Expression by Members of the GATA Family of Transcription Factors*. J Biol Chem 280, 18403–18410 (2005).

130. Malinow, R. A., Ying, P., Koorman, T., Boxem, M., Jin, Y. & Kim, K. W. Functional Dissection of C. elegans bZip-Protein CEBP-1 Reveals Novel Structural Motifs Required for Axon Regeneration and Nuclear Import. Front Cell Neurosci 13, 348 (2019).

131. Li, C., Hisamoto, N. & Matsumoto, K. Axon Regeneration Is Regulated by Ets–C/EBP Transcription Complexes Generated by Activation of the cAMP/Ca2+ Signaling Pathways. Plos Genet 11, e1005603 (2015).

132. Thyagarajan, B., Blaszczak, A. G., Chandler, K. J., Watts, J. L., Johnson, W. E. & Graves, B. J. ETS-4 Is a Transcriptional Regulator of Life Span in Caenorhabditis elegans. Plos Genet 6, e1001125 (2010).

133. Sakai, Y., Hanafusa, H., Pastuhov, S. I., Shimizu, T., Li, C., Hisamoto, N. & Matsumoto, K. TDP2 negatively regulates axon regeneration by inducing SUMOylation of an Ets transcription factor. Embo Rep 20, e47517 (2019).

134. Block, D. H. & Shapira, M. GATA transcription factors as tissue-specific master regulators for induced responses. Worm 4, e1118607 (2015).

135. Hawkins, M. G. & McGhee, J. D. elt-2, a Second GATA Factor from the Nematode Caenorhabditis elegans ∗. J Biol Chem 270, 14666–14671 (1995).

136. Bresnick, E. H., Lee, H.-Y., Fujiwara, T., Johnson, K. D. & Keles, S. GATA switches as developmental drivers. J Biological Chem 285, 31087–93 (2010).

137. Zheng, R. & Blobel, G. A. GATA Transcription Factors and Cancer. Genes Cancer 1, 1178–1188 (2010).

138. Fujiwara, T. GATA Transcription Factors: Basic Principles and Related Human Disorders. Tohoku J Exp Medicine 242, 83–91 (2017).

139. Gould, D. J. & Chernajovsky, Y. Endogenous GATA Factors Bind the Core Sequence of the tetO and Influence Gene Regulation with the Tetracycline System. Mol Ther 10, 127–138 (2004).

140. Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J. & Niwa, H. Differentiation of embryonic stem cells is induced by GATA factors. Gene Dev 16, 784–789 (2002).

141. Chou, J., Provot, S. & Werb, Z. GATA3 in development and cancer differentiation: Cells GATA have it! J Cell Physiol 222, 42–49 (2010).

142. Blobel, G. A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S. H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. Proc National Acad Sci 95, 2061–2066 (1998).

143. Hong, W., Nakazawa, M., Chen, Y., Kori, R., Vakoc, C. R., Rakowski, C. & Blobel, G. A. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. Embo J 24, 2367–2378 (2005).

144. Burda, P., Laslo, P. & Stopka, T. The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. Leukemia 24, 1249–1257 (2010).

145. Rodríguez-López, M., Gonzalez, S., Hillson, O., Tunnacliffe, E., Codlin, S., Tallada, V. A., Bähler, J. & Rallis, C. The GATA Transcription Factor Gaf1 Represses tRNAs, Inhibits Growth, and Extends Chronological Lifespan Downstream of Fission Yeast TORC1. Cell Reports 30, 3240-3249.e4 (2020).

146. Stiernagle, T. Maintenance of C. elegans. Wormbook 1–11 (2006). doi:10.1895/wormbook.1.101.1

147. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890 (2018).

148. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37, 907–915 (2019).

149. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014).

150. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550 (2014).

151. Team, R. C. R: A Language and Environment for Statistical Computing. (2021). at <https://www.R-project.org/>

152. Morgan, M. BiocManager: Access the Bioconductor Project Package Repository. (R package version 1.30.16, 2021). at <https://CRAN.R-project.org/package=BiocManager>

153. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. & Yutani, H. Welcome to the Tidyverse. Journal of Open Source Software 4, 1686 (2019).

154. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849 (2016).

155. Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F. & Manke, T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–W165 (2016).

156. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. (2021).

157. Kamath, R. S. & Ahringer, J. Genome-wide RNAi screening in Caenorhabditis elegans. Methods 30, 313–321 (2003).

158. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682 (2012).

159. Preibisch, S., Saalfeld, S. & Tomancak, P. Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics 25, 1463–1465 (2009).

160. Djira, G., Hasler, M., Gerhard, D. & Schaarschmidt, L. S. and F. mratios: Ratios of Coefficients in the General Linear Model. R package version 1.4.2. (2020). at <https://CRAN.R-project.org/package=mratios>

161. Park, P. J. ChIP–seq: advantages and challenges of a maturing technology. Nat Rev Genet 10, 669–680 (2009).

162. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat Protoc 13, 1006–1019 (2018).

163. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 6, e21856 (2017).

164. Kaya-Okur, H. S., Wu, S. J., Codomo, C. A., Pledger, E. S., Bryson, T. D., Henikoff, J. G., Ahmad, K. & Henikoff, S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10, 1930 (2019).

165. Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R. H. & Murray, J. I. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science 365, (2019).

166. Kim, K. W., Thakur, N., Piggott, C. A., Omi, S., Polanowska, J., Jin, Y. & Pujol, N. Coordinated inhibition of C/EBP by Tribbles in multiple tissues is essential for Caenorhabditis elegans development. Bmc Biol 14, 104 (2016).

167. Irazoqui, J. E., Urbach, J. M. & Ausubel, F. M. Evolution of host innate defence: insights from Caenorhabditis elegans and primitive invertebrates. Nat Rev Immunol 10, 47–58 (2010).

168. Kim, D. H. & Ausubel, F. M. Evolutionary perspectives on innate immunity from the study of Caenorhabditis elegans. Curr Opin Immunol 17, 4–10 (2005).

169. Breitkopf, S. B. & Asara, J. M. Determining In Vivo Phosphorylation Sites Using Mass Spectrometry. Curr Protoc Mol Biology 98, 18.19.1-18.19.27 (2012).

170. Lee, C.-R., Park, Y.-H., Min, H., Kim, Y.-R. & Seok, Y.-J. Determination of protein phosphorylation by polyacrylamide gel electrophoresis. J Microbiol 57, 93–100 (2019).

171. O'Donoghue, L. & Smolenski, A. Analysis of protein phosphorylation using Phos-tag gels. J Proteomics 259, 104558 (2022).

172. Gerstein, M. B., Lu, Z. J., Nostrand, E. L. V., Cheng, C., Arshinoff, B. I., Liu, T., Yip, K. Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., Alves, P., Chateigner, A., Perry, M., Morris, M., Auerbach, R. K., Feng, X., Leng, J., Vielle, A., Niu, W., Rhrissorrakrai, K., Agarwal, A., Alexander, R. P., Barber, G., Brdlik, C. M., Brennan, J., Brouillet, J. J., Carr, A., Cheung, M.-S., Clawson, H., Contrino, S., Dannenberg, L. O., Dernburg, A. F., Desai, A., Dick, L., Dosé, A. C., Du, J., Egelhofer, T., Ercan, S., Euskirchen, G., Ewing, B., Feingold, E. A., Gassmann, R., Good, P. J., Green, P., Gullier, F., Gutwein, M., Guyer, M. S., Habegger, L., Han, T., Henikoff, J. G., Henz, S. R., Hinrichs, A., Holster, H., Hyman, T., Iniguez, A. L., Janette, J., Jensen, M., Kato, M., Kent, W. J., Kephart, E., Khivansara, V., Khurana, E., Kim, J. K., Kolasinska-Zwierz, P., Lai, E. C., Latorre, I., Leahey, A., Lewis, S., Lloyd, P., Lochovsky, L., Lowdon, R. F., Lubling, Y., Lyne, R., MacCoss, M., Mackowiak, S. D., Mangone, M., McKay, S., Mecenas, D., Merrihew, G., III, D. M. M., Muroyama, A., Murray, J. I., Ooi, S.-L., Pham, H., Phippen, T., Preston, E. A., Rajewsky, N., Rätsch, G., Rosenbaum, H., Rozowsky, J., Rutherford, K., Ruzanov, P., Sarov, M., Sasidharan, R., Sboner, A., Scheid, P., Segal, E., Shin, H., Shou, C., Slack, F. J., Slightam, C., Smith, R., Spencer, W. C., Stinson, E. O., Taing, S., Takasaki, T., Vafeados, D., Voronina, K., Wang, G., Washington, N. L., Whittle, C. M., Wu, B., Yan, K.-K., Zeller, G., Zha, Z., Zhong, M., Zhou, X., Consortium, modENCODE, Ahringer, J., Strome, S., Gunsalus, K. C., Micklem, G., Liu, X. S., Reinke, V., Kim, S. K., Hillier, L. W., Henikoff, S., Piano, F., Snyder, M., Stein, L., Lieb, J. D. & Waterston, R. H. Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. Science 330, 1775–1787 (2010).

173. Ummethum, H. & Hamperl, S. Proximity Labeling Techniques to Study Chromatin. Frontiers Genetics 11, 450 (2020).

174. Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M. & Zaret, K. S. Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. Mol Cell 9, 279–289 (2002).

175. Bossard, P. & Zaret, K. S. GATA transcription factors as potentiators of gut endoderm differentiation. Development 125, 4909–4917 (1998).

176. Sommermann, E. M., Strohmaier, K. R., Maduro, M. F. & Rothman, J. H. Endoderm development in Caenorhabditis elegans: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. Dev Biol 347, 154–166 (2010).

177. Rosenfeld, N., Elowitz, M. B. & Alon, U. Negative Autoregulation Speeds the Response Times of Transcription Networks. J Mol Biol 323, 785–793 (2002).

178. Nevozhay, D., Adams, R. M., Murphy, K. F., Josić, K. & Balázsi, G. Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. Proc National Acad Sci 106, 5123–5128 (2009).

179. Hinczewski, M. & Thirumalai, D. Noise Control in Gene Regulatory Networks with Negative Feedback. J Phys Chem B 120, 6166–6177 (2016).

180. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., Rogers, A. J., McElrath, J. M., Blish, C. A., Gottardo, R., Smibert, P. & Satija, R. Integrated analysis of multimodal single-cell data. Cell 184, 3573-3587.e29 (2021).

APPENDIX A

SINGLE-CELL RNA SEQUENCING OF FACS ISOLATED EMBRYO INTESTINE CELLS

A.1 Summary

As demonstrated in our bulk intestine transcriptome data, it appears that the embryo transcriptome is distinct from larval stages. To explore the dynamics of the embryo intestine transcriptome during embryonic development, we performed single cell RNA-seq (scRNA-seq) on FACS isolated embryo intestine cells. The application of scRNA-seq provides an opportunity to assay the complete transcriptome of thousands of individual cells providing an opportunity to resolve distinct cell subtypes within a population of cells [14,15]. Previous work profiling single cell transcriptomes for the C. elegans embryo do not currently provide a complete picture of embryonic intestine development [86,165].

Embryo intestine cells were prepared as previously described (Chapter 2). 60,000 mixed stage embryo intestine cells were captured and scRNA-seq libraries were prepared with the 10x Genomics Chromium Controller with the Chromium Single Cell 3' Reagent Kit (v3.1 Chemistry, Dual Index). Libraries were sequenced on the Illumina NextSeq platform to an average read depth of ~100 million reads per library. Two biological replicates of scRNA-seq libraries for FACS isolated embryo stage intestine cells were generated. Raw sequencing data was analyzed with the *cellranger* pipeline (v 7.0.0, 10x Genomics). Dimensionality reduction and visualization was performed with the R package Seurat (v 4.0) [180].

Analysis of the scRNA-seq libraries generated on FACS isolated embryo stage intestine cells, we identified 6 clusters of intestine cell populations with distinct transcriptome profiles (Figure A.1A), demonstrating additional complexities in the embryonic intestine GRN. Through computational analysis, we identified intestine TF and key intestine marker genes distinct to

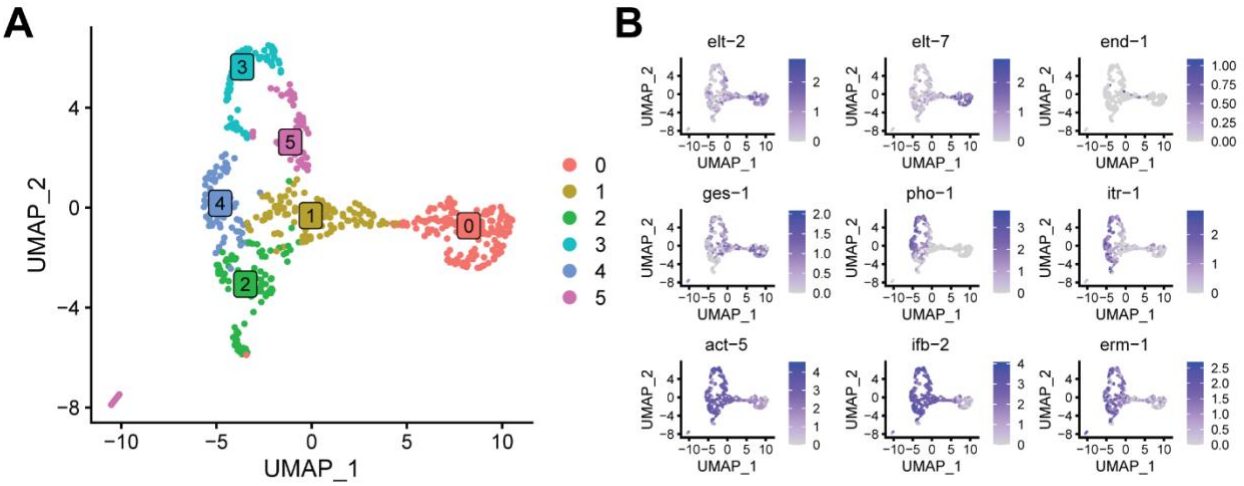

Figure A.1: UMAP projection of embryonic *C. elegans* intestine cells. (A) Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) projection of our embryo intestine scRNA-seq dataset. Each point on the plot represents a cell transcriptome. Cells are color coded and numbered based on assigned transcriptome similarity clusters defined by differential expression analysis with the Seurat package. (B) Visualization of key TFs (top row) and intestine marker genes (bottom two rows) in the UMAP projection shown in (A). Scales to the right of the plot depict the normalized transcript abundance for a given cell on the plot (blue, high transcript abundance; white, low transcript abundance).

each cluster, thereby providing an opportunity to investigate key TFs for the embryonic intestine development (Figure A.2B). An outstanding question we have is if these clusters represent different stages of intestine development or different sub-regions of intestine tissue. Together, this dataset will serve as a rich dataset for future work dissecting the *C. elegans* intestine GRN.

APPENDIX B

PROCEDURES FOR DISSOCIATION AND FACS ISOLATION OF EMBRYONIC AND POST-EMBRYONIC *C. ELEGANS* INTESTINE CELLS FOR RNA-SEQ ANALYSIS[2]

B.1 Summary

Included in this appendix is a collection of protocols for the isolation of *C. elegans* intestine cells through FACS from embryo, L1, and L3 stage. The protocols are focused on collecting material for RNA-seq analysis but could be utilized for other biochemical assays. Troubleshooting suggestions are listed at key steps of the protocols. A single experiment cycle takes approximately two weeks to complete. To perform this experiment, perform these protocols in the following order:

- B.2 Synchronized *C. elegans* culture on NGM plates for FACS isolation of intestine cells
- One of the following:
    - B.3 Embryo stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells
    - B.4 L1 stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells
    - B.5 L3 stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells
- B.6 FACS isolation of intestine-specific *C. elegans* cells

[2] An interactive version of the protocols included in this chapter are published on the protocol.io platform.

Robert TP Williams, Erin Osborne Nishimura 2022. Protocol collection: Dissociation and FACS isolation of embryonic and post-embryonic C. elegans intestine cells for RNA-seq analysis. **protocols.io** https://dx.doi.org/10.17504/protocols.io.5jyl895jdv2w/v1

B.2 Synchronized C. elegans culture on NGM plates for FACS isolation of intestine cells

B.2.1 Abstract

This protocol details the steps necessary to scale-up and synchronize *C. elegans* cultures for FACS isolation of intestine cells. We cultured worms with agar-based NGM plates to reduce any confounding effects that may be introduced by large scale liquid culture. This protocol utilizes two rounds of mixed stage culture growth followed by two rounds of synchronized growth. After scale-up and synchronization, this protocol provides details for culture conditions necessary for intestine FACS of embryo, L1 or L3 stage experiments.

B.2.2 Materials

Strains:

1. OP50 E. coli
2. FACS control *C. elegans* strain, i.e. N2
3. FACS sorting *C. elegans* strain, i.e. JM149 *cals71*[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]

Reagents:

1. LB Broth Mix (Genesee 11-120)
2. M9 buffer
    a. 3 g KH_2PO_4 (Sigma-Aldrich P0662)
    b. 6 g Na_2HPO_4 (Thermo Fisher S373)
    c. 5 g NaCl (Sigma-Aldrich S9888)
    d. 1 ml 1 M $MgSO_4$ (Sigma-Aldrich 208094)
    e. H_2O to 1 liter
3. NGM plates
    a. Complete NGM protocol is available at the following URL: http://www.wormbook.org/chapters/www_strainmaintain/strainmaintain.html#d0e214

b. 3g NaCl (Sigma-Aldrich S9888)

c. 17g agar (Genesee 20-249)

d. 2.5g peptone (VWR 89406-350)

e. 975g H_2O (sterile and deionized)

f. 1ml 1M $CaCl_2$ (Sigma-Aldrich C3306)

g. 1ml 5mg/ml cholesterol (Fisher 501848291)

h. 1ml 1M $MgSO_4$ (Sigma-Aldrich 208094)

i. 25ml 1M KPO_4 Buffer pH 6.0 (108.3 g KH_2PO_4, 35.6 g K_2HPO_4, H_2O to 1 liter) (Sigma-Aldrich P0662, P3786

4. Peptone enriched NGM: in recipe above use 20g peptone instead of 2.5g peptone
5. Bleaching solution
    a. Sodium Hypochlorite Solution, 6% available chlorine (Ricca Chemical, 7495.7-32)
    b. 5N NaOH (Fisher S318-100)

Consumables:

1. 150 mm petri dishes "large plates" (Corning 351058)
2. 15 ml centrifuge tubes (Peak PS-695)

Equipment:

1. Swinging bucket rotor refrigerated centrifuge (Eppendorf 5810R)
2. Pipet-Aid (VWR 89166-464)
3. 20°C incubator (Caron 7001-28-1)

B.2.3 Protocol Steps

B.2.3.1 Prepare OP50 seeded NGM plates

1. Using sterile technique, pick an OP50 colony and inoculate a 250 ml bottle of sterile LB.
2. Incubate OP50 liquid culture at 37°C overnight

3. Make 3 liters of Nematode Growth Media (NGM) with 150 mm petri dish **(hereafter referred to as "large plates")**.

   NOTE: For FACS isolation of post-embryonic worm stages, prepare an additional 1 liter of peptone enriched NGM media. For peptone enriched NGM media, replace the normal 2.5 g peptone mass with 20 g.
4. Pour molten NGM into large plates. Each liter should make 20 plates, for a total of 60 large plates
5. Allow plates to dry overnight
6. Seed each NGM plate with 3 ml OP50 liquid culture.
7. Cover as much agar surface as possible by moving the plate in first a circular pattern, then a figure 8 pattern
8. Dry the OP50 seeded plates at room temperature with the lids on for three to four days until there is no more excess liquid

B.2.3.2 Grow mixed stage cultures of cell sorting strain

9. Identify a 60 mm petri plate culture of the sorting strain that has recently exhausted the E. coli lawn
10. Chunk the plate into 5 equal pieces
11. Transfer each chunk to a fresh large NGM OP50 plate with the worm covered surface facing down
12. Place sorting strain cultures in 20°C incubator for 72-96 hours, until the E. coli lawn is exhausted

B.2.3.3 Expand mixed stage cultures of cell sorting strain

13. Harvest the mixed stage worm population from the 5 plates by washing each plate with ~10 ml of M9
14. Transfer the worm suspension to a 15 ml conical centrifuge tube
15. Pellet the worms by centrifuging for 1 min at 2,000 rcf

16. Wash with additional M9 by aspirating the supernatant and resuspending the worms in fresh M9 to a total volume of 15 ml
17. Repeat the M9 wash until the supernatant is clear
18. Measure the approximate concentration of worms in suspension such that the optimal density of worms are seeded onto the plate
    a. Shake or vortex the tube to ensure the worms are evenly distributed in the suspension
    b. Aspirate 2 ul of worm suspension with a p10 pipette. Pipette the worm suspension up and down at least four times before moving on.
    c. Dispense the worm suspension on a clean microscope slide
    d. With a cell counter, count the number of worms on the slide under a dissection microscope. Dilute the worm suspension if there are too many to count.
    e. Determine the concentration and total number of worms

$$\frac{\text{worms in 2ul drop}}{\text{2ul}} \approx \frac{\text{worms}}{\text{ul}}$$

$$\frac{\text{worms}}{\text{ul}} \times \text{worm suspension ul} \approx \text{total \# of worms}$$

19. Seed 20 fresh large NGM OP50 plates with 5,000 worms per plate

    NOTE: Optimal seeding density may need to be independently determined as lab conditions vary. Optimal worm density should allow for worms to consume the bacterial lawn in time to become gravid without inducing a food deprivation stress response. To determine the optimal seeding density, seed a range of worms on several plates and culture until gravid.
20. Incubate for 72 hours in a 20°C incubator, until there is a large number of gravid adults

B.2.3.4 First embryo synchronization with hypochlorite solution

21. At the beginning for the day, chunk one recently starved 60 mm N2 plate to a fresh large NGM OP50 plate. This is key step and will serve as the negative GFP control for cell sorting.
22. Harvest mixed stage sorting strain worms from all 20 plates by washing each plate individually with ~10 ml of M9
    a. Transfer mixed stage worm suspension to a 15 ml conical centrifuge tube
    b. Pellet the worms by centrifuging for 1 minute at 2,000 rcf
    c. Aspirate the supernatant
    d. Harvest worms from another plate as outline above
    e. Resuspend the worm pellet with worm suspension from the newly washed plate
    f. Repeat this process until worms have been harvested from all 20 plates into a single 15ml tube
23. Once all plates have been harvested, continue washing the worm pellet with fresh M9 to remove excess E. coli by pelleting and resuspending in fresh M9. The final worm pellet yield should be 1 to 2 ml.
24. Once the worm suspension is free of E. coli, centrifuge again and remove all M9 supernatant from the worm pellet
25. Resuspend the worm pellet in 8 ml of H2O
26. Add 0.9 ml of Sodium Hypochlorite Solution (Ricca Chemical, 7495.7-32) and 1.44 ml of 5N NaOH to the worm suspension
27. Resuspend the worm pellet with brief vigorous vortexing
28. Incubate at room temperature for 6 to 8 minutes. While incubating shake the tube or place on a nutator. The time to bleach the worms depends on the worm pellet volume, with larger worm pellets taking longer. Do not incubate for longer than 8 minutes.
29. Monitor the progression of the hypochlorite treatment. Larval worms should dissolve, adult worms will begin to break at the vulva and release embryos. I typically monitor the

treatment by looking through the tube under a dissection microscope. Aliquots of the worm suspension can also be taken throughout the process and viewed on a microscope slide.

30. Once the worms are sufficiently dissolved, centrifuge the tube for 30 seconds at 2,000 rcf
31. Decant the supernatant and wash the embryo pellet by adding 15 ml of M9 to quench the hypochlorite treatment
32. Wash the bleached embryos a second time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
33. Wash the bleached embryos a third time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
34. Wash the bleached embryos a fourth time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
35. Measure the approximate concentration of embryos in suspension
    a. Shake or vortex the tube to ensure the embryos are evenly distributed in the suspension
    b. Aspirate 2 ul of embryo suspension with a p10 pipette. Pipette the embryo suspension up and down at least four times before moving on.
    c. Dispense the embryo suspension on a clean microscope slide
    d. With a cell counter, count the number of embryos on the slide under a dissection microscope. Dilute the embryo suspension if there are too many to count. See step 18.5 for the formula to determine the embryo concentration.

36. Seed 20 large NGM/OP50 plates with 5,000 embryos. Incubate at 20°C for approximately 72 hours until worms are gravid.

B.2.3.5 Second embryo synchronization with hypochlorite solution

37. Harvest both synchronized sorting strain worms and mixed stage N2 worms in parallel by washing individual plates with ~10 ml of M9 and collecting in two separate 15 ml tubes
38. Transfer the worm suspension to a 15 ml conical centrifuge tube
39. Pellet the worms by centrifuging for 1 minute at 2,000 rcf
40. Discard the supernatant
41. Resuspend the worm pellet with worm suspension from another large plate
42. Repeat this process until worms have been harvested from all 20 plates for the sorting strain and 1 N2 plate
43. Once all plates have been harvested, continue washing the worm pellet with fresh M9 to remove excess E. coli. The final worm pellet yield for the sorting strain should be 1 to 2 ml.
44. Once the worm suspension is free of E. coli, centrifuge again and remove all M9 supernatant from the worm pellet
45. Resuspend the worm pellet in 8 ml of H2O
46. Add 0.9 ml of Sodium Hypochlorite Solution (Ricca Chemical, 7495.7-32) and 1.44 ml of 5N NaOH to the worm suspension
47. Resuspend the worm pellet with brief vigorous vortexing
48. Incubate at room temperature for 6 to 8 minutes. While incubating shake the tube or place on a nutator. The time to bleach the worms depends on the worm pellet volume, with larger worm pellets taking longer. Do not incubate for longer than 8 minutes.
49. Monitor the progression of the hypochlorite treatment. Larval worms should dissolve, adult worms will begin to break at the vulva and release embryos. I typically monitor the

treatment by looking through the tube under a dissection microscope. Aliquots of the worm suspension can also be taken throughout the process and viewed on a microscope slide.

50. Once the worms are sufficiently dissolved, centrifuge the tube for 30 seconds at 2,000 rcf
51. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9 to quench the hypochlorite treatment
52. Wash the bleached embryos a second time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
53. Wash the bleached embryos a third time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
54. Wash the bleached embryos a fourth time. Centrifuge the tube for 30 seconds at 2,000 rcf to pellet the embryo suspension. Decant the supernatant and resuspend the embryo pellet in 15 ml of M9.
55. The final embryo yield should be approximately 0.01 ml for the wildtype N2 stain and 0.2ml for the fluorescent sorting strain.
56. For embryo stage FACS experiments, move on to the embryo dissociation protocol B.3
57. For post-embryonic stage FACS experiments, move on to Step 58 for L1 stage experiments or Step 65 for L3 stage experiments

B.2.3.5 L1 Culture

58. Incubate the synchronized embryos in 15 ml M9 overnight rotating in 20°C incubator for 24 hours
59. Feed the synchronized L1 sorting strain worms six hours before beginning the L1 dissociation protocol

NOTE: This step is necessary to reduce any observable starvation-induced responses in the measured transcriptional data. Negative control N2 worms can remain incubating in M9 suspension.

60. Pellet the L1 sorting strain worms and resuspend in 500 ul of M9. Transfer equal volumes of the synchronized L1 sorting strain worms onto two large peptone enriched NGM OP50 plates
61. Feed the synchronized L1 worms for 6 hours in a 20°C incubator.
62. Harvest the synchronized L1 worms by washing the plates with fresh M9. Pellet the worms for 1 min at 2,000 rcf. Discard the supernatant and repeat M9 washes until the supernatant is free of visible E. coli.
63. For both fluorescent sorting strain and wildtype strain, pass the harvested L1 suspension through a 20 micron filter. This will filter any contaminating debris (agar chunks, partially bleached worm chunks) and any unhatched or dead embryos.
64. Move on to the L1 stage dissociation protocol B.4

B.2.3.6 L3 Culture

65. Synchronize the embryos to the L1 stage by incubating the embryos for 24 hours suspended in M9, rotating in a 20°C incubator
66. Measure the approximate concentration of L1 worms in suspension
    a. Shake or vortex the tube to ensure the worms are evenly distributed in the suspension
    b. Aspirate 2 ul of worm suspension with a p10 pipette. Pipette the worm suspension up and down at least four times before moving on.
    c. Dispense the worm suspension on a clean microscope slide
    d. With a cell counter, count the number of worms on the slide under a dissection microscope. Dilute the embryo suspension if there are too many to count.

67. Seed the plates with a worm suspension volume between 100 and 500 ul. Concentrate or dilute the worm suspension accordingly. To avoid generating stressed clumps of developmentally delayed worms, spot the worm suspension evenly across the OP50 lawn.
    a. For the sorting strain worms, seed at least 10 large peptone enriched NGM OP50 plates with 30,000 L1 worms.
    b. For the wildtype worms, seed at least 5 large peptone enriched NGM OP50 plates with 30,000 L1 worms
68. Incubate the worms for 48 hours in a 20C incubator until worms visibly reach the L3 stage.
69. Move on to the L3 stage dissociation protocol B.5

B.3 Embryo stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells

B.3.1 Abstract

This protocol is for generating a single cell suspension suitable for isolation of intestine-specific cells through Fluorescence Activated Cell Sorting (FACS) from embryo stage *C. elegans*. This protocol utilizes treatment with Chitinase and Pronase E to disrupt the cuticle. Embryos are mechanically homogenized with 21G syringe needle.

B.3.2 Materials

Strains:

1. FACS control *C. elegans* strain, i.e. N2
2. FACS sorting *C. elegans* strain, i.e. JM149 *cals71*[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]

Reagents:

1. L15-10 solution

a. 500 ml Leibovitz's L-15 Medium (Thermo 21083027)
b. 50 ml Fetal Bovine Serum (heat inactivated) (Thermo 10438026)
c. 5 ml 100X Penicillin Streptomycin solution (Thermo 15140148)
d. 7.7 g Sucrose powder
e. Filter with 0.2 µm pore filter
f. Store at 4°C

2. Egg Buffer
    a. 2.95 ml of 2M NaCl
    b. 1.2 ml of 2M KCl
    c. 0.1 ml of 1M CaCl2
    d. 0.1 ml of 1M MgCl2
    e. 1.25 ml of 1M HEPES pH 7.2
3. Enzymes
    a. Chitinase from Streptomyces griseus (Sigma C6137-5UN)
        i. Resuspend 5 U of chitinase powder with 5 ml of egg buffer
        ii. Store 1 ml aliquots at -20°C
        iii. Final concentration 1 U/ml
    b. Pronase E, Protease from Streptomyces griseus (Sigma P8811-1G)
        i. Resuspend 150 mg Pronase E powder in 10 ml egg buffer
        ii. Nutate until powder is dissolved
        iii. Store 1 ml aliquots at -20°C
        iv. Final concentration 15 mg/ml

Consumables:

1. standard 1.5 ml tubes
2. Stericup 0.2 micron filter (Fisher S2GPU05RE)
3. 21 gauge 1 inch needle (fisher 14-826C)

4. 1 ml syringe (fisher 14-823-30)
5. 35-micron nylon mesh filter caps (Stellar Scientific FSC-FLTCP)
6. 5 ml sterile polypropylene round-bottom tube (STEMCELL Technologies 38057)
7. Bio-Rad TC20 automated cell counting slide (Bio-Rad 1450011)

Equipment:

1. Fixed angle rotor centrifuge (Eppendorf 5424)
2. Swinging bucket rotor refrigerated centrifuge (Eppendorf 5810R)
3. 15 ml tube and 1.5 ml tube adapter (Eppendorf 022638704, Eppendorf 022638742)
4. Fluorescent microscope
5. Nutating mixer
6. Bio-Rad TC20 automated cell counter

B.3.3 Protocol Steps

B.3.3.1 Before beginning

1. Prepare reagents in advance

   L15-10 Buffer: Mix 500 ml Leibovitz's L-15 Medium, 50 ml Fetal Bovine Serum (heat inactivated), 50 ul of 100x Penicillin-Streptomycin solution and 7.7 g sucrose. Filter with 0.2 micron pore filter. Store at 4ºC.

   Egg Buffer: Mix 29.5 ml of 2M NaCl, 12 ml of 2M KCl, 1 ml of 1M CaCl2, 1 ml of MgCl2, 12.5 ml of 1M HEPES-NaOH pH 7.2 and 435 ml molecular grade water. Filter with 0.2 micron pore filter. Store at 4ºC.

   Chitinase solution (1 U/ml): Dissolve 5 units of Chitinase from Streptomyces griseus (Sigma C6137-5UN) in 5 ml of Egg Buffer. Nutate the solution for approximately 10 minutes until dissolved. Prepare 1 ml aliquots in 1.5 ml tubes. Store aliquots at -20ºC.

   Pronase E solution (15 mg/ml): Weigh 150 mg of Protease from Streptomyces griseus (Sigma P8811-1G) into a 15 ml tube. Dissolve the enzyme in 10 ml of Egg Buffer. Nutate

the solution for approximately 10 minutes until dissolved. Prepare 1 ml aliquots in 1.5 ml tubes. Store aliquots at -20ºC.

2. On day of protocol:

   Cool swinging bucket centrifuge to 4ºC

   Thaw Pronase and chitinase aliquots at room temperature

   Place L15-10 and egg buffer on ice

3. Starting material:

   Worm suspension in 15 ml tube (material generated from protocol B.2)

   Strains: N2, fluorescent sorting strain

   Perform this protocol on both strains in parallel

   NOTE: The volumes for enzymatic treatments in this protocol require an embryo pellet less than 200 ul. If embryo pellet exceeds 200 ul, utilize 2x the embryo pellet volume.

B.3.3.2 Chitinase Treatment

4. Centrifuge embryo suspension at 2000 rcf for 1 minute in swinging bucket centrifuge
5. Resuspend the embryo pellet in 1 ml of M9 and transfer to a 1.5 ml tube.
6. Pellet the embryos at 2000 rcf for 1 minute in a centrifuge
7. Aspirate and discard the supernatant
8. Transfer 10 ul of embryo pellet to 1ml of Qiazol and store at -80°C for downstream RNA analysis
9. Resuspend the embryo pellet from Step 7 in 0.5 ml egg buffer and 1 ml chitinase (1 U/ml)
10. Incubate for 20 min rotating/nutating at room temperature
11. Verify eggshell digestion by visualizing the chitinase treated embryos under a microscope. Early embryos should change shape, and pretzel stage embryos should release from their eggshell.
12. Pellet the embryos at 200 rcf for 5 min in fixed angle rotor centrifuge

13. Aspirate and discard the supernatant

B.3.3.3 Pronase treatment and dissociation

14. Resuspend the chitinase treated embryo pellet in 200 ul Pronase (15 mg/ml) and 500 ul egg buffer
15. Attach a 21 gauge 1¼ inch needle to a sterile 1 ml syringe
16. Disrupt the embryo vitelline membrane and release cells by passing the embryo suspension through the needle 100 times, generating a worm slurry
17. Visually confirm embryo dissociation by viewing a 2 ul sample of worm slurry on a fluorescent microscope
18. Quench the Pronase treatment by adding 800 ul of L15-10 media to worm slurry
19. Store the sample on ice until all strains are completed

B.3.3.4 Wash and harvest single cells

20. Wash away excess Pronase from the worm slurry
    a. Pellet the worm slurry at 500 rcf for 5 mins in swinging bucket centrifuge cooled to 4ºC
    b. Aspirate and discard the supernatant
    c. Resuspend the worm slurry in 1 ml of L15-10 media
    d. Pellet the worm slurry at 500 rcf for 5 mins in swinging bucket centrifuge cooled to 4ºC
    e. Aspirate and discard the supernatant
    f. Resuspend the worm slurry in 1 ml of L15-10 media
21. Harvest the cells
    a. Pellet undissociated embryos at 100 rcf for 1 minute in swinging bucket centrifuge cooled to 4ºC

       NOTES:

       - This step will separate the dissociated cells from intact embryos

- Cells will remain in the supernatant

- Ensure your cell type of interest is not lost during this step.

- Visually confirm fluorescent cells **are present** in the supernatant.

- Visually confirm fluorescent cells **are not present** in the pellet.

- You may need to reduce the centrifuge speed and/or time if fluorescent cells are in the pellet of this step.

b. Aspirate 1 ml of the cell-containing supernatant. Keep the pipette away from the pelleted worm debris.

c. Dispense the cell suspension though a 35-micron nylon mesh filter into a 5 ml flow cytometry tube

d. Pellet undissociated embryos at 100 rcf for 1 minute in swinging bucket centrifuge cooled to 4ºC

22. Perform an additional round of cell harvest for the sorting strain only (Step 21)

    Total cell suspension volumes:

    - Control strain = 1ml

    - Sorting strain = 2ml

23. - Transfer 70 ul of cells to 1 ml of Qiazol and store at -80°C for downstream RNA analysis

    - Continue to step 24

    - Retain the remaining ~2ml of cells for FACS protocol B.6

B.3.3.4 Measure approximate cell concentration

24. Load 10 ul of cell suspension to a Bio-Rad TC20 automated cell counting slide

25. This protocol should yield between $2x10^6$ to $4x10^6$ total cells

26. Dilute the sample to $1x10^6$ cells/ml if above this concentration with L15-10

27. Microscopically confirm fluorescent cells are present in the cell suspension

28. Move on to FACS protocol B.6

B.4 L1 stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells

B.4.1 Abstract

This protocol is for generating a single cell suspension suitable for isolation of intestine-specific cells through Fluorescence Activated Cell Sorting (FACS) from L1 stage *C. elegans*. This protocol utilizes treatment with SDS-DTT solution and Pronase E to disrupt the cuticle. Worms are mechanically homogenized with a Dounce homogenizer.

B.4.2 Materials

Strains:

1. FACS control *C. elegans* strain, i.e. N2
2. FACS sorting *C. elegans* strain, i.e. JM149 *cals71*[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]

Reagents:

1. L15-10 solution
    a. 500 ml Leibovitz's L-15 Medium (Thermo 21083027)
    b. 50 ml Fetal Bovine Serum (heat inactivated) (Thermo 10438026)
    c. 5 ml 100X Penicillin Streptomycin solution (Thermo 15140148)
    d. 7.7 g Sucrose powder
    e. Filter with 0.2 µm pore filter
    f. Store at 4°C
2. Egg Buffer
    a. 2.95 ml of 2M NaCl
    b. 1.2 ml of 2M KCl
    c. 0.1 ml of 1M CaCl2
    d. 0.1 ml of 1M MgCl2

e. 1.25 ml of 1M HEPES pH 7.2

3. Enzymes
    a. Pronase E, Protease from Streptomyces griseus (Sigma P8811-1G)
        i. Resuspend 150 mg Pronase E powder in 10 ml egg buffer
        ii. Nutate until powder is dissolved
        iii. Store 1 ml aliquots at -20°C
        iv. Final concentration 15 mg/ml

Consumables:

1. standard 1.5 ml tubes
2. Stericup 0.2 micron filter (Fisher S2GPU05RE)
3. 20-micron mesh filter (Fisher Scientific NC1004201)
4. 35-micron nylon mesh filter caps (Stellar Scientific FSC-FLTCP)
5. 5 ml sterile polypropylene round-bottom tube (STEMCELL Technologies 38057)
6. Bio-Rad TC20 automated cell counting slide (Bio-Rad 1450011)

Equipment:

1. Fixed angle rotor centrifuge (Eppendorf 5424)
2. Swinging bucket rotor refrigerated centrifuge (Eppendorf 5810R)
3. 15 ml tube and 1.5 ml tube adapter (Eppendorf 022638704, Eppendorf 022638742)
4. Fluorescent microscope
5. Nutating mixer
6. Bio-Rad TC20 automated cell counter
7. 2ml Dounce homogenizer with pestle A (Sigma-Aldrich D8938)

B.4.3 Protocol Steps

B.4.3.1 Before beginning

1. Prepare reagents in advance

    L15-10 Buffer: Mix 500 ml Leibovitz's L-15 Medium, 50 ml Fetal Bovine Serum (heat

inactivated), 50 ul of 100x Penicillin-Streptomycin solution and 7.7 g sucrose. Filter with 0.2 micron pore filter. Store at 4ºC.

Egg Buffer: Mix 29.5 ml of 2M NaCl, 12 ml of 2M KCl, 1 ml of 1M CaCl2, 1 ml of MgCl2, 12.5 ml of 1M HEPES-NaOH pH 7.2 and 435 ml molecular grade water. Filter with 0.2 micron pore filter. Store at 4ºC.

Pronase E solution (15 mg/ml): Weigh 150 mg of Protease from Streptomyces griseus (Sigma P8811-1G) into a 15 ml tube. Dissolve the enzyme in 10 ml of Egg Buffer. Nutate the solution for approximately 10 minutes until dissolved. Prepare 1 ml aliquots in 1.5 ml tubes. Store aliquots at -20ºC.

2. On day of protocol:

   Cool swinging bucket centrifuge to 4ºC

   Thaw Pronase aliquots at room temperature

   Place L15-10 and egg buffer on ice

   Wash and sterilize Dounce homogenizer and pestle A (1ml H20, 1ml 70% EtOH)

3. Starting material:

   Worm suspension in 15 ml tube (material generated from protocol B.2)

   Strains: N2, fluorescent sorting strain

   Perform this protocol on both strains in parallel

   NOTE: The volumes for chemical and enzymatic treatments in this protocol require an L1 pellet less than 200 ul. If L1 pellet exceeds 200 ul, utilize 2x the L1 pellet volume.

B.4.3.2 Harvest L1 Worms

4. Harvest the synchronized L1 worms by washing the plates with fresh M9.
5. Pass the harvested L1 suspension through a 20 micron filter. This will filter any contaminating debris (agar chunks, partially bleached worm chunks) and any unhatched or dead embryos.

6. Pellet the worms in a 15 ml tube for 1 min at 2,000 rcf. Discard the supernatant and repeat M9 washes until the supernatant is free of visible E. coli.

B.4.3.3 SDS-DTT Treatment

7. Once the worm suspension is free of E. coli, centrifuge worm suspension at 2,000 rcf for 1 minute
8. Resuspend the worm pellet in 1 ml of M9 and transfer to a 1.5 ml tube.
9. Pellet the worms again at 2,000 rcf for 1 minute
10. Transfer 10 ul of L1 worm pellet to 1ml of Qiazol and store at -80°C for downstream RNA analysis
11. Resuspend the worm pellet in 200 ul of fresh SDS-DTT solution
12. Incubate for 2 min room temperature to digest the cuticle

    NOTE:

    - Worms should look ruffled, and pharynx should change from thin and elongated to short and round

    - Incubation exceeding 2 min will damage the sample
13. Quench reaction by adding 800 ul of ice cold egg buffer
14. Pellet worms at 13,000 rcf for 15 seconds at 4ºC
15. Decant the supernatant and replace with 1 ml of ice-cold egg buffer
16. Repeat the egg buffer wash for a total of 5 washes

B.4.3.4 Pronase E Treatment

17. Resuspend the SDS-DTT treated worms in 200 ul of 15 mg/ml Pronase E
18. Incubate for 30 minutes at room temperature nutating to digest the cuticle
19. Quench the Pronase E treatment by adding 1 ml of L15-10
20. Pellet the worms at 13,000 rcf for 15 seconds at 4°C
21. Decant the supernatant and resuspend the worms in 1 ml of L15-10
22. Pellet the worms at 13,000 rcf for 15 seconds at 4°C

23. Resuspend the worms in 1ml of L15-10

B.4.3.5 Cell dissociation

24. Transfer the worm suspension to a 2 ml glass Dounce homogenizer
25. Perform 100 strokes with Dounce pestle A to generate a worm slurry
26. Visually confirm worm dissociation by viewing a 2 ul sample of worm slurry on a fluorescent microscope
27. Harvest the cells
    a. Transfer worm slurry to 1 ml centrifuge tube
    b. Pellet undissociated worms at 100 rcf for 1 minute at 4ºC in swinging bucket centrifuge

       NOTES:

       - This step will separate the dissociated cells from intact worms
       - Cells will remain in the supernatant
       - Intact and partially dissociated worms will remain in the pellet
       - Ensure your cell type of interest is not lost during this step
       - Visually confirm fluorescent cells **are present** in the supernatant
       - Visually confirm fluorescent cells **are not present** in the pellet
       - You may need to reduce the centrifuge speed and/or time if fluorescent cells are in the pellet of this step

    c. Aspirate 1 ml of the cell-containing supernatant. Keep the pipette away from the pelleted worm debris
    d. Dispense the cell suspension though a 35-micron nylon mesh filter into a 5 ml flow cytometry tube
28. Resuspend remaining worm slurry in 1ml of L15-10
29. For the sorting strain, perform an additional round of dissociation and cell harvest (Steps 24-28). You will perform a total of two to three homogenization cycles, until few intact

worms remain.

Total cell suspension volumes:

- Control strain = 1ml

- Sorting strain = 2-3ml

30. Transfer 70 ul of cells to 1 ml of Qiazol and store at -80°C for downstream RNA analysis
    a. Continue to step 31
    b. Retain the remaining ~2ml of cells for FACS protocol B.6

B.4.3.6 Measure approximate cell concentration

31. Load 10 ul of cell suspension to a Bio-Rad TC20 automated cell counting slide. A hemocytometer can also be used for cell counting.
32. This protocol should yield between $2x10^6$ to $4x10^6$ total cells
33. Dilute the sample to $1x10^6$ cells/ml if above this concentration with L15-10
34. Microscopically confirm fluorescent cells are present in the cell suspension
35. If the total cell yield is less than $1x10^6$ cells and many intact worms remain, repeat the SDS-DTT, Pronase E and Dounce homogenization steps
36. Move on to FACS protocol B.6

B.5 L3 stage *C. elegans* dissociation for FACS isolation and RNA-seq analysis of intestine-specific cells

B.5.1 Abstract

This protocol is for generating a single cell suspension suitable for isolation of intestine-specific cells through Fluorescence Activated Cell Sorting (FACS) from L3 stage *C. elegans*. This protocol utilizes treatment with SDS-DTT solution and Pronase E to disrupt the cuticle. Worms are mechanically homogenized with a Dounce homogenizer.

B.5.2 Materials

Strains:

3. FACS control *C. elegans* strain, i.e. N2
4. FACS sorting *C. elegans* strain, i.e. JM149 *cals71*[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]

Reagents:

4. L15-10 solution
    a. 500 ml Leibovitz's L-15 Medium (Thermo 21083027)
    b. 50 ml Fetal Bovine Serum (heat inactivated) (Thermo 10438026)
    c. 5 ml 100X Penicillin Streptomycin solution (Thermo 15140148)
    d. 7.7 g Sucrose powder
    e. Filter with 0.2 µm pore filter
    f. Store at 4°C
5. Egg Buffer
    a. 2.95 ml of 2M NaCl
    b. 1.2 ml of 2M KCl
    c. 0.1 ml of 1M CaCl2
    d. 0.1 ml of 1M MgCl2
    e. 1.25 ml of 1M HEPES pH 7.2
6. Enzymes
    a. Pronase E, Protease from Streptomyces griseus (Sigma P8811-1G)
        i. Resuspend 150 mg Pronase E powder in 10 ml egg buffer
        ii. Nutate until powder is dissolved
        iii. Store 1 ml aliquots at -20°C
        iv. Final concentration 15 mg/ml

Consumables:

7. standard 1.5 ml tubes
8. Stericup 0.2 micron filter (Fisher S2GPU05RE)

9. 20-micron mesh filter (Fisher Scientific NC1004201)
10. 35-micron nylon mesh filter caps (Stellar Scientific FSC-FLTCP)
11. 5 ml sterile polypropylene round-bottom tube (STEMCELL Technologies 38057)
12. Bio-Rad TC20 automated cell counting slide (Bio-Rad 1450011)

B.5.3 Protocol steps

B.5.3.1 Before beginning

1. Prepare reagents in advance

   L15-10 Buffer: Mix 500 ml Leibovitz's L-15 Medium, 50 ml Fetal Bovine Serum (heat inactivated), 50 ul of 100x Penicillin-Streptomycin solution and 7.7 g sucrose. Filter with 0.2 micron pore filter. Store at 4ºC.

   Egg Buffer: Mix 29.5 ml of 2M NaCl, 12 ml of 2M KCl, 1 ml of 1M CaCl2, 1 ml of MgCl2, 12.5 ml of 1M HEPES-NaOH pH 7.2 and 435 ml molecular grade water. Filter with 0.2 micron pore filter. Store at 4ºC.

   Pronase E solution (15 mg/ml): Weigh 150 mg of Protease from Streptomyces griseus (Sigma P8811-1G) into a 15 ml tube. Dissolve the enzyme in 10 ml of Egg Buffer. Nutate the solution for approximately 10 minutes until dissolved. Prepare 1 ml aliquots in 1.5 ml tubes. Store aliquots at -20ºC.

2. On day of protocol:

   Cool swinging bucket centrifuge to 4ºC

   Thaw Pronase aliquots at room temperature

   Place L15-10 and egg buffer on ice

   Wash and sterilize Dounce homogenizer and pestle A (1ml H20, 1ml 70% EtOH)

3. Starting material:

   Worm suspension in 15 ml tube (material generated from protocol B.2)

   Strains: N2, fluorescent sorting strain

   Perform this protocol on both strains in parallel

NOTE: The volumes for chemical and enzymatic treatments in this protocol require an L3 pellet less than 200 ul. If L3 pellet exceeds 200 ul, utilize 2x the L3 pellet volume.

B.5.3.2 Harvest L3 worms

4. Harvest the synchronized L3 worms by washing the plates with fresh M9
5. Pass the worm suspension through a 20 micron filter. All L3 stage worms will be retained on the filter, while any contaminating worms younger than the L3 stage will pass through the filter. This step will also quickly remove a majority of contaminating E. coli.
6. Invert the filter and collect the retained L3 worms into a fresh 50 ml tube
7. Transfer the filtered L3 worms to a 15 ml tube
8. Pellet the worms in a 15 ml tube for 1 min at 2,000 rcf. Discard the supernatant and repeat M9 washes until the supernatant is free of visible E. coli.

B.5.3.3 SDS-DTT Treatment

9. Once the worm suspension is free of E. coli, centrifuge worm suspension at 2,000 rcf for 1 minute
10. Resuspend the worm pellet in 1 ml of M9 and transfer to a 1.5 ml tube.
11. Pellet the worms again at 2000 rcf for 1 minute.
12. Transfer 10 ul of L1 worm pellet to 1ml of Qiazol and store at -80°C for downstream RNA analysis
13. Resuspend the worm pellet in 200 ul of fresh SDS-DTT solution
14. Incubate for 2 min room temperature

    NOTE:

    - Worms should look ruffled, and pharynx should change from thin and elongated to short and round

    - Incubation exceeding 2 min will damage the sample
15. Quench reaction by adding 800 ul of ice cold egg buffer
16. Pellet worms at 13,000 rcf for 15 seconds at 4ºC

17. Decant the supernatant and replace with 1 ml of ice-cold egg buffer
18. Repeat the egg buffer wash for a total of 5 washes

B.5.3.4 Pronase E Treatment

19. Resuspend the SDS-DTT treated worms in 200 ul of 15 mg/ml Pronase E
20. Incubate for 30 minutes at room temperature nutating to digest the cuticle
21. Quench the Pronase E treatment by adding 1 ml of L15-10
22. Pellet the worms at 13,000 rcf for 15 seconds at 4°C
23. Decant the supernatant and resuspend the worms in 1 ml of L15-10
24. Pellet the worms at 13,000 rcf for 15 seconds at 4°C
25. Resuspend the worms in 1ml of L15-10

B.5.3.5 Cell dissociation

26. Transfer the worm suspension to a 2 ml glass Dounce homogenizer
27. Perform 100 strokes with Dounce pestle A to generate a worm slurry
28. Visually confirm worm dissociation by viewing a 2 ul sample of worm slurry on a fluorescent microscope
29. Harvest the cells
    a. Transfer worm slurry to 1 ml centrifuge tube
    b. Pellet undissociated worms at 20 rcf for 1 minute at 4ºC in swinging bucket centrifuge

    NOTES:

    - This step will separate the dissociated cells from intact worms
    - Cells will remain in the supernatant
    - Intact and partially dissociated worms will remain in the pellet
    - Ensure your cell type of interest is not lost during this step.
    - Visually confirm fluorescent cells **are present** in the supernatant.
    - Visually confirm fluorescent cells **are not present** in the pellet.

- You may need to reduce the centrifuge speed and/or time if fluorescent cells are in the pellet of this step.

c. Aspirate 1 ml of the cell-containing supernatant. Keep the pipette away from the pelleted worm debris.

d. Dispense the cell suspension though a 35-micron nylon mesh filter into a 5 ml flow cytometry tube

30. Resuspend remaining worm slurry in 1ml of L15-10
31. For the sorting strain, perform an additional round of dissociation and cell harvest (Steps 26-30). You will perform a total of two to three homogenization cycles, until few intact worms remain.

    Total cell suspension volumes:

    - Control strain = 1ml

    - Sorting strain = 2-3ml

32. Transfer 70 ul of cells to 1 ml of Qiazol and store at -80°C for downstream RNA analysis
    a. Continue to step 31
    b. Retain the remaining ~2ml of cells for FACS protocol B.6

B.5.3.6 Measure approximate cell concentration

33. Load 10 ul of cell suspension to a Bio-Rad TC20 automated cell counting slide. A hemocytometer can also be used for cell counting.
34. This protocol should yield between $2x10^6$ to $4x10^6$ total cells
35. Dilute the sample to $1x10^6$ cells/ml if above this concentration with L15-10
36. Microscopically confirm fluorescent cells are present in the cell suspension
37. If the total cell yield is less than $1x10^6$ cells and many intact worms remain, repeat the SDS-DTT, Pronase E and Dounce homogenization steps
38. Move on to FACS protocol B.6

B.6 FACS isolation of intestine-specific *C. elegans* cells

B.6.1 Abstract

This protocol describes the steps necessary for FACS isolation of *C. elegans* intestine cells. Section 1 contains general guidelines and considerations to be made before performing a FACS experiment. Sections 2-4 detail how to prepare cells for a specific developmental stage. Sections 5-7 detail how to prepare the FACS instrument for sorting and evaluating the purity of isolated cells. Section 8 describes how to handle cells once they have been isolated.

B.6.2 Materials

Strains:

1. FACS control *C. elegans* strain, i.e. N2
2. FACS sorting *C. elegans* strain, i.e. JM149 *cals71*[elt-2p::GFP::HIS-2B::unc-54 3'UTR + rol-6(su1006)]

Consumables:

1. DRAQ5 stain (Thermo 62251)
2. ReadyFlow Propidium Iodide (Invitrogen R37169)
3. 5 ml sterile polypropylene round-bottom tube (STEMCELL Technologies 38057)
4. Qiagen miRNEasy Micro kit (Qiagen 217084)

Equipment:

1. BD FACSAria III Cell sorter

B.6.3 Protocol Steps

B.6.3.1 Before beginning

1. The starting material required to perform this procedure was generated with either of the following protocols:
    a. Embryo Stage Dissociation, protocol B.2
    b. L1 Stage Dissociation, protocol B.3
    c. L3 Stage Dissociation, protocol B.4

2. Ensure your cells do not exceed a concentration of $1x10^6$. Cell suspensions exceeding this concentration are challenging to analyze on the FACS instrument for cells of this size and may increase the probability of doublets/hitchhiking cells.
3. This protocol utilizes dye combinations determined optimal for the corresponding stage and cell suspension concentration. Alternative dye combinations can be used but ensure that the dyes used do not spectrally overlap. Consult resources such as FluoroFinder (https://fluorofinder.com/). When using new dyes, perform a staining index experiment to determine the optimal stain concentration to use.

   Dye rationale:

   a. Embryo stage - Stain with viability dye Propidium Iodide to separate live cells (PI-) from dead cells (PI+)
   b. L1 stage - Viability dyes are not used, as intestine cells preferentially take up viability dyes and will confound sorting
   c. L3 stage - Similar to L1, viability dyes are not used. L3 cell preps have a high degree of debris, so a cell permeable nucleic acid dye such as DRAQ5 is used to distinguish cells (DRAQ5+) from debris (DRAQ5-).
4. When first performing this assay, aim to microscopically visualize samples isolated through FACS as a primary endpoint for assay success. Intestine samples should be GFP+, resemble intestine cell morphology, and contain few to no GFP- cells. Alternatively, non-intestine samples should be GFP- and contain few to no GFP+ cells. Once confident in the sorting procedure, samples do not need to be microscopically visualized for each experiment to preserve material for RNA-seq analysis.
5. The following steps describe how to prepare samples necessary for FACS isolation of intestine cells. If using embryo samples, follow Steps 6-9 for cell preparation. If using L1 samples, follow Steps 10-11. If using L3 samples, follow Steps 12-15. Once cells are prepared, proceed to step 18 for FACS.

B.6.3.2 Embryo stage cell prep

6. Divide samples into flow tubes with the volumes and labels indicated in Table B.1
7. To the appropriate tubes indicated in Table B.1, Add two drops of ReadyFlow Propidium Iodide per $1x10^6$ cells/ml
8. Incubate for 15 mins on ice protected from light
9. Proceed to FACS (Step 18)

B.6.3.3 L1 stage cell prep

10. Divide samples into flow tubes with the volumes and labels indicated in Table B.2
11. Proceed to FACS (Step 18)

B.6.3.4 L3 stage cell prep

12. Divide samples into flow tubes with the volumes and labels indicated in Table B.3
13. To the appropriate tubes indicated in Table B.3, add 1 ul of DRAQ5 per ml of cell volume
14. Incubate for 30 minutes on ice protected from light
15. Wash cells to remove excess dye. Pellet cells by centrifuging for 5 minutes at 530 rcf in 4ºC swinging bucket centrifuge.
16. Resuspend cells in L15-10 equal to their starting volume
17. Proceed to FACS (Step 18)

Table B.1: Sample table for embryo stage FACS

Sample type	Worm strain	Dye	Cell volume (ml)
Unstained	N2	None	0.5
GFP single	GFP strain	None	0.5
PI single	N2	Propidium Iodide	0.5
Sorting sample	GFP strain	Propidium Iodide	2-4

Table B.2: Sample table for L1 stage FACS

Sample type	Worm strain	Dye	Cell volume (ml)
Unstained	N2	None	1
Sorting sample	GFP strain	None	2-4

Table B.3: Sample table for L3 stage FACS

Sample type	Worm strain	Dye	Cell volume (ml)
Unstained	N2	None	0.5
GFP single	GFP strain	None	0.5
PI single	N2	Propidium Iodide	0.5
Sorting sample	GFP strain	Propidium Iodide	2-4

B.6.3.5 FACS Setup

18. The following protocol steps were developed on a BD FACSAria III instrument. The settings and recommendations detailed here may be adjusted as necessary.
    a. Setup the instrument:

       - Set laser gain and area scaling factor on the FACS instrument to ensure that collected data falls within the plotting area.

       - Set plots to log-log axis

    b. **NOTE:** *C. elegans* cells are smaller than material typically run on a FACS instrument. Compared to mammalian cell lines, *C. elegans* cells require a lower area scaling factor and lower laser voltages. Additionally, set plot axes to log-log transformed.
19. Run the unstained control cell samples through the FACS instrument. Adjust laser voltages and area scaling factors to ensure that collected data (i.e. SSC and FSC) falls within the plotting area. Once values are set properly, record data from ~10,000 events
20. Run all appropriate single stain control samples through the FACS instrument and record data from ~10,000 events. Set gates based on single stain controls.
21. Run the sorting sample through the FACS instrument and record data from ~10,000 events
22. Verify that gates are set appropriately for the sorting sample and adjust gates as needed
23. Figure B.1 provides example gating strategies utilized for intestine isolation. Simultaneously collect cells within the "intestine cells" and "non-intestine cells" gates.

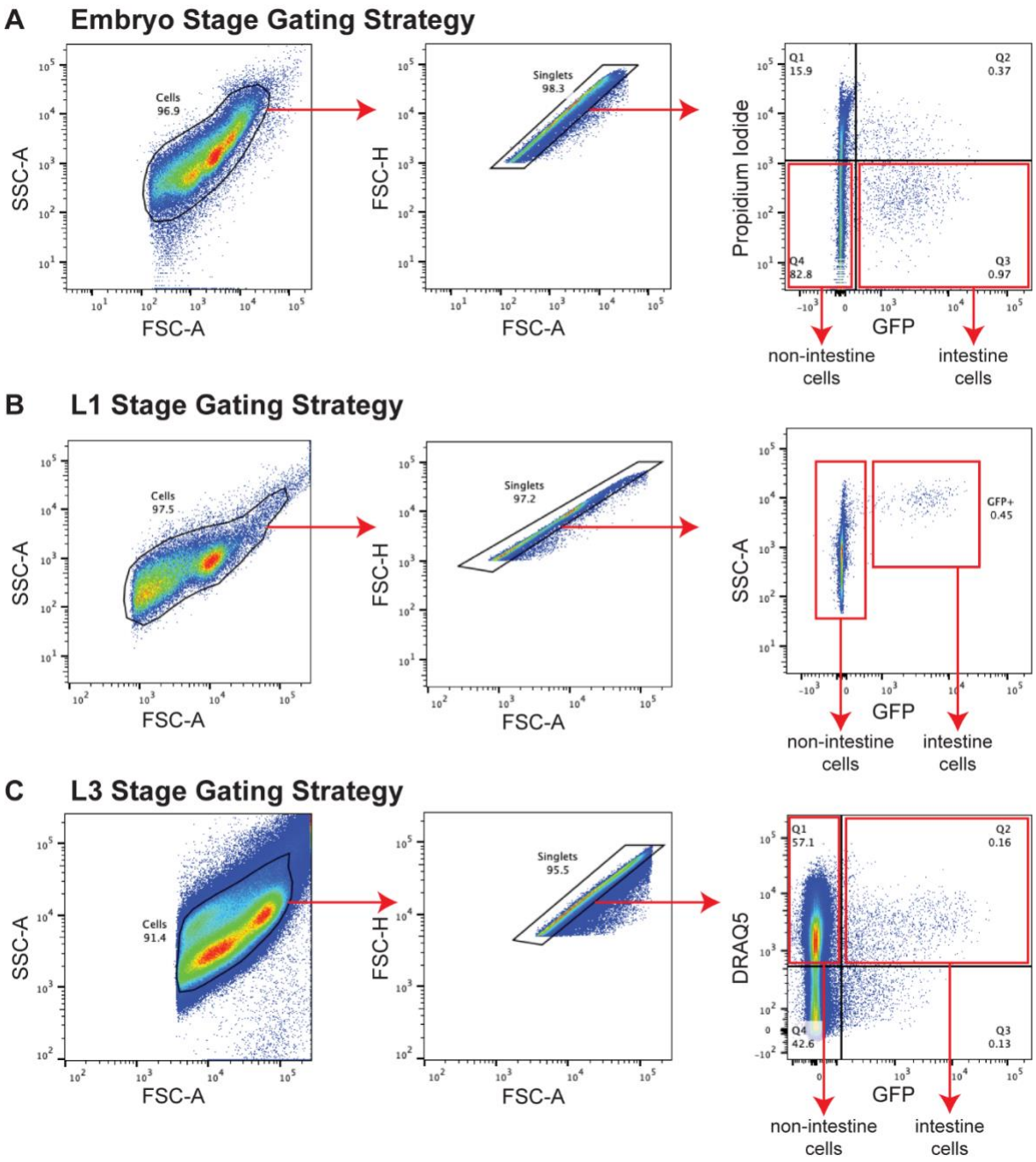

Figure B.1: FACS gating strategy for isolation of intestine cells. Diagram of gating strategy used for embryo (A), L1 (B), and L3 (C) stage FACS intestine isolation.

B.6.3.6 Sort Cells

24. Once gates and gating strategy are established, prepare two flow tubes to collect the sorted cells by aliquoting 0.2ml of L15-10 into them. Place the collection tubes into the tube holder inside the FACS instrument.
25. Prepare the FACS instrument for a simultaneous two-way collection of cells from the "non-intestine cells" gate and "intestine cells" gate outlined in Step 23
26. Collect $1x10^6$ cells from the non-intestine cells gate. This should yield a total volume of 3.5-4ml of cell suspension. Store sorted cells on ice.
27. Collect cells from the "intestine cells" gate continuously until there is no more sorting sample remaining. This should yield between 50,000 to 300,000 cells in a total volume of 0.2-0.5ml. Store sorted cells on ice.

B.6.3.7 Post-sort purity analysis

28. Once sorting is complete, perform a post-sort purity analysis. This will determine if the FACS run was successful.
29. The microfluidic line must be cleaned after sorting and before measuring post-sort purity for each sample. Clean the line by running a sample of fresh 10% bleach for ~5 mins. Follow this by running a sample of filtered H_20 for ~2 mins. The FACS instrument should record fewer than 1 event/sec.
30. Run the intestine cell sample back through the instrument. The intestine cell post-sort purity should be 80-90%.
31. Clean the microfluidic line following the procedure outlined in Step 29. Run the non-intestine cell sample back through the instrument. The non-intestine cell post-sort purity should be 100%.

B.6.3.8 Sample preparation for RNA extraction

32. Prepare the cells for downstream RNA analysis: Pellet intestine and non-intestine cells at 10,000 rcf for 5 mins in 4ºC cooled centrifuge. There should be a small but visible pellet in the intestine cell population.
33. Decant supernatant and resuspend cells in 1 ml Qiazol. Store at -80°C until ready for RNA isolation and quantification.
34. We utilize the Qiagen miRNEasy kit for RNA isolation. Quantification is performed with High Sensitivity RNA kits from Qubit and Tapestation. By following the protocols in this collection, expect approximately 10 ng total RNA for intestine cell samples (50-300,000 cells) and >100 ng total RNA for non-intestine cell samples ($1x10^6$ cells).

APPENDIX C

smiFISH PROBE SET SEQUENCES

Table C.1: smiFISH probe sets used in this dissertation

Gene name	Sequence name	WormBase ID	Fluorophore	smiFISH FLAP	Probe Sequence
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	GTAGTTGTTCTTGAAACACTTTTCCACGTCGCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	ACTTGAAATGGAATGTCCTTGTACTTCAGCCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	AGTTCTTGTAGACATTCCAGTCAGCCTTTCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	GGTTTGCCAGTGTCACATCATTACTCCATCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TATTCTGACGCCAAGTTGGTGAATCCACATCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TCTTATGAAACATTCGTTCTTCGTGTGTCTCCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TTTCATTTCAAGTTCTCGAAGAACCTGGCTTCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TTGTCACTTCCAATGACTCCAGCCACGCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TCCAGCAAGTCCACTTGCCACCATATCCTCCTAAGTTTCGAGCTGGACTCAGTG

Table C.1: smiFISH probe sets used in this dissertation

Gene name	Sequence name	WormBase ID	Fluorophore	smiFISH FLAP	Probe Sequence
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TCTCATTGAGCAGTCGAAGGCATTCCACCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	TTCGCTTCATTTGAAGTTGTTCGTCTAGAGCCCCTCCTAAGTTTCGAGCTGGACTCAGTG
acy-4	T01C2.1	WBGene00000071	Cal Fluor 610	X	CAAGTCGGGCGTGTTTGATGAGGGCTCTTCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	TCGACACGATTTACAACGGTTCAGTTTCAGGCCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	TCAAAGGCCTCTATATATTGTTGGTCCCTCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	GGTCGTTTGTCGCTTCCTCTATCATATACACCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	TGTATTTCCTAGGTTGTTGTTCTCTTGCCACCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	GGAAGTTGATTTTTTTTGATCCCAACGTCGCCCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	TATCATTGATTAGTGGGGTGAGTTGATCCGGACCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	ACGTCTTCATCGATTTTGCCAATTAGATTGGGCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	CAATTTCTCGTCTGCTCCATTTATTCCGACCCTCCTAAGTTTCGAGCTGGACTCAGTG

Table C.1: smiFISH probe sets used in this dissertation

Gene name	Sequence name	WormBase ID	Fluorophore	smiFISH FLAP	Probe Sequence
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	GAAATAATGCCTTCATCCTTCCATCGCCCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	TCCAAGTTTTTCGTAGAATGTTGCGACGAGCGCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	GTCTTGTATACACTCTCGATCGATTTGGGTGCCTCCTAAGTTTCGAGCTGGACTCAGTG
bre-2	Y39E4B.9	WBGene00000267	Cal Fluor 610	X	GGCTTTGCCGCATTGGTGGAGAGGTAACCCTCCTAAGTTTCGAGCTGGACTCAGTG
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TTACAGCTCGTCCTTGTTTGACGAATGCATTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	CTACTTTTGGCATATCTTCTGGTGACAATCCTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	CCAATCAATGTTGGCTTTACTGGAGACTCCTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TGCGAGTTGCTTCGAAGTTTCCACGACGTTTTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	ACCATGTTGACTAGCCGAACCGATTTCATTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	ATGGATCAGGGGCGACTGGCTTCTCATTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TCTCAAATCATCGACTGGTGGCTTGGCTTACACTCGGACCTCGTCGACATGCATT

Table C.1: smiFISH probe sets used in this dissertation

Gene name	Sequence name	WormBase ID	Fluorophore	smiFISH FLAP	Probe Sequence
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	CCATTCTACATCTTCTATTTGGGGAACTCCTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TGAGATCCTTCACTTCTGCCGGTTTTCCTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	CAGTTGTCCATCCCTGAACTTCAATTGCGTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TTCCTTCGCTGAAGAATCCCATGAATCCTTACACTCGGACCTCGTCGACATGCATT
ges-1	R12A1.4	WBGene00001578	Quasar 670	Y	TTCGGTATTGGGTTGCTGGGTAGACATCTTCTTACACTCGGACCTCGTCGACATGCATT