APPROVAL SHEET

Title of Dissertation: **Identification of targets of the heterochronic protein LIN-29 and their implications for the larva-to-adult transition in *C. elegans***

Name of Candidate: **Patricia Abete-Luzi**
Doctor of Philosophy, 2018

Dissertation and Abstract Approved: _____
Dr. David M. Eisenmann
Professor
Department of Biological Sciences

Date Approved: 9/24/2018

Title of dissertation: **Identification of targets of the heterochronic protein LIN-29 and their implications for the larva-to-adult transition in *C. elegans***

Author: **Patricia Abete-Luzi, Ph.D. 2018**

Dissertation directed by: **Dr. David M. Eisenmann**

Professor

Department of Biological Sciences

The precise temporal control of gene expression is an essential aspect of metazoan development. The ecdysozan *C. elegans* goes through six life stages (embryo, L1 to L4, and adult) separated by molts, and analyses of modENCODE data showed that most members of the large cuticle collagen (*col*) gene family express predominantly at a single stage. This temporal preference motivated us to investigate the mechanisms underlying *col* gene expression and here we focus on a subset of *col* genes expressed in the L4 stage. We identified minimal promoter regions of <300 bp for *col-38*, *col-49*, and *col-63* in which we predicted *cis*-regulatory sequences and evaluated their function *in vivo* via mutagenesis of a *col-38p::yfp* reporter. Regulatory sequences that turned out to be crucial for *col* expression correspond to LIN-29 binding sites. The zinc finger transcription factor LIN-29 is the most downstream effector of the heterochronic pathway and is known to regulate the larva-to-adult switch in worms. We used RNAi to study the requirements for candidate transcription regulators and found LIN-29 to be necessary for the expression of four L4-

specific genes (*col-38*, *col-49*, *col-63* and *col-138).* Temporal misexpression of LIN-29 was also sufficient to activate these genes at a different developmental stage. The LIN-29 DNA-binding domain bound three minimal promoters *in vitro.* For *col-38* we showed that the LIN-29 sites necessary for reporter expression *in vivo* are also bound *in vitro*. To expand our knowledge of the L4 regulatory network, we used RNA-Seq to identify genes differentially expressed between animals overexpressing LIN-29 and control animals. After ectopic expression of LIN-29 in the early L3 stage, we identified 230 and 350 genes that were upregulated and downregulated, respectively. Hypodermal genes encoding cuticle components that normally peak in the L4 stage were overrepresented among our LIN-29 activated targets; whereas genes with intestinal expression and roles in fat metabolism where enriched among our LIN-29-repressed targets. Moreover, we identified signaling molecules WRT-6 and INS-37 that act as mediators of a cell non-autonomous LIN-29-dependent regulation of genes that exclusively express in the intestine and encode fat metabolic enzymes and vitellogenins. Our results reveal a new role for LIN-29 and provide new insights for the exploration of a potentially conserved gene regulatory network controlling a metabolic shift during the developmental transition to adult life.

Identification of Targets of The Heterochronic Protein LIN-29 and Their Implications for The Larva-To-Adult Transition In *Caenorhabditis elegans*

by

Patricia Abete-Luzi

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, Baltimore County in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2018

This work is dedicated to my beloved family, who always believed in me:

To my creative and overachieving mother, Mariela, whose daily fighting the good cause brings fairness and quality to this world; to my adventurous and clever father, Gustavo, who never misses a second of happiness and always showed me how to see the bright side; and to my incredibly smart and amazing lil' brother, Diego, who I saw grow up and become that guy who will always earn everyone's hearts, mine included.

Acknowledgements

My deepest appreciation goes to my advisor, Dave Eisenmann, for being such a remarkable mentor and human being, who shared his time, his space, his wisdom and his heart with me, to help me grow, and to become a better scientist and a better person. The mentor-mentee relationship is a unique and exceptionally challenging partnership that relies on a considerable amount of trust, patience and open-mindedness. I was, and I still am, very fortunate to have formed such a bond with Dave whom as a *boss*, I will never be able to replace. I may own the chapters of this thesis, but he owns this chapter of my life, and he will always have a place in my world. Also, he has the best haircut on campus.

Second, I want to thank my advisory committee: Steve Miller, Phyllis Robinson, Phil Farabaugh and Andy Golden, for saying *yes* to me, for their time and their dedication to ensure my academic development, for their vision, and their fair judgement. I know I am not the most typical student, yet I think they saw me for who I could become. I thank Andy for taking the task of joining our academic endeavors at UMBC to help me become a doctor. I thank Phyllis for her teachings since my first years, for fighting the good causes and for always being upfront. I also thank Phil for committing to be on my committee, even when the department needs his full attention, and for his constant push to make me smarter. My qPCR data representation wouldn't look that beautiful without him. Finally, I would like to give a special thanks to my favorite former Graduate Program Director (GPD) of all, Steve, whose sense of fairness will always be unmatched. I experienced his GPDness as a newbie, his teaching as a student, his teaching as a TA, and his advice as a whining

student. All I can say, is that Steve is quality. No matter what task you give him, he will always do his absolute best.

Life as a graduate student at UMBC was not particularly easy, however, amazing characters inhabit(ed) our department and made our days brighter with their uniqueness and their good hearts. Examples include Cynthia Wagner, Steve Caruso, Ivan Erill, Jeff Leips, Mike Swierczewski, Sarah Jackson and Tom Cronin. In fact, as for the UMBC community, I want to give a shout out to the crowd of the third floor, in particular the Brewster Lab for the constant lively vibe, our Graduate Program Coordinator, Brandy Darcey, and Michelle Massey from the International Education Services who has been my angel all these years.

Last but not least, big thanks to those who became my Maryland family over the last seven years: first, my cat JB, who always kept my heart warm and my oxytocin levels really high; and second, my partners in crime and truly amazing friends: Sharlene Brown, Cassie Nelson, Maggie Correa-Mendez, Brian Gregory, Dan Beury, Rob Nunn, Maki Negesse and the rest of the Poetry gang. Like Sharlene said to me '*you came with nothing'*, and now I leave with a PhD and some of the greatest friendships, which I plan to keep until the end of days.

Table of contents

Chapter 1: Introduction

I. Temporal control of gene expression during animal development

The spatiotemporal regulation of cellular and molecular events is a crucial aspect of metazoan development. After fertilization, a zygote must go through a series of cell division, cell differentiation, cell migration and cell death events in order to generate the tissues and organs that will provide the final structures and connections characteristic of a functional adult body. For this orchestration to be successful, events must occur not only in the right place but also at the right time. While the spatial control of gene expression, the subsequent cellular decision-making, and its impact on developmental patterning have been extensively studied, the temporal aspect of such regulation is of equal relevance yet less understood. There are a number of different variables when considering the complexity of developmental timing, from the time allotted for specific processes to happen, to the rate at which they happen and the way they take place in synchrony with other events. Full coordination of tissues and systems is required when architectural modifications take place, either once or repeated times such as during molting cycles. Proper spatial patterning in the wrong temporal context can sometimes dramatically affect structure, size and functionality of the overall body or sections of it. Failure of developmental synchrony can result in fatal birth defects and malformations, such as human holoprosencephaly in which the forebrain fails to divide partially or completely (the most severe case results in *cyclopia*) as a result of the delay of prechordal mesoderm differentiation with respect to other parts of the prosencephalon (Wilson, 1988).

Several different mechanisms that regulate developmental timing have been found among different phyla, including the establishment of molecular gradients of key regulators and their diffusion, the on and off response to self-regulatory signaling mechanisms, and the response to environmental cues. All of these mechanisms ultimately lead to changes in gene expression. One example is embryonic development in Drosophila, in which sequential waves of transcription factors exhibit different expression patterns at different timepoints: a first wave of factors derived from maternal transcripts will locally regulate their targets based on their concentration gradients and promote the accumulation of a new set of factors, which once active will regulate the next wave of factors, and so forth (see overview by Phillips, 2008). The developmental timing in this scenario is determined by the rate of protein synthesis, diffusion and degradation. A second example in Drosophila, as well as in most insects, involves pulses of the steroid hormone *ecdysone* which drive the transition between developmental stages by promoting molting and metamorphosis. The duration of juvenile stages, however, is determined by the accumulation of another hormone, the *juvenile hormone* which represses metamorphosis. Only when levels of juvenile hormone are low enough during the last instar, is ecdysone able to trigger the structural changes that will lead to the final adult form. Mutations that result in the decrease of juvenile hormone titers cause premature metamorphosis in many insects (see review by Di Cara & King-Jones, 2013). A third example of temporal control is seen during vertebrate embryonic development in a process called somitogenesis. Mesodermal somites are the precursors of the spinal skeletal system (including vertebrae and muscle) and their formation consists of a temporally-regulated segmentation process that takes place in a periodic manner. The rhythm (rate) of somite formation is unique for most vertebrate

species and it is controlled by the oscillation of gene expression driven by Notch signaling and its own negative feedback loop (see review by Pourquie, 2003). Finally, while the default developmental timing tends to be regulated by intrinsic genetic programs, changes in environmental or nutritional conditions may also affect developmental timing by either delaying or speeding up developmental transitions. Examples of developmental stasis are the diapause states adopted by insects and free-living nematodes (i.e. dauer stage) or mammal hibernation, all in response to adverse conditions or lack of nutrients. However, the way in which external stimuli affect development is a large topic and beyond the scope of this thesis.

The broad goal of my dissertation is to better understand the temporal regulation of gene expression and its role in *C. elegans* development. Interestingly, *C. elegans* possesses a time-keeping mechanism that regulates lineage-specific cell fate decisions in the skin and other tissues; it is called the *heterochronic gene pathway* and I will describe it later in this chapter. In the remainder of this introduction, I will provide an overview of *C. elegans* as a model organism (II), then I will discuss the worm's cuticle and cuticle collagen gene expression as a model to study temporal regulation (III), the heterochronic pathway (IV), the heterochronic protein LIN-29 in the larva-to-adult transition (V), and I will finish this chapter with a summary of my thesis work (VI).

II. *C. elegans* as a model system

The nematode *Caenorhabditis elegans* is a free-living, 1 mm-long roundworm naturally found in soil and rotting fruit, that is widely used as a model system in disciplines such as

genetics, developmental biology and neurobiology. Populations are composed mostly of self-fertilizing hermaphrodites, whereas males occur at a much lower frequency (0.1-0.2%, unless they are induced in the laboratory; Lints & Hall, 2009). In the absence of males to mate with, a single hermaphrodite can generate about 300 offspring in its lifetime. Since only hermaphrodites were used in my work, only hermaphrodite properties will be discussed henceforth.

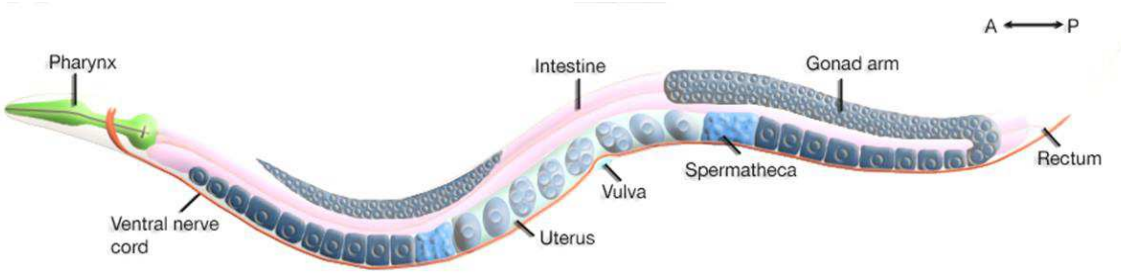


Figure 1. Representation of *C. elegans* anatomy. A cartoon of the adult hermaphrodite is shown; main body structures are as indicated. Figure from Altun & Hall (2009)

While *C. elegans* has less morphological complexity than vertebrates, it shares cell and tissue types and molecular developmental mechanisms with more complex species. The adult hermaphrodite body has a total of 959 somatic cells. The major structures in the adult are a digestive system, nervous system, body-wall and pharyngeal muscles, skin, and a somatic gonad that produces fertilized eggs which are internally-fertilized before they are laid through a specialized structure called the *vulva*. (Figure 1; Altun & Hall, 2009). The embryonic cells that give rise to all these structures have been traced and, consequently, the entire cell lineage from zygote to adult is known in this organism (Sulston & Horvitz, 1977; Sulston, Schierenberg, White, & Thomson, 1983). This fact has made *C. elegans* a

great system to study developmental timing in the context of cellular decision-making, as I will explain later in this chapter.

The worm is also a powerful system for genetic analysis: its genome was one of the first to be fully sequenced revealing ~22,000 genes, a number that is similar to that in humans (Antoshechkin & Sternberg, 2007; Xu & Kim, 2011). The genomes of other species of *Caenorhabditis* have also been sequenced and this provides a platform for evolutionary and comparative studies at the genomic level (Xu & Kim, 2011). *C. elegans* is very suitable for both forward and reverse genetic assays and screens. In particular, RNA interference assays are widely used in this organism, and genome editing by CRISPR is easily performed (Dickinson & Goldstein, 2016). Other properties that make *C. elegans* a convenient model organism include: 1) larvae can be frozen allowing long term storage of strains, 2) it is easy to grow and inexpensive, using *E. coli* as a source of food on small Petri plates, 3) it has a transparent body which allows *in vivo* imaging as well as cell lineage analysis, and 4) it has a relatively short life cycle, growing from embryo to adult in about 3.5 days at 20℃.

The first stage of the worm's life cycle is embryogenesis, which occurs inside an eggshell: after leaving the mother's uterus, most of embryogenesis occurs in the external environment. After embryogenesis is complete, a small worm-shaped L1 stage larva hatches from the eggshell and starts postembryonic larval development. During this time, the worm goes through four larval stages (L1, L2, L3 and L4) separated by molting events, before becoming an adult that is capable of laying eggs (Figure 2). An exception to the

usual order of stages occurs when conditions are unfavorable, when an alternative quiescent *dauer* larva forms after the L1 stage. A worm can stay in the dauer state for months until conditions are good again, when it molts into an L4 and continues the normal cycle (Figure 2; for more detail about this topic, see Hu, 2008).

Like in other ecdysozoans, molting is a crucial aspect of development and body growth. The outer covering of the worm is called the cuticle; every time the worm molts an old cuticle is shed and a new cuticle is produced around a now bigger larva. The temporal regulation of this cyclic process in terms of cuticle synthesis inspired the work described in Chapter 2, for the reasons I will discuss in the following section.

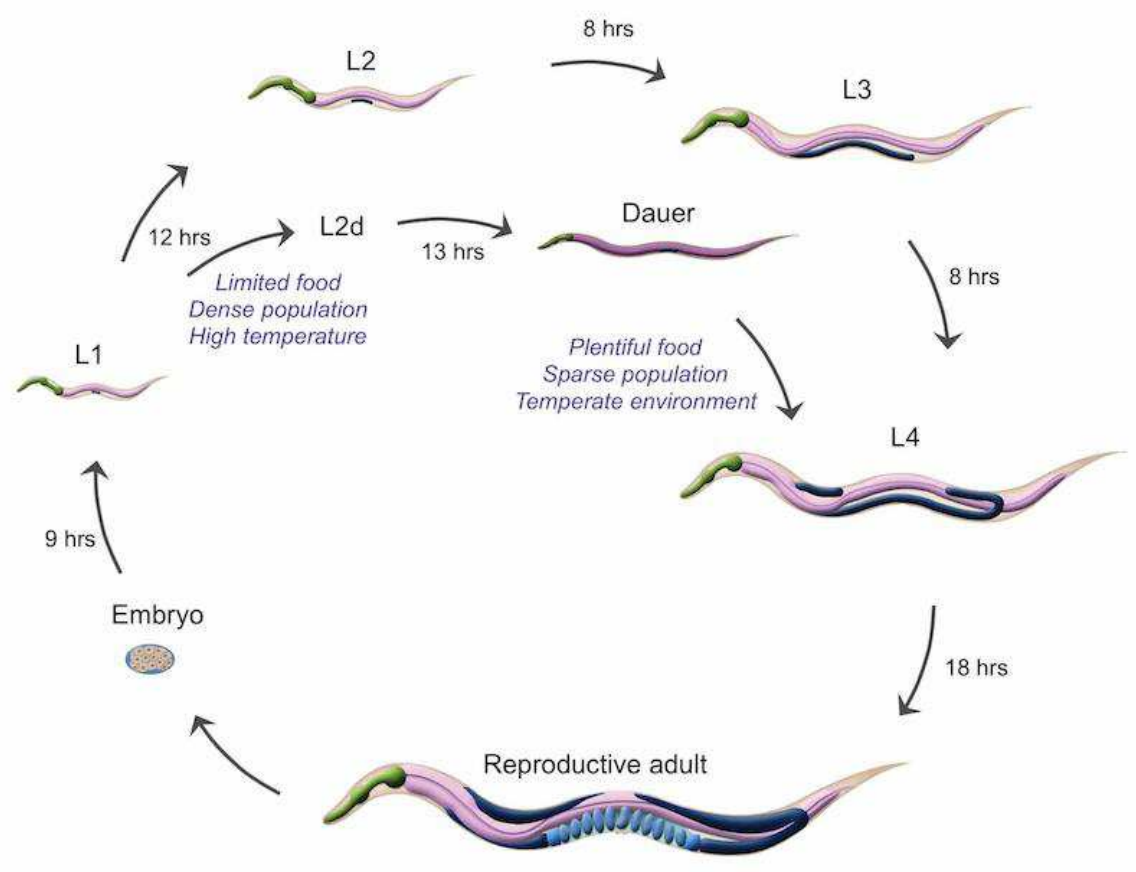


Figure 2. *C. elegans* life cycle. A cartoon of the different life stages and the time between them at 22°C is shown. Body sizes are not proportionally scaled. After embryonic development, the worm hatches and goes through 4 larval stages (L1-L4) before becoming adult, molting its outer cuticle at each transition. The full cycle takes 2-2.5 days at 22°C and 3.5 days at 20°C. An exception to this life cycle takes place when in harsh conditions, the L1 larva arrests and enters a dauer stage which can survive up to four months until conditions improve. Figure from Wolkow & Hall (2015).

III. The *C. elegans* cuticle and cuticle collagens genes as a model to study temporal regulation of gene expression

The outer covering of the worm, known as the cuticle, is an exoskeletal structure that defines the animal's size and shape as well as protects it from the environment. Beneath the cuticle lies a single cell layer-thick epithelium called the *hypodermis*, which is the skin of the worm and which is responsible for cuticle secretion.

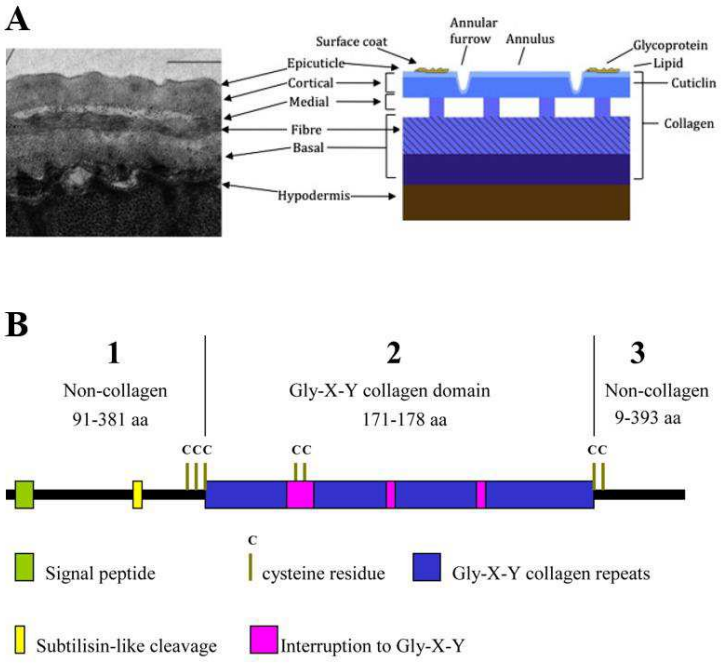


Figure 3. *C. elegans* cuticle is primarily composed of cuticle collagen. A) On the left: a transmission electron micrograph (TEM) of a cross-section of the adult cuticle; on the right: a cartoon highlighting the adult cuticle layers as indicated. Scale bar: 1µm. Figure from Page, Stepek, Winter, & Pertab (2014). B) Generic representation of the main elements of a cuticle collagen protein. From Page & Johnstone (2007)

Five cuticles are synthesized *de novo* during the life of this nematode: once before hatching as an L1, and once before each of the four larval molts (at the end of the L1, L2, L3 and L4 stages). On the external face, all five cuticles have annular indentations on their surface, but the L1, dauer and adult cuticles are distinguished by the presence of 2-4 lateral protruding ridges extending from head to tail on each side, called *alae*, that are not found in other stages. The larval and adult cuticles all share a layered architecture: in each stage there is an epicuticle, consisting of lipids and glycoproteins, as well as external, and internal cortical layers, all of which are composed mainly of cuticle collagens (~80% of cuticle is made of collagen; Figure 3A; Cox, Kusch, & Edgar, 1981). These collagens are secreted by underlying hypodermal cells, and like other collagens, their peptides carry Gly-X-Y collagen-repeat domains necessary for trimerization before they leave the cell (Figure 3B). As opposed to the typically long vertebrate collagens, cuticle collagens are shorter, nematode-specific collagens of about ~ 300 amino acids. Once secreted to the extracellular matrix, multiple triple helixes become hyper-crosslinked, a feature that confers the characteristic sturdiness and insolubility of the worm's cuticle.

The cuticle collagens of *C. elegans* are encoded by a large family of at least 187 genes (Johnstone, 2000). Previous work on a small number of cuticle collagen genes showed that they express in a cyclic pattern such that a peak preceded each molt (Figure 4A; Cox et al., 1981; Johnstone, 2000; Johnstone & Barry, 1996). However, when members of the Eisenmann Lab examined the expression patterns for all cuticle collagen (*col*) genes by comparing RNA-seq data obtained at different developmental timepoints (available from the modENCODE project) they found that 62% of these genes show a strong peak of

expression in only a single developmental stage (Figure 4B; Gerstein et al., 2010; Jackson, Abete-Luzi, Krause, & Eisenmann, 2014). To our knowledge, *col* genes do not show evident functional differences between their final products, most appear to express in hypodermal tissues and serve the same purpose (i.e. to constitute the cuticle); yet many of them are specific to a given stage. Ideas have been proposed to explain constantly changing cuticle constituents as a strategy to avoid the host immune system in parasitic species or for the evasion of pathogens in free-living nematodes. We do not know the evolutionary advantage for such stage-specific expression in *C. elegans*, however we believe that this behavior makes this family of small, compact genes, some of which are known to have minimal amounts of upstream DNA for correct temporal and spatial gene expression (Abete-Luzi & Eisenmann, 2018; Rougvie & Ambros, 1995), a formidable system for the study of temporal regulation of gene expression.

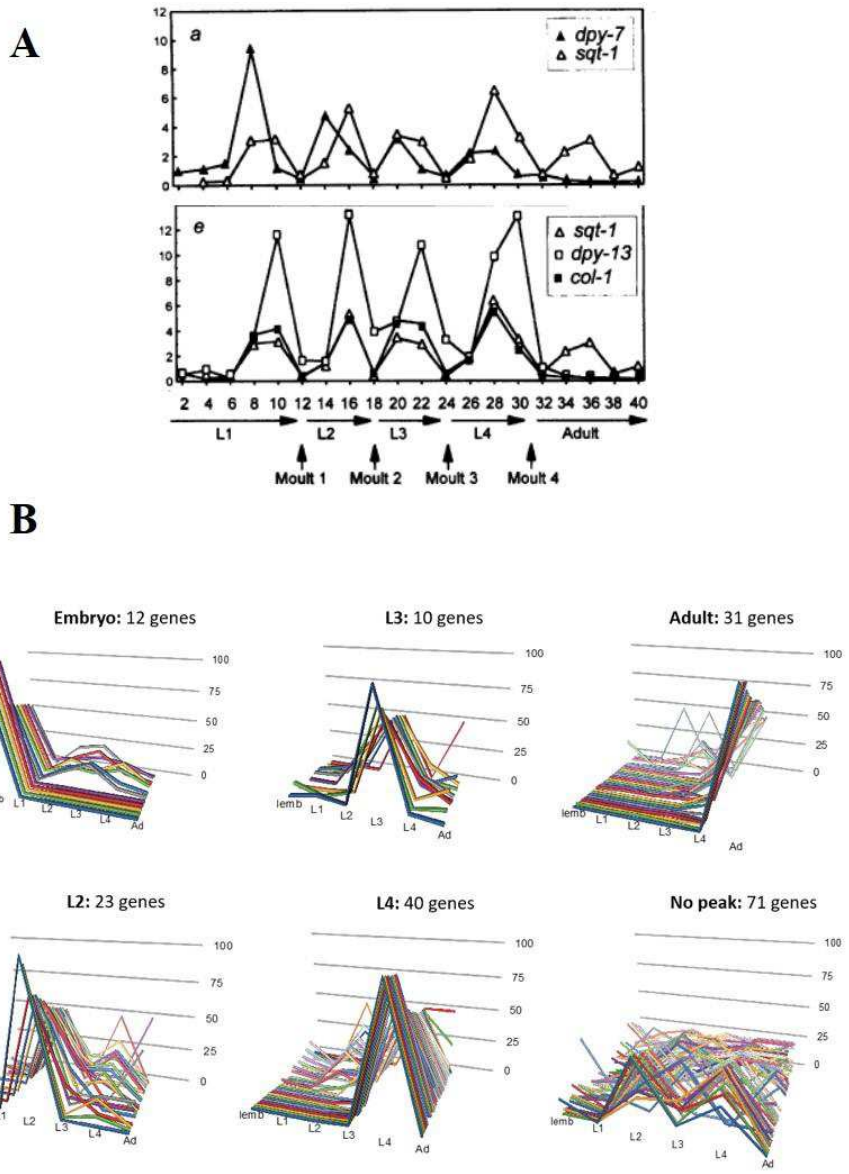


Figure 4. Temporal expression data for cuticle collagen (*col*) genes in *C. elegans*. A) Relative expression of the indicated cuticle collagen genes over time. Adapted from Johnstone & Barry (1996). B) The percentages of total gene expression observed at each of six developmental timepoints (embryo, L1−L4 larval stages and adult; data from modENCODE project (Gerstein et al., 2010) are shown. *col* genes were grouped based on their temporal expression patterns (i.e. genes that peak at the same time or show no specific peak) as indicated. From Jackson et al. (2014).

Having the *col* gene family divided into subsets based on temporal co-expression patterns (as seen in Figure 4B) set the ground for one of the main goals of the Eisenmann lab, which is to examine stage−specific subsets of *col* genes and identify the *cis* and *trans* acting requirements for their temporal regulation. My contribution towards this goal consists of the study of *col* genes that are specific to the L4 larval stage and their temporal regulation by the heterochronic protein LIN-29, which is work I will discuss in detail in Chapter 2.

IV. The heterochronic pathway regulates developmental timing in *C. elegans*

Over the past few decades, developmental studies and genetic screens have revealed a molecular timing mechanism in *C. elegans*, known today as the *heterochronic pathway*. This pathway consists of a complex network of novel proteins, transcription factors and microRNAs (miRNAs) that interact with each other to determine the temporal component of developmentally-programmed cellular events. Essentially, each larval stage is characterized by the presence of one or more heterochronic proteins, and the presence of these proteins instructs stage-specific events and processes in individual cells. When the levels of specific heterochronic proteins decline due to miRNA regulation, cells can then transition into the next developmental stage, which is now determined by a different set of heterochronic proteins (Moss & Romer-Seibert, 2014; Nimmo & Slack, 2009; Rougvie & Moss, 2013). This type of interaction and its prevalence in the pathway has been termed *the microRNA-target paradigm* (see Figure 5; Moss, 2007). It is important to keep in mind that this pathway does not regulate developmental timing in every cell of the worm but

rather in a few tissues, especially in hypodermal cells, where heterochronic genes are expressed (Bettinger, Lee, & Rougvie, 1996; Lall et al., 2006; Lin et al., 2003; Martinez et al., 2008; Perales, King, Aguirre-Chen, & Hammell, 2014; Reinhart & Ruvkun, 2001).

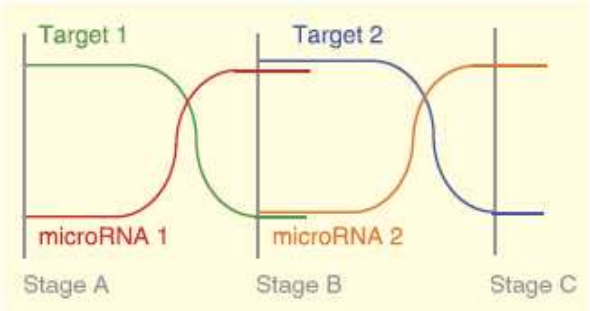


Figure 5. The *microRNA-target paradigm*. The figure shows an oversimplified representation of the heterochronic pathway interactions in *C. elegans*. From Moss (2007).

A particularly well-studied case of heterochronic regulation takes place during postembryonic development in the lateral hypodermal (seam) cells, where the pathway was discovered. Hypodermal cells can be divided into two groups: 1) those forming the hypodermis, which consist of hyp7, a large syncytium surrounding the main body, as well as other hypodermal cells located in the head and in the tail (hyp1 through hyp11; Figure 6A); and 2) epithelial cells with specialized roles such as providing neuronal protection and growth guidance, forming connections within skin cells as well as with other tissues; and forming specialized structures in the cuticle (Altun & Hall, 2005). Among the latter group are the seam cells, which are arranged in two lateral rows which extend from head to tail on each side of the body (Figure 6A). One distinctive feature of the seam cells is their active cell division cycle during larval development, before they finally differentiate and fuse as the worm transitions into adulthood (Figure 6B). As shown in Figure 6C, the

seam cell lineage follows a stage-dependent cell division pattern. During the L1, most seam cells divide once in an asymmetric stem cell-like fashion by generating one differentiated hypodermal daughter cell which later fuses with the hypodermal syncytium, and another seam cell which retains cell division potential. During the L2, some seam cells undergo a symmetric division; this time both daughter cells keep the seam cell fate. These cells then divide asymmetrically, again generating one hypodermal cell and one seam cell each. In the L3 stage, seam cells divide once again in a stem cell-like fashion. Finally, in the L4 stage seam cells divide for the last time and following this division these cells terminally differentiate: they exit the cell cycle, fuse with each other to form a single long syncytial cell, and they start producing the final cuticle and the alae. This process is part of the 'larva-to-adult' switch for the seam cells.

When heterochronic genes are mutated, the seam cell developmental program is altered such that cellular events occur either earlier or later than they normally would, phenotypes that are defined either as *precocious*, or *retarded*, respectively (Figure 6D-F). By identifying genes that caused retarded and precocious phenotypes in the seam cells and other cells, and by identifying additional genes that suppressed these phenotypes, the heterochronic pathway was uncovered (V Ambros, 1989; V Ambros & Horvitz, 1984, 1987; Nimmo & Slack, 2009; Rougvie & Moss, 2013).

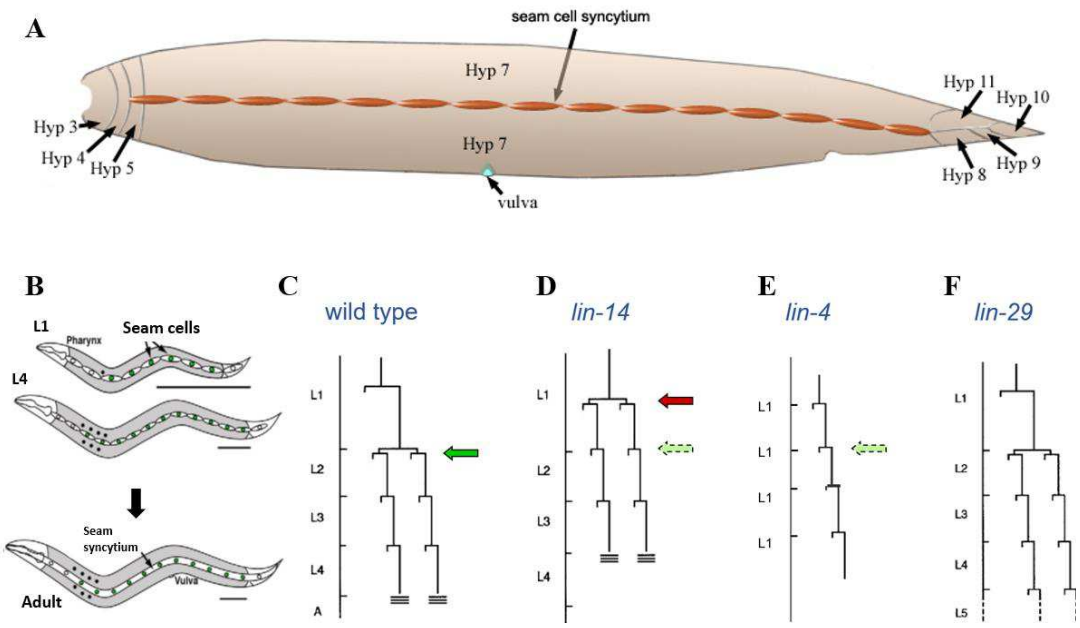


Figure 6. Seam cell postembryonic development. A) Schematic representation of hypodermal cells in the adult hermaphrodite; from Altun & Hall (2005). B) Transition of larval seam cells to their syncytial adult form. Adapted from Harris & Horvitz (2011). C-F) Seam cell division pattern over time in wild type, and heterochronic mutants *lin-14*, *lin-4* and *lin-29* (adapted from Rougvie & Ambros, 1995). Green arrows indicate when symmetric division occurs or should have occurred (dashed arrow). A precocious phenotype is observed in *lin-14* mutants since both the symmetric division (red arrow) and the terminal differentiation take place at earlier stages (D). In retarded mutants, however, either the L1 division keeps happening (*lin-4*; E) or the final differentiation never takes place and molting continues (*lin-29*; F).

The core members of the heterochronic pathway and their interactions are depicted in Figure 7A (for extensive review of this pathway, see Moss, 2007; Moss & Romer-Seibert, 2014; Nimmo & Slack, 2009; Rougvie & Moss, 2013). One of the first genes identified was *lin-14*, which encodes a transcription factor that is required for the L1 fate. At the end of the L1 stage, this gene is downregulated by *lin-4*, a microRNA (miRNA) that binds the *lin-14* 3'-UTR and interferes with *lin-14* translation. *lin-4* was the first miRNA gene to be discovered and characterized for its regulatory role (for more detail about miRNA regulation see reviews by Victor Ambros & Ruvkun, 2018; Cai, Yu, Hu, & Yu, 2009). In *lin-14* mutants, the symmetric seam cell division event observed during the L2 occurs in the L1 stage (a precocious phenotype; Figure 6D), while in *lin-4* mutants, the L1 stage is reiterated and the L2 symmetric division never takes place (a retarded phenotype; Figure 6F). Another target of *lin-4* miRNA is *lin-28*, which encodes a cytoplasmic protein with zinc and cold shock RNA-binding domains. LIN-28 plays a crucial role in defining the L2 and the L3 fates by activating the heterochronic gene *hbl-1* and by downregulating the miRNA *let-7*. In *lin-28* mutants the symmetric L2 seam cell division is skipped and the L3 division occurs instead, followed by the L4 division and the adult fate transition. These precocious mutants end up with fewer seam cells given their premature differentiation. Three miRNAs from the *let-7* family (*miR-48*, *miR-84*, and *miR-241*) inhibit expression of the *hunchback* transcription factor HBL-1, which is another heterochronic regulator of the L2 fate. Once LIN-28 and HBL-1 levels are brought low due to downregulation by miRNAs, seam cells enter the L3 stage, which is determined by expression of LIN-41 (Figure 7B), a RING finger beta-box, coiled-coil (RBCC) protein. Subsequently, downregulation of LIN-41 expression by rising levels of miRNA *let-7* allows the transition

to the L4 stage and the accumulation of LIN-29, which was previously repressed by HBL-1 and LIN-41 (Figure 7). LIN-29, the most downstream heterochronic component of the pathway, is a major regulator of the transition of the seam cells to their adult fate and other larva-to-adult related events, which I will discuss in the next section. *lin-29* turned out to be of special relevance to my work given my primary interest in temporal regulation in the L4 stage.

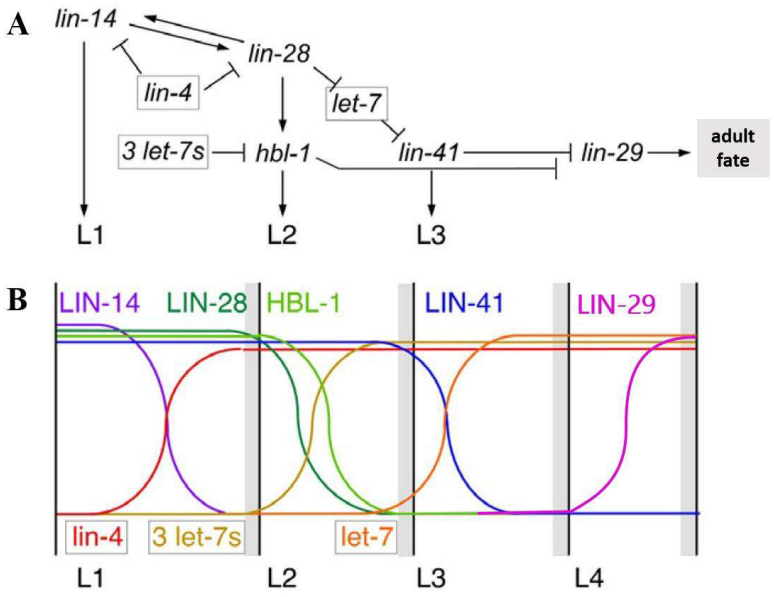


Figure 7. The heterochronic gene pathway. A) A model of the interactions between core members of the heterochronic pathway throughout development is displayed. miRNAs are represented in boxes B) Similarly, heterochronic protein and miRNA accumulation levels are shown over time (adapted from Rougvie & Moss, 2013)

V. LIN-29 and the larva-to-adult switch

V.a. The LIN-29 protein

lin-29 was first identified due to the egg-laying defective (Egl) phenotypes of its loss-of-function alleles *n333*, *n482*, *n546*, and *n836* (Trent et al., 1983); and soon after, it was found to be part of the heterochronic network acting as the temporal regulator of the seam cell adult fate (V Ambros & Horvitz, 1984). *lin-29* encodes a C2H2-class zinc finger transcription factor for which there are three isoforms, LIN-29*a*, *b* and *c*. The isoforms *a* and *b* are very similar and almost identical in sequence with 459 and 455 amino acids, respectively. Their transcripts have eleven exons and encode a total of five zinc fingers located in exons 6, 7 and 8 (Figure 8A). The isoform *c* also has the five-zinc finger DNA-binding domain, yet it uses a start codon in the 5th exon, thus the final product is a smaller protein of 317 amino acids (Figure 8A). All isoforms contribute to the known LIN-29 requirements and to the proper protein accumulation levels (Bettinger, Euling, & Rougvie, 1997; Rougvie & Ambros, 1995). Protein binding array analysis by Narasimhan et al. showed that the LIN-29 DNA-binding domain binds sequences with five to six consecutive As (Narasimhan et al., 2015). We later showed (Chapter 2) that such sites are present in the promoter regions of several LIN-29-regulated cuticle collagen genes, that they bind the LIN-29 DNA binding domain *in vitro,* and that they are required for proper reporter gene expression *in vivo* (Abete-Luzi & Eisenmann, 2018). In addition to the zinc finger domain, all three LIN-29 isoforms contain a conserved, C-terminal R1 domain (Figure 8A) that interacts with the transcription cofactor MAB-10 (described below).

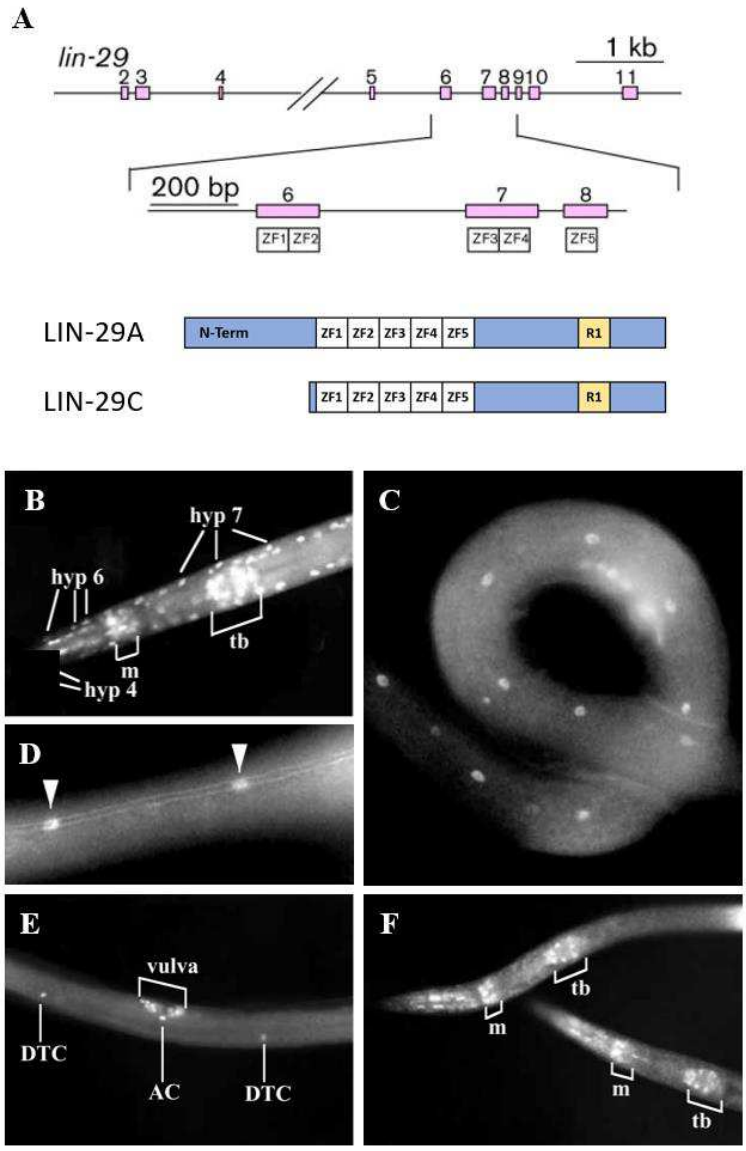


Figure 8. Zinc finger transcription factor LIN-29 and its expression pattern. A) Top: *lin-29* gene structure is shown with exons (pink boxes; Newman et al., 2000). Bottom: LIN-29 protein isoforms A and C with zinc finger and R1 domains. B-F) anti-LIN-29 immunostaining and epifluorescence show LIN-29 accumulation in the nuclei of: hyp cells in the L4 (B), seam cells both in the L4 (C) and in the adult (D), the AC and DTC in the L3 (E) and pharyngeal cells in the L2 (F) and in the L4 (B). Adapted from Bettinger & Rougvie (1996).

V.b. LIN-29 global expression patterns and functions

LIN-29 is known to be expressed in multiple tissues where it acts in the generation of structures needed to become a reproductively-capable adult. The temporal and spatial expression pattern of LIN-29 has been determined with an antibody that recognizes all isoforms as well as a translational reporter fusion (Bettinger et al., 1996; Harris & Horvitz, 2011). This transcription factor was first characterized for its role in hypodermal cells during the larva-to-adult switch. Naturally, its major accumulation pattern is in all hypodermal nuclei beginning in the early L4 in the seam cells, immediately followed by expression in hyp cells and the hyp syncytium (Figure 8B-D), and remaining through adulthood (Bettinger et al., 1997; Harris & Horvitz, 2011). LIN-29 also accumulates in the vulval cells, as well as in the anchor cell (AC) and distal tip cells (DTCs) of the gonad in the L3 stage (Figure 8E). Consistent with this expression pattern, LIN-29 was found to be required for vulva morphogenesis—as I will discuss in more detail in the following sections—and for the regulation of the gonad arm elongation, which is directed by the DTC. In this process, LIN-29 acts along with DAF-12 and BLMP-1 as a redundant regulator of *unc-5*, a gene necessary for the DTC turn during the phase 2 of DTC migration (Bettinger et al., 1997; Cecchetelli & Cram, 2017; Fielenbach et al., 2007; Anna P. Newman, Inoue, Wang, & Sternberg, 2000).

In males, LIN-29 is expressed in the linker cell (LC)—the cell that controls gonad migration in this sex—during the L3 stage and it disappears in the late L4, when this cell dies facilitating the connection of the gonad with the cloaca (Euling, Bettinger, & Rougvie,

1999) LIN-29 has been shown to be required for both proper gonad migration and the cell-autonomous programmed LC death (Abraham, Lu, & Shaham, 2007; Euling et al., 1999). Also in the male, LIN-29 is expressed in a cluster of cells of the preanal ganglion as well as in the B cell progeny—from which male spicules derive—starting in the L3 and persisting through adult stage. Once again, in agreement with its expression pattern, LIN-29 is required for proper development of the male reproductive apparatus, specifically for spicule differentiation (Euling et al., 1999).

Finally, it should be mentioned that LIN-29 shows steady expression in the pharynx of both males and hermaphrodites, beginning in the L1 stage and persisting through adulthood (Bettinger et al., 1996; Euling et al., 1999). However, to my knowledge, the role of LIN-29 in pharyngeal cells remains to be determined.

Because of the relevance to my work, in the following subsections I will review LIN-29 function in hypodermal cells (V.c.) as well as in the vulva (V.d.); I will introduce a recently found role in the intestine promoting vitellogenesis (V.e.), which is another process that occurs during adult life; and I will discuss the potential conservation of LIN-29 heterochronic regulation (V.f.).

V.c. LIN-29 function in the hypodermal cells

As mentioned above, LIN-29 is the most downstream effector of the heterochronic pathway and its role in the seam cell maturation is well- studied. After programmed rounds of cell division during each of the four larval stages (see section IV), seam cells terminally differentiate by 1) exiting the cell division cycle, 2) fusing with each other to form a long seam syncytium on each side of the body, and 3) secreting the alae—a cuticular ridge present only in the L1, dauer and adult stages. None of these events happen in *lin-29* mutants (Figure 9; V Ambros & Horvitz, 1984; Rougvie & Ambros, 1995). Subsequent work identified targets of LIN-29 that act in each of these processes.

LIN-29 was shown to regulate seam cell division cycle exit in the L4 stage through the downregulation of the cyclin-dependent kinase coding gene *cdk-1* (Ecsedi, Rausch, & Großhans, 2015), as well as the activation of *cki-1*, a member of the CIP/KIP family of cyclin-dependent kinase inhibitors which promotes cell cycle quiescence (Hong, Roy, & Ambros, 1998). LIN-29 also promotes seam cell fusion by activating the expression of the fusogen AFF-1, a cell-surface protein that mediates cell-to-cell fusion (Friedlander-shani & Podbilewicz, 2011). Another aspect of the seam cell larva-to-adult transition for which LIN-29 is required is the exit of the molting cycle. *lin-29* mutants complete their development, yet these animals undergo extra molts in the adult life(Harris & Horvitz, 2011; Rougvie & Ambros, 1995). The nuclear hormone receptors NHR-23 and NHR-25, both highly expressed in the hypodermis, were shown to be major regulators of molting (Asahina et al., 2000; Frand, Russel, & Ruvkun, 2005; Gissendanner & Sluder, 2000; G.

D. Hayes, Frand, & Ruvkun, 2006; Kostrouchova, Krause, Kostrouch, & Rall, 1998), and a model was proposed in which LIN-29 represses *nhr-25* gene expression thus preventing further molting (G. D. Hayes & Ruvkun, 2006). Supporting this model, Harris & Horvitz showed by single molecule FISH, that both transcription factors (NHR-23 and NHR-25) were upregulated in *lin-29* mutant adults (Harris & Horvitz, 2011). Another protein, the transcriptional co-factor MAB-10, was shown to co-express and interact with LIN-29 in the regulation of seam cell terminal differentiation and the exit of the molting cycle via the adult co-repression of *nhr-23* and *nhr-25* in males, and a more modest co-repression of *nhr-25* in hermaphrodites (Harris & Horvitz, 2011). Finally, LIN-29 plays a role in the secretion of the final adult cuticle since it appears to be a major activator of specific cuticle collagen genes for which whose normal expression is restricted to the time around the last molt: LIN-29 was first shown to regulate adult-specific cuticle collagen genes *col-19* and *col-7* (Liu, Kirch, & Ambros, 1995; Rougvie & Ambros, 1995), and later we showed that it also regulates a set of four L4-specific cuticle collagen genes (e.g. *col-38*, *col-49*, *col-63* and *col-138*; Abete-luzi & Eisenmann, 2018/Chapter 2). Consistent with this idea, our more recent genomic studies suggest that LIN-29 also activates 20 additional *col* genes that normally express in the L4 stage, as I will discuss in Chapter 3.

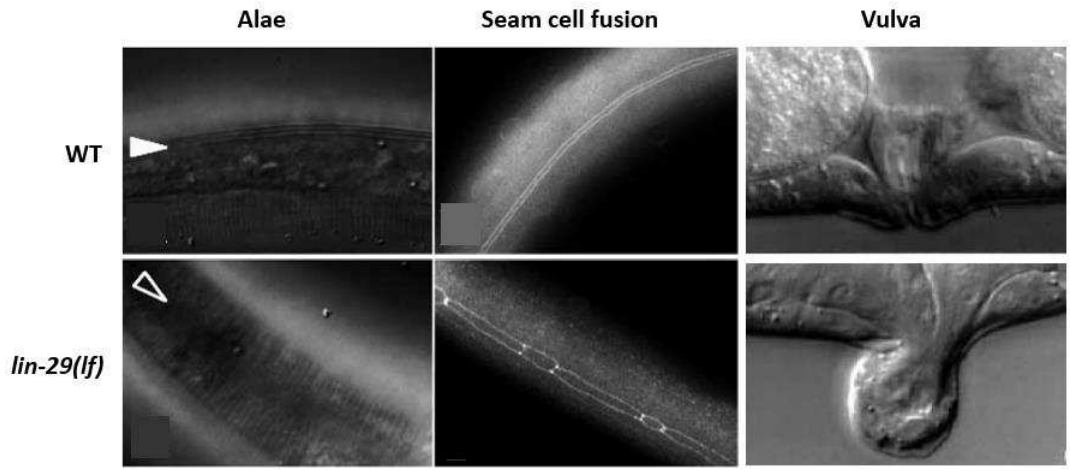


Figure 9. Three *lin-29* mutant phenotypes. Adult alae, seam cell junctions and vulva are shown for wild type and *lin-29* mutants as indicated. Arrows point to the alae (solid white arrowhead) or its absence (open white arrowhead). Alae and vulva are observed using Normarski contrast microscopy. Seam cell junctions are seen with anti-MH27 staining and epifluorescence. Adapted from Bettinger et al. (1997).

V.d. LIN-29 function in development of the egg-laying apparatus

Another developmental process that appears to be temporally regulated by the heterochronic pathway is the formation of the egg-laying apparatus. This system consists of a uterus connected to a specialized hypodermal opening through which embryos exit the body (the vulva). The formation of this system is a multistep process that requires the coordination of two separate cell lineages: the ventral uterine precursor (VU) cells and the ventral hypodermal (P cells), which give rise to the vulval precursor cells (VPCs). Several events must be coordinated during this process, including several specific cell-cell interactions, cell fate specification by induction and signaling cascades, as well as cell

division, cell migration and cell fusion events. The whole process starts in the early L3 and by the late L4, both uterus and vulva as well as the uterus-vulva-seam cell connections are fully formed. Like in most developmental processes, every step of vulva organogenesis is tightly regulated both spatially and temporally; and the failure of one or more of those steps will most likely result in vulva abnormalities, such as in heterochronic mutants. For example, the loss of *lin-14* or *lin-28* results in the early accumulation of LIN-29 in the VPCS in the L2 stage, leading to their precocious induction and division and vulva abnormalities in the adult (V Ambros & Horvitz, 1984; Ding, Spencer, Morita, & Han, 2005; Euling & Ambros, 1996; Gabriel D Hayes, Riedel, & Ruvkun, 2011). Vulval abnormalities in *lin-29* loss-of-function mutants include the protruding vulva phenotype (Pvl; Figure 9) and the consequent inability to lay eggs (egg-laying defective or Egl). Since LIN-29 was shown to regulate fusion of the seam cells, it was thought that perhaps LIN-29 was also responsible for the extensive cell fusions that take place during vulva development, however, vulval cell fusions were unaltered in *lin-29* loss-of-function animals (Bettinger et al., 1997). Subsequently it was shown that LIN-29 plays at least two roles in development of hermaphrodite egg-laying system: 1) induction of pi fates in uterus formation, and 2) expression of target genes in vulval cells.

● *lin-29 function in induction of uterine cell fates:* The anchor cell (AC) is probably the most important cell during vulva morphogenesis since it acts in the specification of both the VU and VPC lineages by expressing Notch and Ras ligands that induce cell fates in these cells that are needed to eventually form the vulval system (Gupta, Hanna-Rose, & Sternberg, 2012). One of these inductions allows the formation of the uterus. During the

late L3 stage, a subset of six uterine cells come in close contact with the AC and adopt the pi cell fate via activation of Notch signaling (Figure 10A). After one round of pi cell divisions, four of the twelve daughter cells adopt uv1 cell fates and the remaining eight cells fuse together and fuse with the AC to form the uterine-seam cell (*utse*; Figure 10B-D), an H-shaped syncytium that connects both vulval and uterine lumens to the body wall (Figure 10E).

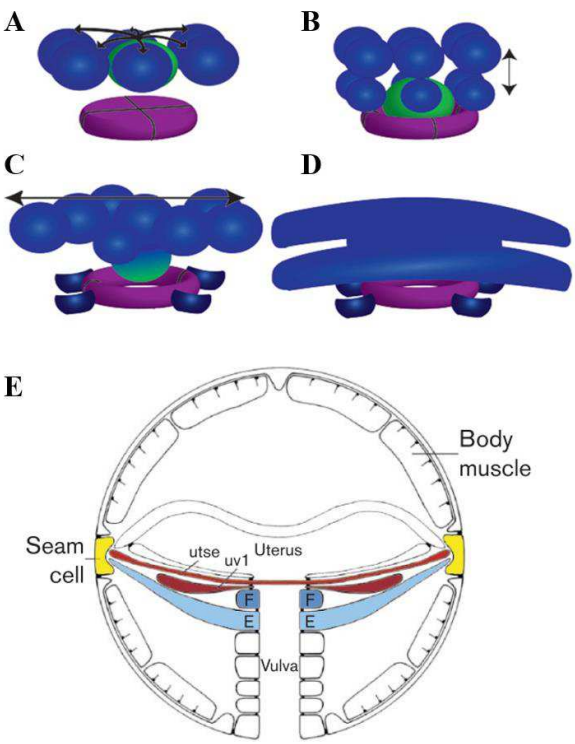


Figure 10. Morphogenesis of the vulva-uterine connection. The formation of the utse is depicted in four steps. A-D adapted from Gupta et al. (2012): A) The AC (green), which sits right on top of vulF (purple), induces six surrounding uterine cells to adopt pi fate (blue). After one round of division (B), pi cells align on one plane (C) and fuse with each other as well as with the AC to form the H-shaped utse syncytium (D). Differentiated cells and final vulva-uterus-seam connections in the late L4 as seen from a cross-section (E), from Newman et al. (2000)

LIN-29 is required for the AC expression of *lag-2,* which encodes the Notch pathway ligand that binds Notch-like receptor LIN-12 on the uterine cells to mediate pi cell fate

induction during the L3 (Gupta et al., 2012; Anna P. Newman et al., 2000). Part of this inductive signaling involves the activation of LIN-11 and EGL-13, two transcription factors required for fusion events and the utse formation (Figure 11; Hanna-Rose & Han, 1999; A P Newman, Acton, Hartwieg, Horvitz, & Sternberg, 1999).

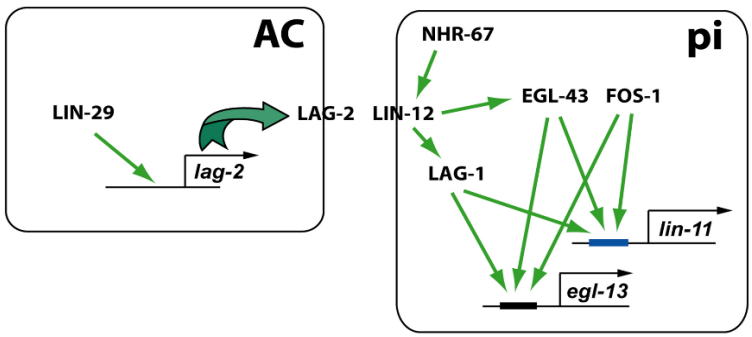


Figure 11. LIN-29 activation of Notch ligand *lag-2* in the AC promotes pi cell fate induction. Model of the gene interactions downstream of LIN-29 activation of Notch signaling via LAG-2 in the AC-mediated induction of uterine cells to adopt and maintain the pi fate. From Gupta et al. (2012).

● *lin-29 regulates gene expression in vulval cells in the L4:* As shown in Figure 12, six hypodermal cells (P3.p – P8.p), present on the ventral side of animal and known as VPCs, respond to inductive signals and divide to generate 22 cells in the early L3 stage. Then, during the L4 stage, the 22 cells organize themselves into seven ring-shaped cells—called vulA through F—which stack to form the vulva opening and attach it to the uterus. By the late L4, the vulval-uterine apparatus is fully formed (For further detail on vulva development see review by Gupta et al., 2012).

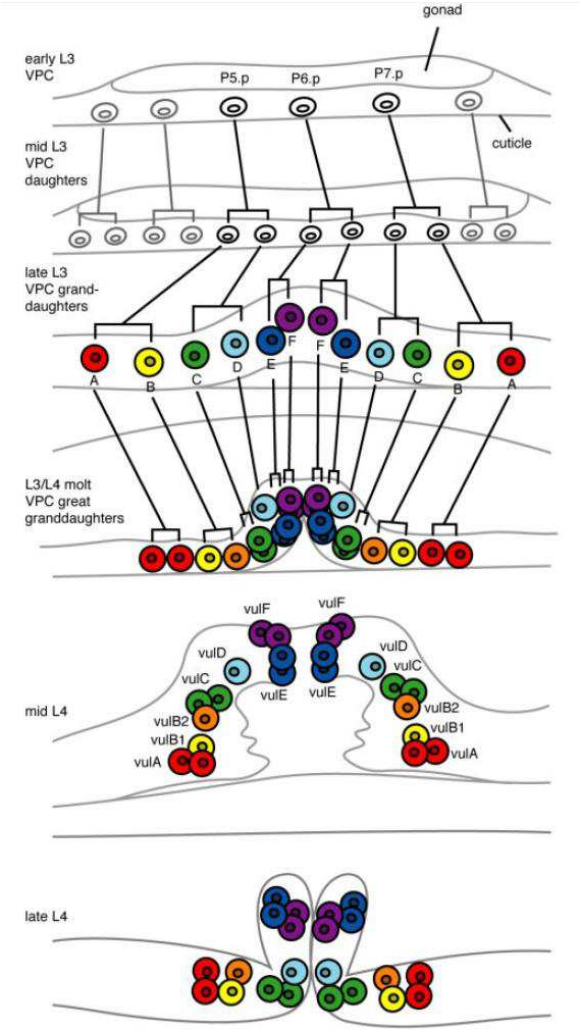


Figure 12. Vulval development through the L3 and the L4 stages.

Shown here is a temporal sequence (from top to bottom) of cell divisions and inductions that give rise to the vulva apparatus (as seen from longitudinal sections, ventral side at the bottom). After two rounds of division, vulva precursor cells (VPCs) become vulval cells, vulA through F (color coded). Further division and fusion events result in a stack of toroidal cells and a vulval lumen characteristic of the mid L4 stage. The eversion and narrowing of this lumen is the last step towards the adult morphology. From Ririe, Fernandes, & Sternberg (2008).

LIN-29 is also expressed in vulval cells (Bettinger et al., 1997), where it is required for the L4 expression of several genes in a cell-specific manner, such as *egl-17* in vulC and vulD, *ceh-2* in vulC, *zmp-1* in vulD and vulE (Inoue et al., 2002; Inoue, Wang, Ririe, Fernandes, & Sternberg, 2005) and *lin-11* in all vulA-D vulval cells (Ririe et al., 2008). In the absence of LIN-29, expression levels of *lin-11* are significantly reduced yet not completely abolished, suggesting redundancy of upstream regulators for this gene. LIN-11 is a LIM

homeodomain-type of protein and a major vulval transcriptional regulator along with EGL-38, COG-1, and NHR-67 (Ririe et al., 2008). All these transcription factors, their targets (e.g. *zmp-1*, *egl-17*, *lin-3*) and their interactions form a gene regulatory network that is responsible for the specification of vulval cell fates; LIN-29 is part of this network and may contribute to the temporal component of such specifications (Gupta et al., 2012; Inoue et al., 2005; Ririe et al., 2008).

V.e. LIN-29 and vitellogenesis

A more recent and surprising discovery includes the requirement of LIN-29 for the transport of lipid from the intestine to the germline, in a process called *vitellogenesis*. One crucial step towards the adult reproductive life as an egg-laying organism is the initiation of the yolk deposition or vitellogenesis, a process that requires fat mobilization from the intestine to oocytes in the gonad. This process is facilitated by *vitellogenins*, low-density lipoproteins that can bind and transport lipids and lipid-like molecules, such as cholesterol, and deposit them into the oocytes (Kimble & Sharrock, 1983; Spieth & Blumenthal, 1985). In *C. elegans,* vitellogenins are encoded by a family of six *vit* genes (*vit-1* to *vit-6*) and their spatiotemporal expression pattern is restricted to the intestine and to the adult stage in hermaphrodites; *vit* genes are repressed by the sexual development regulator MAB-3 in males (Kimble & Sharrock, 1983; Spieth & Blumenthal, 1985; Yi & Zarkower, 1999). *vit* activation was first shown to be controlled by insulin-like growth factor (IGF) signaling, which in *C. elegans* is mediated by the IGF receptor DAF-2/IGFR and its downstream

effectors: the kinases AGE-1/PI3K, AKT-1 and SGK-1, and the transcription factor DAF-16/FoxO (DePina et al., 2011; Halaschek-Wiener et al., 2005; Kimura, Tissenbaum, Liu, & Ruvkun, 1997; Morris, Tissenbaum, & Ruvkun, 1996; Murphy et al., 2003). This pathway is a key player in the regulation of major life events in the worm; it mediates changes in metabolism, ageing, reproduction, and the dauer arrest in response to environmental cues and the availability of resources (Baumeister, Schaffitzel, & Hertweck, 2006; Gems et al., 1998; Kenyon, Chang, Gensch, Rudner, & Tabtiang, 1993; Kimura et al., 1997; McElwee, Bubb, & Thomas, 2003; Murphy et al., 2003). As a general rule, egg production is an energy-consuming process and it contributes to the trade-off between reproduction and survival. Consistent with this idea, Murphy et al. showed that silencing *vit* gene expression results in longer lifespan in *C. elegans* (Murphy et al., 2003). Thus, it is not surprising to find vitellogenesis to be a target of IGF signaling integrative regulation.

In more recent studies, additional *vit* regulators were identified. The homeodomain protein UNC-62/Homothorax was shown to modulate lifespan specifically from the intestine and the hypodermis, and to directly bind and activate *vit* genes in the adult (Van Nostrand, Sánchez-Blanco, Wu, Nguyen, & Kim, 2013). Interestingly, in a separate RNAi screen for *vit* regulators more genes were found; of particular interest were the heterochronic miRNAs *lin-4* and *let-7*, which turned out to activate *vit* expression through their downstream regulator, LIN-29 (Dowen, Breen, Tullius, Conery, & Ruvkun, 2016). Both tissue-specific RNAi and rescue experiments showed that LIN-29 regulates *vit* genes from the hypodermis (Figure 13A and B), suggesting for the first time the existence of a non-autonomous regulatory mechanism by which the heterochronic pathway affects the timing of gene

expression in the intestine (Dowen et al., 2016). In the same study, the authors identified the intestinal *vit* regulators downstream of LIN-29 as members of the mTOR signaling pathway, the receptor LET-363/mTORC2 and the kinase SKG-1, as well as the transcriptional repressor, PQM-1. The mTOR pathway is a major integrator of intracellular and extracellular signals (i.e. IGF, EGF) and participates in the regulation of cell growth, cell metabolism, cell proliferation and survival, which makes it highly relevant in the stem-cell, cancer and ageing fields (see review by Laplante & Sabatini, 2009). In the worm, the zinc finger transcription factor PQM-1 was found to be responsive to this pathway and a downstream negative regulator of *vit* gene expression (Dowen et al., 2016). In wild type animals, PQM-1 is found in the nuclei of intestinal cells during larval stages but localizes the cytoplasm in adult life (Dowen et al., 2016; Tepper et al., 2013). The nuclear localization of PQM-1, as well as its activity, was first shown to be controlled by IGF signaling and antagonized by the insulin signaling downstream effector DAF-16/FoxO (Tepper et al., 2013). Interestingly, knockdown of *lin-29* or *skg-1* results in nuclear PQM-1 throughout adulthood and the continued repression of *vit* genes (Figure 13; Dowen et al., 2016). Dowen et al. proposed a model to explain *vit* gene regulation as the result of two independent yet collaborative mechanisms (Figure 14). From the hypodermis, LIN-29 activates a signal that diffuses to the intestine and activates the mTORC2 pathway, which in parallel with IGF signaling, represses PQM-1 and allows *vit* expression (Dowen et al., 2016). The LIN-29-dependent signal mediating the intestinal activation of the mTOR pathway remains unknown.

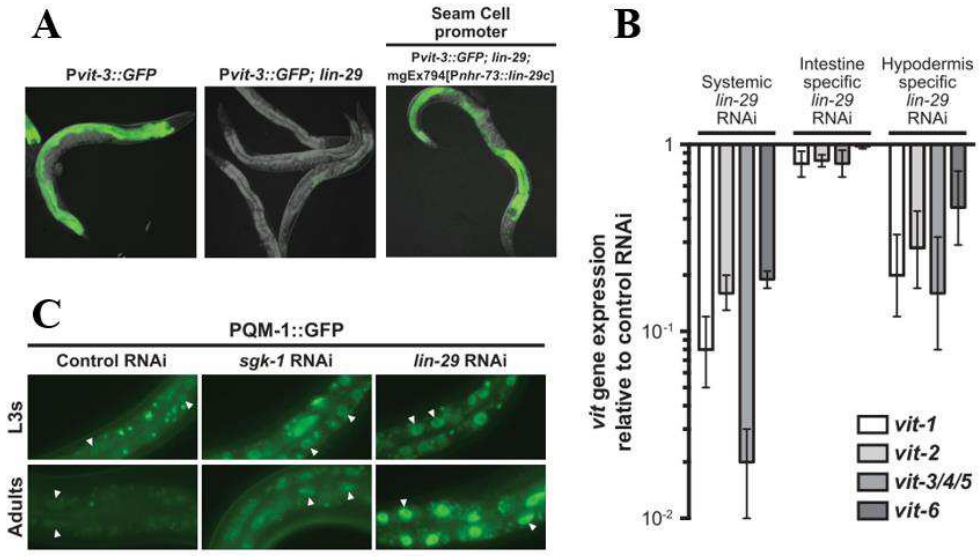


Figure 13. Hypodermal LIN-29 regulates *vit* gene expression through mTOR-mediated PQM-1 inactivation. A) Epifluorescence shows the expression of a *vit-3* GFP transcriptional reporter being abrogated in *lin-29*(*lf*), yet recovered when adding a seam cell specific *lin-29* rescue construct. B) qPCR results reveal endogenous *vit* expression is compromised in animals treated with either systemic or hypodermic-specific *lin-29* RNAi. C) Epifluorescence of PQM-1::GFP protein fusion shows *skg-1* and *lin-29* are required for it to exit intestinal nuclei in adults. From Dowen et al. (2016)

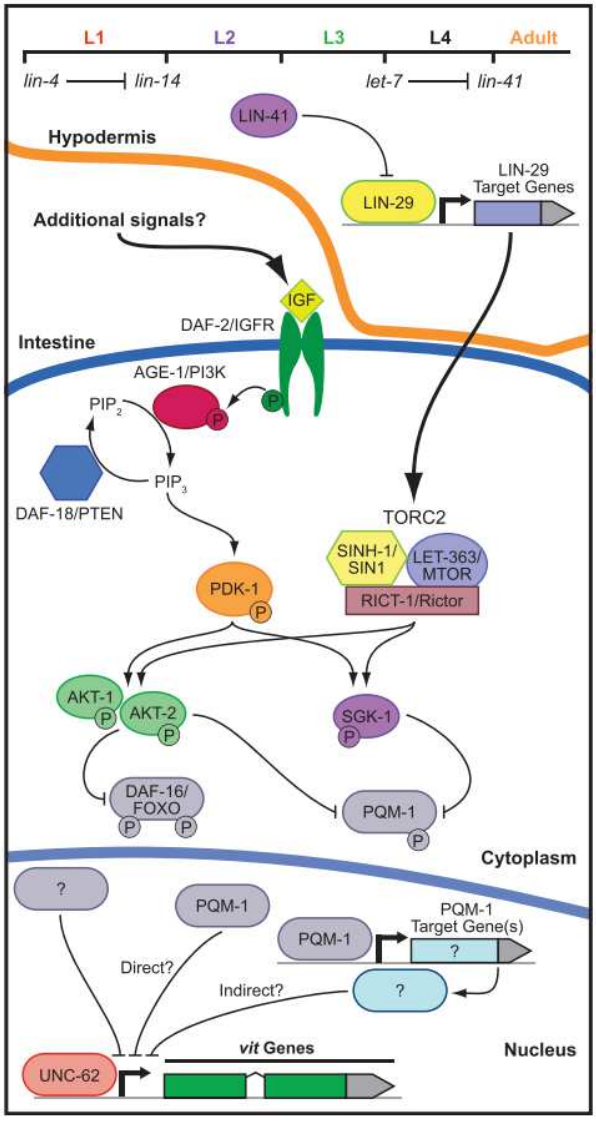


Figure 14. Dowen *et al.* model for the regulation of *vit* regulation. Two independent mechanisms cooperate to control *vit* gene expression: the DAF-2 insulin signaling pathway and the LIN-29 triggered mTORC2/PQM-1 pathway. From Dowen et al. (2016).

In addition, given the role of LIN-29 and SKG-1 in fat mobilization via *vit* gene regulation, the authors investigated whether these genes also affect body fat accumulation by doing Oil Red O staining (which allows the quantification of lipids; Wählby et al., 2012, 2014), and found that in both *lin-29* and *skg-1* mutants, total fat content was decreased compared to wild type values (Figure 15; Dowen et al., 2016). The mechanisms causing such body fat reduction when *lin-29* function is reduced remain to be determined.

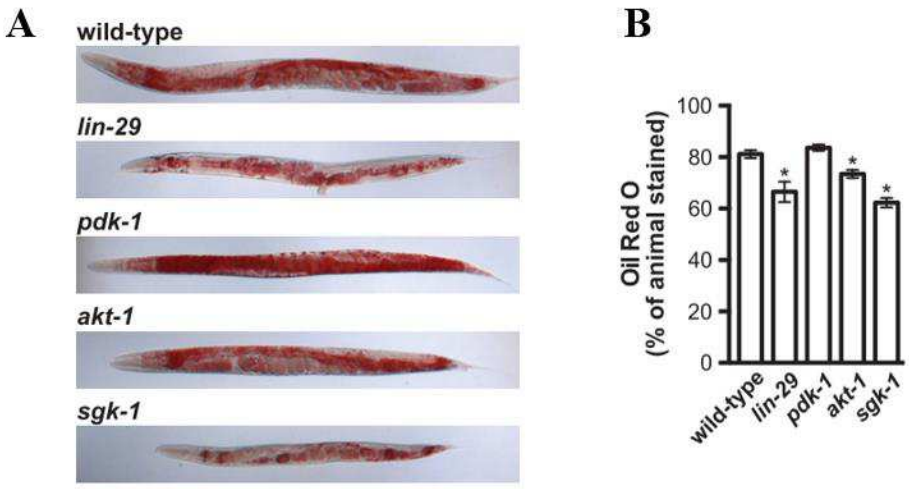


Figure 15. LIN-29 and mTORC2 components are required for body fat accumulation. Images of adults stained with Oil Red O (A) and the staining percentages (B) are shown for the indicated genotypes. From Dowen et al. (2016).

All in all, the fact that LIN-29 can affect intestinal fat dynamics revealed 1) a new aspect of heterochronic regulation, from one tissue to another, and 2) a potential role for LIN-29 in the balance between energy storage, energy expenditure and fertility during the transition to adulthood. In light of the latter idea, there are three questions I would like to answer. Could LIN-29 be a temporal regulator of other metabolic decisions, besides promoting fertility? Could LIN-29 influence longevity? Could this type of regulation be conserved among other organisms?

V.f. Conservation of heterochronic genes, and EGR1 as the mammalian homolog of LIN-29

While the heterochronic pathway as described above for *C. elegans* seems to be restricted to nematodes, many of the core heterochronic genes and their interactions are found in larger animals as well. As defined based on their seed sequences, both miRNAs *lin-4* and *let-7* are also present in flies and vertebrates; the *lin-4* family is called *mir-125*, whereas *let-7*, the most conserved and similar in all species, is the only miRNA that kept the original worm name as *let-7* in most species (Pasquinelli et al., 2000; Roush & Slack, 2008). Both mammalian *mir-125* and *let-7* play major roles in stem cells, in regulating the transition from pluripotency to differentiated states (Houbaviy, Murray, & Sharp, 2003). In humans, *let-7* has 9 mature forms encoded by 12 loci and it is well known for its implications in cancer given its role as a tumor suppressor (see review by Su, Chen, Johansson, & Kuo, 2012). LIN-28 also has homologs in vertebrates. For example, in Xenopus, Lin28 is required for proper temporal expression of genes involved in mesoderm specification in the early embryo (Faas et al., 2013). Also in this species, Lin28 regulates developmental transitions by inhibiting thyroid hormone and negatively regulating metamorphosis (Faunes, Gundermann, Muñoz, Bruno, & Larraín, 2017). As for humans, LIN28 is known as a pluripotency factor which, like in *C. elegans*, represses *let-7* biogenesis and is downregulated by *lin-4/mir-125* in stem cells (Nam, Chen, Gregory, Chou, & Sliz, 2011; M. A. Newman, Thomson, & Hammond, 2008; Rybak et al., 2008; Wu & Belasco, 2005). In fact, LIN28 and miRNAs are key elements of a major stem cell regulatory network in which different pathways interact to determine cell growth, cell cycle and even cell metabolism during stem cell renewal (see review by Shyh-Chang & Daley, 2013).

Moreover, LIN28, in conjunction with other transcription factors, is capable of reprogramming differentiated human cells into induced pluripotent stem cells, a fact that made this gene very attractive over the past decade (Yu et al., 2007). LIN-41 also shows conservation among species. This gene is a member of the TRIM-NHL family, and both TRIM proteins as well as their downregulation by *let-7*, are conserved in flies, worms, zebrafish, chick, mice and humans (see review by Ecsedi & Großhans, 2013). Like LIN-28, LIN-41/TRIM71 has "pro-stem cell" and reprogramming capacities and is downregulated by "pro-differentiation" regulator *let-7* in mammals (Worringer et al., 2014).

Interestingly, in the case of LIN-29 the proposed mammalian homolog is the Early Growth Response-1 (EGR1; Harris & Horvitz, 2011), also known as KROX24, NGFIA, and ZIF268. EGR1 is a transcription factor with three C2H2 zinc fingers, known for its rapid activation in response to growth factors, cytokines and stress signals. It is involved in the regulation of several inflammatory genes as well as growth factors, such as insulin-like growth factor-II (IGF-II) and transforming growth factor beta 1 (TGF-ß1), cell cycle regulators (i.e. retinoblastoma gene Rb and cyclin D1), and the luteinizing hormone beta-subunit (see EGR1 review by Pagel & Deindl, 2011). One important aspect of EGR1 is that, in both mice and humans, the expression of this factor specifically in white adipose tissue is highly correlated with diet-induced obesity and obesity-related pathologies such as insulin resistance and hyperlipidemia; whereas the loss of Egr-1 in mice protects those animals from such pathologies (Zhang et al., 2013). It turns out that these phenotypes are most likely doe to a role of Egr-1 as a negative regulator of energy expenditure in

adipocytes (Zhang et al., 2013), a fact that is particularly relevant to some of my findings on LIN-29 which I will discuss in Chapter 3.

While the sequence similarity between LIN-29 and EGR1 is rather low (41% identity), the resemblance between their interactions with other proteins and their regulation suggests homology at least to some extent. LIN-29 interacts with MAB-10, a NGFI-A-binding (NAB) protein in the same way that EGR proteins interact with mammalian NAB proteins. MAB-10, like other NAB proteins, encodes a transcriptional co-factor with two canonical NAB-conserved domains, NCD1 and NCD2, whereas LIN-29 possesses a NAB-interaction domain that is conserved among EGR and EGR-like proteins from humans and *Drosophila*, respectively (R1 domain, see Figure 8A; Harris & Horvitz 2011). Curiously, like LIN-29 and MAB-10 interact to regulate the terminal differentiation of seam cells in the worm, mammalian EGR and NAB proteins interact to activate or repress genes required for cell differentiation in several cell types including chondrocytes, Schwann cells, keratinocytes and macrophages (Harris & Horvitz, 2011; Laslo et al., 2006; Le et al., 2005; Sevetson, Svaren, & Milbrandt, 2000; Swirnoff et al., 1998). These interactions—in both, worms and mammals—can result in either co-activation or co-repression, apparently depending on the cellular context (Collins, Wolfraim, Drake, Horton, & Powell, 2006; Harris & Horvitz, 2011; Kumbrink, Kirsch, & Johnson, 2010; Sevetson et al., 2000). EGR1 was shown to activate expression of its co-factor NAB2 through eleven EGR-binding sites present in the *Nab2* promoter (Kumbrink, Gerlinger, & Johnson, 2005). Harris and Horvitz (2011) reported that LIN-29 does not regulate *mab-10* at the transcriptional level. However, as described in Chapter 3, our work along with work done by Hunter et al., indicates that

LIN-29 does activate *mab-10* transcription (Hunter et al., 2013; Chapter 3). Additional evidence supporting the idea of EGR1 as the LIN-29 mammalian homolog comes from recent work showing conservation of LIN-29/EGR1 negative regulation by LIN-41 (TRIM71 in mammals). Strikingly, in both *C. elegans* and human cells, LIN-41/TRIM71 downregulates LIN-29/EGR1 protein by translational repression (Aeschimann et al., 2017; Worringer et al., 2014). A model for the conservation of a LIN-29/EGR mediated pathway first proposed by Harris & Horvitz (2011) is summarized in Figure 16.

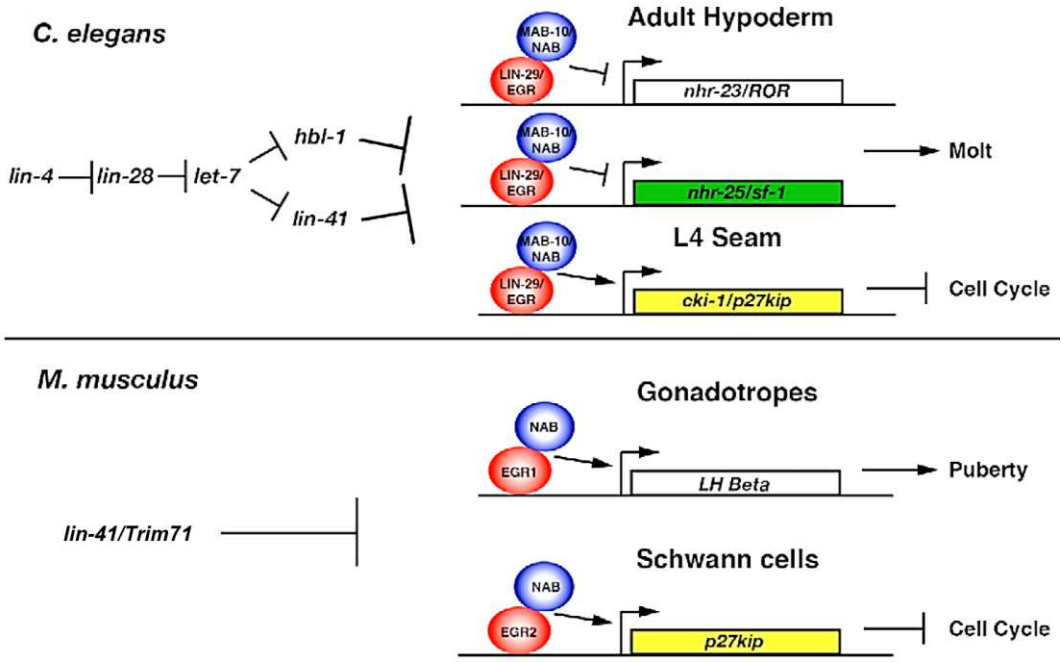


Figure 16. Model of conservation between pathways that regulate terminal differentiation and the onset of adulthood. Regulated by the heterochronic genes, LIN-29 and NAB co-factor MAB-10 repress molting genes and promote cell cycle inhibitor *cki-1/p27kip* (top). In mice, LIN-41/TRIM71 regulates EGR proteins, which together with NAB co-factor activate β-luteinizing hormone and the *cki-1* homolog, *p27kip* (bottom). Adapted from Harris & Horvitz (2011).

VI. Summary

Broadly, one of the goals of the Eisenmann lab is to better understand cell fate specification during animal development by studying the spatial and temporal regulation of gene expression. My work involves the latter aspect of this process. While at first my work was focused on the elucidation of the temporal regulation of *C. elegans* larval development by identifying key regulators of stage-specific cuticle collagen (*col*) genes, the results I obtained for the L4-stage *col* genes revealed conspicuous regulation by the heterochronic effector LIN-29 and provided me with a new venue for investigation.

As described in the sections above, findings on the conserved heterochronic miRNAs, *lin-4*/*mir-125* and *let-7*, and the proteins LIN-28 and LIN-41, have had major impact in developmental studies as well as in cancer research, making these genes targets of considerable attention in comparison to other members of the heterochronic pathway, such as LIN-29. Nevertheless, recently revealed roles for LIN-29 in the temporal control of events pertinent to *C. elegans* puberty (the transition to adulthood), and its plausible relation to a conserved developmental regulatory network, motivated me to further investigate this gene and its downstream targets.

● **In Chapter 2**, I will discuss my research on *col* gene expression and the investigation of putative *cis*-regulatory elements as well as *trans*-acting factors for a subset of L4-specific

cols, work that was published on March 31st of 2018 in *Genesis* with the title **'***Regulation of C. elegans L4 cuticle collagen genes by the heterochronic protein LIN-29'*

● **In Chapter 3**, I will describe the work I have done on *C. elegans* to further explore the developmental impact of LIN-29 on the larva-to-adult transition as well as the search for LIN-29 target genes using a *lin-29* gain-of-function approach. The manuscript is currently in preparation for submission.

● **In Chapter 4**, I will provide a final discussion of the overall results in terms of their relevance as well as future directions for each project.

References

Abete-Luzi, P., & Eisenmann, D. M. (2018, April 23). Regulation of C. elegans L4 cuticle collagen genes by the heterochronic protein LIN-29. *Genesis*. Wiley-Blackwell. https://doi.org/10.1002/dvg.23106

Abraham, M. C., Lu, Y., & Shaham, S. (2007). A morphologically conserved nonapoptotic program promotes linker cell death in Caenorhabditis elegans. *Developmental Cell*, *12*(1), 73–86. https://doi.org/10.1016/j.devcel.2006.11.012

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., & Großhans, H. (2017). LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms. *Molecular Cell*, *65*(3), 476–489.e4. https://doi.org/10.1016/j.molcel.2016.12.010

Altun, Z. F., & Hall, D. H. (2005). Hypodermis. In *Handbook of C. elegans Anatomy*. WormAtlas. Retrieved from http://www.wormatlas.org/ver1/handbook/hypodermis/hypodermis.htm

Altun, Z. F., & Hall, D. H. (2009). Introduction to C. elegans anatomy. In *WormAtlas.*

Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. *Cell*, *57*(1), 49–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2702689

Ambros, V., & Horvitz, H. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. *Science*, *226*(4673), 409–416. https://doi.org/10.1126/science.6494891

Ambros, V., & Horvitz, H. R. (1987). The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. *Genes & Development*, *1*(4), 398–414. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3678829

Ambros, V., & Ruvkun, G. (2018). Recent Molecular Genetic Explorations of Caenorhabditis elegans MicroRNAs. *Genetics*, *209*(3), 651–673. https://doi.org/10.1534/genetics.118.300291

Antoshechkin, I., & Sternberg, P. W. (2007). The versatile worm: genetic and genomic resources for Caenorhabditis elegans research. *Nature Reviews Genetics*, *8*(7), 518–532. https://doi.org/10.1038/nrg2105

Asahina, M., Ishihara, T., Jindra, M., Kohara, Y., Katsura, I., & Hirose, S. (2000). The conserved nuclear receptor Ftz-F1 is required for embryogenesis, moulting and reproduction in Caenorhabditis elegans. *Genes to Cells*, *5*(9), 711–723. https://doi.org/10.1046/j.1365-2443.2000.00361.x

Baumeister, R., Schaffitzel, E., & Hertweck, M. (2006). Endocrine signaling in Caenorhabditis elegans controls stress response and longevity. *The Journal of Endocrinology*, *190*(2), 191–202. https://doi.org/10.1677/joe.1.06856

Bettinger, J. C., Euling, S., & Rougvie, A. E. (1997). The terminal differentiation factor LIN-29 is required for proper vulvar morphogenesis and egg laying in Caenorhabditis elegans. *Development*, *124*(21), 4333–4342.

Bettinger, J. C., Lee, K., & Rougvie, a E. (1996). Stage-specific accumulation of the terminal differentiation factor LIN-29 during Caenorhabditis elegans development.

Development, *122*(8), 2517–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8756296

Cai, Y., Yu, X., Hu, S., & Yu, J. (2009). A brief review on the mechanisms of miRNA regulation. *Genomics, Proteomics & Bioinformatics*, *7*(4), 147–54. https://doi.org/10.1016/S1672-0229(08)60044-3

Cecchetelli, A. D., & Cram, E. J. (2017). Regulating distal tip cell migration in space and time. *Mechanisms of Development*, *148*, 11–17. https://doi.org/10.1016/J.MOD.2017.04.003

Collins, S., Wolfraim, L. A., Drake, C. G., Horton, M. R., & Powell, J. D. (2006). Cutting Edge: TCR-induced NAB2 enhances T cell function by coactivating IL-2 transcription. *Journal of Immunology (Baltimore, Md. : 1950)*, *177*(12), 8301–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17142725

Cox, G. N., Kusch, M., & Edgar, R. S. (1981). Cuticle of Caenorhabditis elegans: its isolation and partial characterization. *The Journal of Cell Biology*, *90*(1), 7–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7251677

DePina, A. S., Iser, W. B., Park, S.-S., Maudsley, S., Wilson, M. A., & Wolkow, C. A. (2011). Regulation of Caenorhabditis elegans vitellogenesis by DAF-2/IIS through separable transcriptional and posttranscriptional mechanisms. *BMC Physiology*, *11*, 11. https://doi.org/10.1186/1472-6793-11-11

Di Cara, F., & King-Jones, K. (2013). How Clocks and Hormones Act in Concert to Control the Timing of Insect Development. *Current Topics in Developmental Biology*, *105*, 1–36. https://doi.org/10.1016/B978-0-12-396968-2.00001-4

Dickinson, D. J., & Goldstein, B. (2016). CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering. *Genetics*, *202*(3), 885–901. https://doi.org/10.1534/genetics.115.182162

Ding, L., Spencer, A., Morita, K., & Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans. *Molecular Cell*, *19*(4), 437–47. https://doi.org/10.1016/j.molcel.2005.07.013

Dowen, R. H., Breen, P. C., Tullius, T., Conery, A. L., & Ruvkun, G. (2016). A microRNA program in the C. elegans hypodermis couples to intestinal mTORC2/PQM-1 signaling to modulate fat transport. *Genes and Development*, *30*(13), 1515–1528. https://doi.org/10.1101/gad.283895.116

Ecsedi, M., & Großhans, H. (2013). Lin-41/Trim71: Emancipation of a miRNA target. *Genes and Development*. https://doi.org/10.1101/gad.207266.112

Ecsedi, M., Rausch, M., & Großhans, H. (2015). The let-7 microRNA directs vulval development through a single target. *Developmental Cell*, *32*(3), 335–44. https://doi.org/10.1016/j.devcel.2014.12.018

Euling, S., & Ambros, V. (1996). Heterochronic genes control cell cycle progress and developmental competence of C. elegans vulva precursor cells. *Cell*, *84*(5), 667–76. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8625405

Euling, S., Bettinger, J. C., & Rougvie, a E. (1999). The LIN-29 transcription factor is required for proper morphogenesis of the Caenorhabditis elegans male tail. *Developmental Biology*, *206*(2), 142–56. https://doi.org/10.1006/dbio.1998.9063

Faas, L., Warrander, F. C., Maguire, R., Ramsbottom, S. A., Quinn, D., Genever, P., & Isaacs, H. V. (2013). Lin28 proteins are required for germ layer specification in Xenopus. *Development (Cambridge, England)*, *140*(5), 976–86. https://doi.org/10.1242/dev.089797

Faunes, F., Gundermann, D. G., Muñoz, R., Bruno, R., & Larraín, J. (2017). The heterochronic gene Lin28 regulates amphibian metamorphosis through disturbance of thyroid hormone function. *Developmental Biology*, *425*(2), 142–151. https://doi.org/10.1016/j.ydbio.2017.03.026

Fielenbach, N., Guardavaccaro, D., Neubert, K., Chan, T., Li, D., Feng, Q., … Antebi, A. (2007). DRE-1: An Evolutionarily Conserved F Box Protein that Regulates C. elegans Developmental Age. *Developmental Cell*, *12*(3), 443–455. https://doi.org/10.1016/J.DEVCEL.2007.01.018

Frand, A. R., Russel, S., & Ruvkun, G. (2005). Functional Genomic Analysis of C. elegans Molting. *PLoS Biology*, *3*(10), e312. https://doi.org/10.1371/journal.pbio.0030312

Friedlander-shani, L., & Podbilewicz, B. (2011). Cell Fusion in Health and Disease, *950*, 5–11. https://doi.org/10.1007/978-94-007-0782-5

Gems, D., Sutton, A. J., Sundermeyer, M. L., Albert, P. S., King, K. V, Edgley, M. L., … Riddle, D. L. (1998). Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in Caenorhabditis elegans. *Genetics*, *150*(1), 129–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9725835

Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., …

Trifonov, T. (2010). Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. *Science*, *330*(6012), 1775 LP-1787. https://doi.org/10.2436/20.7010.01.53

Gissendanner, C. R., & Sluder, A. E. (2000). nhr-25, the Caenorhabditis elegans Ortholog of ftz-f1, Is Required for Epidermal and Somatic Gonad Development. *Developmental Biology*, *221*(1), 259–272. https://doi.org/10.1006/dbio.2000.9679

Gupta, B. P., Hanna-Rose, W., & Sternberg, P. W. (2012). Morphogenesis of the vulva and the vulval-uterine connection*. *WormBook The C. elegans Research Community*, *11521*(10). https://doi.org/10.1895/wormbook.1.152.1

Halaschek-Wiener, J., Khattra, J. S., McKay, S., Pouzyrev, A., Stott, J. M., Yang, G. S., … Riddle, D. L. (2005). Analysis of long-lived C. elegans daf-2 mutants using serial analysis of gene expression. *Genome Research*, *15*(5), 603–615. https://doi.org/10.1101/gr.3274805

Hanna-Rose, W., & Han, M. (1999). COG-2, a sox domain protein necessary for establishing a functional vulval-uterine connection in Caenorhabditis elegans. *Development (Cambridge, England)*, *126*(1), 169–79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9834196

Harris, D. T., & Horvitz, H. R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. *Development*, *138*(18), 4051–4062. https://doi.org/10.1242/dev.065417

Hayes, G. D., Frand, A. R., & Ruvkun, G. (2006). The mir-84 and let-7 paralogous microRNA genes of Caenorhabditis elegans direct the cessation of molting via the

conserved nuclear hormone receptors NHR-23 and NHR-25. *Development*, *133*(23), 4631–4641. https://doi.org/10.1242/dev.02655

Hayes, G. D., Riedel, C. G., & Ruvkun, G. (2011). The Caenorhabditis elegans SOMI-1 zinc finger protein and SWI/SNF promote regulation of development by the mir-84 microRNA. *Genes & Development*, *25*(19), 2079–92. https://doi.org/10.1101/gad.17153811

Hayes, G. D., & Ruvkun, G. (2006). Misexpression of the Caenorhabditis elegans miRNA let-7 is sufficient to drive developmental programs. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 71, pp. 21–27). https://doi.org/10.1101/sqb.2006.71.018

Hong, Y., Roy, R., & Ambros, V. (1998). Developmental regulation of a cyclin-dependent kinase inhibitor controls postembryonic cell cycle progression in Caenorhabditis elegans, *3597*, 3585–3597.

Houbaviy, H. B., Murray, M. F., & Sharp, P. A. (2003). Embryonic stem cell-specific MicroRNAs. *Developmental Cell*, *5*(2), 351–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12919684

Hu, P. J. (2008). *Dauer*. WormBook , ed. The C. elegans Research Community. https://doi.org//10.1895/wormbook.1.144.1

Hunter, S. E., Finnegan, E. F., Zisoulis, D. G., Lovci, M. T., Melnik-Martinez, K. V., Yeo, G. W., & Pasquinelli, A. E. (2013). Functional Genomic Analysis of the let-7 Regulatory Network in Caenorhabditis elegans. *PLoS Genetics*, *9*(3). https://doi.org/10.1371/journal.pgen.1003353

Inoue, T., Sherwood, D. R., Aspöck, G., Butler, J. A., Gupta, B. P., Kirouac, M., … Sternberg, P. W. (2002). Gene expression markers for Caenorhabditis elegans vulval cells. *Gene Expression Patterns*, *2*(3–4), 235–241. https://doi.org/10.1016/S1567-133X(02)00055-8

Inoue, T., Wang, M., Ririe, T. O., Fernandes, J. S., & Sternberg, P. W. (2005). Transcriptional network underlying Caenorhabditis elegans vulval development. *Proc Natl Acad Sci U S A*, *102*(14), 4972–4977. https://doi.org/0408122102 [pii]\r10.1073/pnas.0408122102

Jackson, B. M., Abete-Luzi, P., Krause, M. W., & Eisenmann, D. M. (2014). Use of an Activated Beta-Catenin to Identify Wnt Pathway Target Genes in Caenorhabditis elegans , Including a Subset of Collagen Genes Expressed in Late Larval Development. *G3 Genes|Genomes|Genetics*, *4*(4), 733–747. https://doi.org/10.1534/g3.113.009522

Johnstone, I. L. (2000). Cuticle collagen genes. Expression in Caenorhabditis elegans. *Trends in Genetics : TIG*, *16*(1), 21–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10637627

Johnstone, I. L., & Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development. *The EMBO Journal*, *15*(14), 3633–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8670866

Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461–464. https://doi.org/10.1038/366461a0

Kimble, J., & Sharrock, W. J. (1983). Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans. *Developmental Biology*, *96*(1), 189–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6825952

Kimura, K. D., Tissenbaum, H. A., Liu, Y., & Ruvkun, G. (1997). daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science (New York, N.Y.)*, *277*(5328), 942–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9252323

Kostrouchova, M., Krause, M., Kostrouch, Z., & Rall, J. E. (1998). CHR3: a Caenorhabditis elegans orphan nuclear hormone receptor required for proper epidermal development and molting. *Development (Cambridge, England)*, *125*(9), 1617–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9521900

Kumbrink, J., Gerlinger, M., & Johnson, J. P. (2005). Egr-1 induces the expression of its corepressor Nab2 by activation of the Nab2 promoter thereby establishing a negative feedback loop. *Journal of Biological Chemistry*, *280*(52), 42785–42793. https://doi.org/10.1074/jbc.M511079200

Kumbrink, J., Kirsch, K. H., & Johnson, J. P. (2010). EGR1, EGR2, and EGR3 activate the expression of their coregulator NAB2 establishing a negative feedback loop in cells of neuroectodermal and epithelial origin. *Journal of Cellular Biochemistry*, *111*(1), 207–217. https://doi.org/10.1002/jcb.22690

Lall, S., Grün, D., Krek, A., Chen, K., Wang, Y.-L., Dewey, C. N., … Rajewsky, N. (2006). A genome-wide map of conserved microRNA targets in C. elegans. *Current Biology : CB*, *16*(5), 460–71. https://doi.org/10.1016/j.cub.2006.01.050

Laplante, M., & Sabatini, D. M. (2009). mTOR signaling at a glance. *Journal of Cell Science*, *122*(Pt 20), 3589–94. https://doi.org/10.1242/jcs.051011

Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H.-J., Sciammas, R., … Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, *126*(4), 755–66. https://doi.org/10.1016/j.cell.2006.06.052

Le, N., Nagarajan, R., Wang, J. Y. T., Svaren, J., LaPash, C., Araki, T., … Milbrandt, J. (2005). Nab proteins are essential for peripheral nervous system myelination. *Nature Neuroscience*, *8*(7), 932–940. https://doi.org/10.1038/nn1490

Lin, S.-Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., … Slack, F. J. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Developmental Cell*, *4*(5), 639–50. https://doi.org/10.1016/S1534-5807(03)00124-2

Lints, R., & Hall, D. H. (2009). Male introduction. In *WormAtlas.* https://doi.org/10.3908/wormatlas.2.1

Liu, Z., Kirch, S., & Ambros, V. (1995). The Caenorhabditis elegans heterochronic gene pathway controls stage-specific transcription of collagen genes. *Development (Cambridge, England)*, *121*(8), 2471–2478.

Martinez, N. J., Ow, M. C., Reece-Hoyes, J. S., Barrasa, M. I., Ambros, V. R., & Walhout, A. J. M. (2008). Genome-scale spatiotemporal analysis of Caenorhabditis elegans microRNA promoter activity. *Genome Research*, *18*(12), 2005–15. https://doi.org/10.1101/gr.083055.108

McElwee, J., Bubb, K., & Thomas, J. H. (2003). Transcriptional outputs of the Caenorhabditis elegans forkhead protein DAF-16. *Aging Cell*, *2*(2), 111–121. https://doi.org/10.1046/j.1474-9728.2003.00043.x

Morris, J. Z., Tissenbaum, H. A., & Ruvkun, G. (1996). A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in Caenorhabditis elegans. *Nature*, *382*(6591), 536–539. https://doi.org/10.1038/382536a0

Moss, E. G. (2007). Heterochronic Genes and the Nature of Developmental Time. *Current Biology*, *17*(11), 425–434. https://doi.org/10.1016/j.cub.2007.03.043

Moss, E. G., & Romer-Seibert, J. (2014). Cell-intrinsic timing in animal development. *Wiley Interdisciplinary Reviews: Developmental Biology*, *3*(5), 365–377. https://doi.org/10.1002/wdev.145

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., … Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, *424*(6946), 277–284. https://doi.org/10.1038/nature01789

Nam, Y., Chen, C., Gregory, R. I., Chou, J. J., & Sliz, P. (2011). Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. *Cell*, *147*(5), 1080–1091. https://doi.org/10.1016/j.cell.2011.10.020

Narasimhan, K., Lambert, S. A., Yang, A. W. H., Riddell, J., Mnaimneh, S., Zheng, H., … Hughes, T. R. (2015). Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. *ELife*, *2015*(4), 1–53. https://doi.org/10.7554/eLife.06967

Newman, A. P., Acton, G. Z., Hartwieg, E., Horvitz, H. R., & Sternberg, P. W. (1999). The lin-11 LIM domain transcription factor is necessary for morphogenesis of C. elegans uterine cells. *Development (Cambridge, England)*, *126*(23), 5319–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10556057

Newman, A. P., Inoue, T., Wang, M., & Sternberg, P. W. (2000). The Caenorhabditis elegans heterochronic gene lin-29 coordinates the vulval-uterine-epidermal connections. *Current Biology*, *10*(23), 1479–1488. https://doi.org/10.1016/S0960-9822(00)00827-7

Newman, M. A., Thomson, J. M., & Hammond, S. M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA (New York, N.Y.)*, *14*(8), 1539–49. https://doi.org/10.1261/rna.1155108

Nimmo, R. A., & Slack, F. J. (2009). An elegant miRror: microRNAs in stem cells, developmental timing and cancer. *Chromosoma*. https://doi.org/10.1007/s00412-009-0210-z

Page, A. P., & Johnstone, I. L. (2007). The Cuticle. In *WormBook The C. elegans Research Community*.

Page, A. P., Stepek, G., Winter, A. D., & Pertab, D. (2014). Enzymology of the nematode cuticle: A potential drug target? *International Journal for Parasitology: Drugs and Drug Resistance*, *4*(2), 133–141. https://doi.org/10.1016/J.IJPDDR.2014.05.003

Pagel, J.-I., & Deindl, E. (2011). Early growth response 1--a transcription factor in the crossfire of signal transduction cascades. *Indian Journal of Biochemistry & Biophysics*, *48*(4), 226–35. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/22053691

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., … Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, *408*, 86. Retrieved from http://dx.doi.org/10.1038/35040556

Perales, R., King, D. M., Aguirre-Chen, C., & Hammell, C. M. (2014). LIN-42, the Caenorhabditis elegans PERIOD homolog, Negatively Regulates MicroRNA Transcription. *PLoS Genetics*, *10*(7), e1004486. https://doi.org/10.1371/journal.pgen.1004486

Pourquie, O. (2003). The Segmentation Clock: Converting Embryonic Time into Spatial Pattern. *Science*, *301*(5631), 328–330. https://doi.org/10.1126/science.1085887

Reinhart, B. J., & Ruvkun, G. (2001). Isoform-specific mutations in the Caenorhabditis elegans heterochronic gene lin-14 affect stage-specific patterning. *Genetics*, *157*(1), 199–209. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11139502

Ririe, T. O., Fernandes, J. S., & Sternberg, P. W. (2008). The Caenorhabditis elegans vulva: a post-embryonic gene regulatory network controlling organogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20095–9. https://doi.org/10.1073/pnas.0806377105

Rougvie, A. E., & Ambros, V. (1995). The Heterochronic Gene Lin-29 Encodes a Zinc-Finger Protein that Controls a Terminal Differentiation Event in Caenorhabditis-Elegans. *Development*, *121*(8), 2491–2500. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7671813

Rougvie, A. E., & Moss, E. G. (2013). Developmental Transitions in C. elegans Larval Stages (pp. 153–180). https://doi.org/10.1016/B978-0-12-396968-2.00006-3

Roush, S., & Slack, F. J. (2008). The let-7 family of microRNAs. *Trends in Cell Biology*, *18*(10), 505–516. https://doi.org/10.1016/j.tcb.2008.07.007

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., & Wulczyn, F. G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nature Cell Biology*, *10*(8), 987–993. https://doi.org/10.1038/ncb1759

Sevetson, B. R., Svaren, J., & Milbrandt, J. (2000). A novel activation function for NAB proteins in EGR-dependent transcription of the luteinizing hormone beta gene. *The Journal of Biological Chemistry*, *275*(13), 9749–57. https://doi.org/10.1074/JBC.275.13.9749

Shyh-Chang, N., & Daley, G. Q. (2013). Lin28: Primal regulator of growth and metabolism in stem cells. *Cell Stem Cell*, *12*(4), 395–406. https://doi.org/10.1016/j.stem.2013.03.005

Spieth, J., & Blumenthal, T. (1985). The Caenorhabditis elegans vitellogenin gene family includes a gene encoding a distantly related protein. *Molecular and Cellular Biology*, *5*(10), 2495–501. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3841791

Su, J.-L., Chen, P.-S., Johansson, G., & Kuo, M.-L. (2012). Function and Regulation of Let-7 Family microRNAs. *MicroRNA E*, *1*(1), 34–39. https://doi.org/10.2174/2211536611201010034

Sulston, J. E., & Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology*, *56*(1), 110–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/838129

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

Swirnoff, A. H., Apel, E. D., Svaren, J., Sevetson, B. R., Zimonjic, D. B., Popescu, N. C., & Milbrandt, J. (1998). Nab1, a corepressor of NGFI-A (Egr-1), contains an active transcriptional repression domain. *Molecular and Cellular Biology*, *18*(1), 512–24. https://doi.org/10.1128/MCB.18.1.512

Tepper, R. G., Ashraf, J., Kaletsky, R., Kleemann, G., Murphy, C. T., & Bussemaker, H. J. (2013). PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity. *Cell*, *154*(3), 676–690. https://doi.org/10.1016/j.cell.2013.07.006

Van Nostrand, E. L., Sánchez-Blanco, A., Wu, B., Nguyen, A., & Kim, S. K. (2013). Roles of the developmental regulator unc-62/homothorax in limiting longevity in Caenorhabditis elegans. *PLoS Genetics*, *9*(2), e1003325. https://doi.org/10.1371/journal.pgen.1003325

Wählby, C., Kamentsky, L., Liu, Z. H., Riklin-Raviv, T., Conery, A. L., O'Rourke, E. J., … Carpenter, A. E. (2012). An image analysis toolbox for high-throughput C. elegans assays. *Nature Methods*. https://doi.org/10.1038/nmeth.1984

Wählby, C., Lee Conery, A., Bray, M.-A., Kamentsky, L., Larkins-Ford, J., Sokolnicki,

K. L., … O'Rourke, E. J. (2014). High- and low-throughput scoring of fat mass and body fat distribution in C. elegans. *Methods*, *68*(3), 492–499. https://doi.org/10.1016/j.ymeth.2014.04.017

Wilson, G. N. (1988, February). Heterochrony and human malformation. *American Journal of Medical Genetics*. Wiley-Blackwell. https://doi.org/10.1002/ajmg.1320290210

Wolkow, C. A., & Hall, D. H. (2015). Introduction to the Dauer Larva, Overview. In *WormAtlas.*

Worringer, K. A., Rand, T. A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., … Yamanaka, S. (2014). The let-7/LIN-41 Pathway Regulates Reprogramming to Human Induced Pluripotent Stem Cells by Controlling Expression of Prodifferentiation Genes. *Cell Stem Cell*, *14*(1), 40–52. https://doi.org/10.1016/J.STEM.2013.11.001

Wu, L., & Belasco, J. G. (2005). Micro-RNA Regulation of the Mammalian lin-28 Gene during Neuronal Differentiation of Embryonal Carcinoma Cells Micro-RNA Regulation of the Mammalian lin-28 Gene during Neuronal Differentiation of Embryonal Carcinoma Cells. *Molecular and Cellular Biology*, *25*(21), 9198–9208. https://doi.org/10.1128/MCB.25.21.9198

Xu, X., & Kim, S. K. (2011). The early bird catches the worm: new technologies for the Caenorhabditis elegans toolkit. *Nature Reviews Genetics*, *12*(11), 793–801. https://doi.org/10.1038/nrg3050

Yi, W., & Zarkower, D. (1999). Similarity of DNA binding and transcriptional regulation

by Caenorhabditis elegans MAB-3 and Drosophila melanogaster DSX suggests conservation of sex determining mechanisms. *Development (Cambridge, England)*, *126*(5), 873–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9927589

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., … Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526

Zhang, J., Zhang, Y., Sun, T., Guo, F., Huang, S., Chandalia, M., … Fu, M. (2013). Dietary obesity-induced Egr-1 in adipocytes facilitates energy storage via suppression of FOXC2. *Scientific Reports*, *3*, 1–10. https://doi.org/10.1038/srep01476

Chapter 2: Regulation of *C. elegans* L4 cuticle collagen genes by the heterochronic protein LIN-29

The work in this chapter was published in *genesis* in March 2018:

Regulation of *C. elegans* L4 cuticle collagen genes by the heterochronic protein LIN-29. (2018) P. Abete-Luzi and D. M. Eisenmann. *genesis* 56: e23106. https://doi.org/10.1002/dvg.23106

I. Introduction

The temporal regulation of gene expression is an essential aspect of metazoan development. After embryogenesis, the ecdysozoan nematode *C. elegans* goes through four larval stages (L1 - L4), molting its outer cuticle between each transition before becoming an adult. The major components of the cuticle are nematode-specific cuticle collagen proteins, which are expressed from a large gene family (Cox, 1992; A. Page & Johnstone, 2007). Previously, it was shown that several cuticle collagen (*col)* genes show a peak of expression in each larval stage (Johnstone, 2000; Johnstone & Barry, 1996), but examination of modENCODE project RNA-Seq temporal development data showed that many *col* genes (116/187) display a strong peak of expression in only one developmental stage (Gerstein et al., 2010; Jackson, Abete-Luzi, Krause, & Eisenmann, 2014). This set of temporally co-regulated related genes provides a powerful system to study temporal regulation of gene expression. In *C. elegans*, the heterochronic pathway controls the timing of several developmental events (Moss, 2007). LIN-29, the most downstream effector of this pathway, is a zinc finger transcription factor that accumulates in hypodermal cells in the L4 stage and is required for a subset of developmental events at the larva-to-adult transition, including the expression of adult specific collagens *col-7* and *col-19* (Liu, Kirch, & Ambros, 1995). Our work is focused on a subset of *col* genes that peak during the L4 larval stage for which we identify regulatory sequences, transcription factor requirements and *in vitro* DNA-protein interactions with the heterochronic transcription factor LIN-29.

II. Methods

II.a. *C. elegans* growth and strains used

C. elegans animals were cultured using standard methods (Brenner, 1974). Worms were grown on NGM plates and fed with *E. coli* OP50, or HT115 in the case of RNAi experiments. Experiments were performed at 20°C unless indicated otherwise. The following strains were used in this work:

NL2099: *rrf-3(pk1426) II*

EG669: *ttTi5605 II; unc-119(ed3) III*

EG8078: *oxTi185 I; unc-119(ed3) III*

RB2025: *eri-1(ok2683) IV*.

II.b. Molecular cloning and mutagenesis

All *col* reporters and subsequent *col* promoter deletions were cloned upstream of *2XNLS::yfp* in pDE350. pDE350 was created by removing the region containing [*multicloningsite::2XNLS::yfp::unc-54-3'UTR*] from pBJ101 (Jackson et al., 2014) and inserting it between the SbfI and SpeI sites of the MosSCI (*ttTi5605*) targeting vector pCFJ350 (Frøkjær-Jensen, Davis, Ailion, & Jorgensen, 2012; Addgene plasmid #34866). Different *col* promoter fragments were specified by PCR amplification using the primers

listed in (Table 1) and inserted into pDE350 digested with AvrII and SbfI, via Gibson Assembly® by New England Biolabs (NEB). All mutations of *col-38p* were generated with Q5® Site-Directed Mutagenesis Kit (NEB) except for *col-38p(-262G1G2G3m)*, *col-38p(-262G3m)* and *col-38p(-262L1L2L3L4m)* which were built using gBlocks® from Integrated DNA Technologies (IDT) containing the desired sequence. Mutations of predicted LIN-29 binding sites were all 5'-AAAAA-3' → 5'-AGGGA-3' and mutations of predicted GATA binding sites were 5'-TTATC-3' → 5'-GCAGC-3'. To express heat-shock inducible LIN-29 *(hs::lin-29)*, we first used PCR to remove a fragment from heat shock vector pPD48.79 (a gift from Andrew Fire, Addgene plasmid # 1447) containing [*hsp-16.2p:multicloning site:unc-54-3'UTR*] and cloned it into MosSCI (*ttTi5605*) targeting vector pCFJ350 to create plasmid pPA4. We then introduced a *lin-29a* cDNA (gBlocks®, IDT) into pPA4 via Gibson Assembly® (NEB) to generate pPA5 (*hsp-16.2p::lin-29::unc-54*-3'UTR). Worms were injected with either these plasmids to create the single-copy integrated strains *hs::control* (pPA4) and *hs::lin-29* (pPA5). Sequences of plasmids are available upon request.

Table 1. Primers used for the amplification of promoter regions selected for the deletion analyses of *col* genes. Capital letters indicate the 5' end section carrying plasmid homology for Gibson assembly into pDE350 digested with AvrII and SbfI.

Primer name	Description	Sequence
OPA01	*col-38p* -1070 Fwd	GAGGGTACCAGAGCTCACagacgctttctggcatagtagtc
OPA17	*col-38p* -502 Fwd	GAGGGTACCAGAGCTCACtttgcttcaagtccgcaatc
OPA20	*col-38p* -262 Fwd	GAGGGTACCAGAGCTCACcactcaagaagattatccggc

ODE580	*col-38p* -231 Fwd	GAGGGTACCAGAGCTCACggacggaaaagcaggttttcc
ODE575	*col-38p* -199 Fwd	GAGGGTACCAGAGCTCACggttttttgcaccaaaaacttag
OPA02	*col-38p* Rev	CTAGAGTCGACCTGCAGGttttagacgatgtctgagtgttcc
OPA03	*col-49p* -2009 Fwd	GAGGGTACCAGAGCTCACgcagataaaaattgaaaagcagac
OPA18	*col-49p* -1223 Fwd	GAGGGTACCAGAGCTCACtgagtttgaaaaactagtacaacagg
OPA21	*col-49p* -591 Fwd	GAGGGTACCAGAGCTCACctattttccaaacggattaaattg
OPA24	*col-49p* -282 Fwd	GAGGGTACCAGAGCTCACctagttttgagcagctacgaca
OPA32	*col-49p* -120 Fwd	GAGGGTACCAGAGCTCACgcttgttgtatccagttcatcaaa
OPA04	*col-49p* Rev	CTAGAGTCGACCTGCAGGgatgatgaggaagaatgagatttt
OPA05	*col-63p* -916 Fwd	GAGGGTACCAGAGCTCACcaatcttccttcggcacaat
OPA19	*col-63p* -414 Fwd	GAGGGTACCAGAGCTCACgccaccaatgagttttcaga
OPA22	*col-63p* -222 Fwd	GAGGGTACCAGAGCTCACtcttgaaattctaagccacataca
OPA25	*col-63p* -113 Fwd	GAGGGTACCAGAGCTCACgatcctttccaagcgaaaaac
OPA06	*col-63p* Rev	CTAGAGTCGACCTGCAGGcgttggatgatgtctattcgtaa

II.c. Generation of transgenic strains

All single copy insertion strains carrying YFP transcriptional reporters, as well as inducible LIN-29 (*hsp-16.2p::lin-29::unc-54*-3'UTR) and control (*hsp-16.2p::unc-54*-3'UTR) were generated by microinjection of MosSCI targeting vectors specific for the *ttTi5605* site, into insertion strain EG6699 (Mos site in LGII) following standard protocols for injection (Mello & Fire, 1995) and selection (Frøkjær-Jensen, 2015). The *hsp-16.2p::lin-29::unc-54*-3'UTR vector was also microinjected into EG8078 (Mos site on LGI) and integrated to facilitate crossing with the strain carrying *col-38p(-262)* reporter integrated on LGII.

II.d. Ectopic induction of LIN-29

Strains carrying either *hsp-16.2p::lin-29::unc-54*-3'UTR or *hsp-16.2p::unc-54*-3'UTR (control) were grown at 20°C and induced by heat shock exposure of 30 minutes at 37°C followed by 60-minute recovery at 20°C before sample collection or imaging. Inductions in the embryo were done in a mixed population of eggs. Inductions in the L1 stage, in the L2/L3 molt and in the adult, were done at 3, 26 and 66 hours-post-feeding, respectively.

II.e. Imaging and YFP expression recording

All transgenic animals carrying *yfp* reporters were imaged on a Zeiss Axioplan 2 and recorded with a Lumenera Infinity 3 camera and Infinity Analyze software. Every construct was assessed in at least two independent lines. Specific developmental stages in which animals were assessed were determined by the extent of gonad migration or vulval morphology. YFP expression results were evaluated in terms of penetrance and recorded as either positive or negative. In general, the intensity (expressivity) of the L4-stage YFP reporters used in this work showed minor variation in the mid L4 with a tendency to increase in brightness and reached a stable maximum in the late L4. The expression of *col-38p::yfp* across all stages was as follows: late embryogenesis 0%; L1 stage 0%; L2 stage 0%; L3 stage 0%; early L4 stage 0%; mid L4 stage (Christmas tree) 81%, late L4 stage 100%; newly-gravid adult 0% ($n \geq 15$ in all cases). The developmental expression of *col-49p::yfp* and *col-63p::yfp* reporters was substantially similar.

II.f. RNAi

Synchronized L1-staged worms were incubated at 20°C (or at 25°C when indicated) and RNAi treated by feeding as described (Kamath, Martinez-Campos, Zipperlen, Fraser, & Ahringer, 2000). The *lin-29* RNAi clone used in this work was obtained from the Ahringer RNAi library (Kamath & Ahringer, 2003). RNAi clones for *elt-1*, *elt-3* and *mab-10* were obtained from the Vidal library (Rual et al., 2004). The RNAi control was empty 'feeding' vector L4440, a gift from Andrew Fire (Addgene plasmid # 1654). Effectiveness of *lin-29* RNAi was monitored by *lin-29*(*lf*) related adult phenotypes of abnormal vulva and egg-laying defective (Kamath & Ahringer, 2003; Rual et al., 2004; Trent, Tsuing, & Horvitz, 1983) which were greater than 80% penetrant. Effectiveness of *elt-1* + *elt-3* RNAi was assessed by the somatic phenotype of herniation through the vulva at the L4 molt (Smith, McGarr, & Gilleard, 2005; penetrance 50%) and embryonic lethality of progeny (Baugh et al., 2005; B. D. Page, Zhang, Steward, Blumenthal, & Priess, 1997; penetrance >80%). Lastly, *mab-10* RNAi causes no visible phenotypes in hermaphrodites, however it was shown that *mab-10*(*lf*) leads to a 3-fold increase of *nhr-25* expression in the adult (Harris & Horvitz, 2011). We corroborated the effectiveness of *mab-10(RNAi)* experiments by including qPCR analysis of *nhr-25* (three replicates) and observed a 2-fold increase of the latter in the late L4, indicating that the treatment was effective to some degree. In all cases, vector control animals did not display these phenotypes. Maternal feeding of *col-38p(-262)::yfp; eri-1* (RNAi hypersensitive) animals was also performed: P0 hermaphrodites were grown from the L1 stage on *elt-1/3* RNAi plates and then surviving newly-hatched L1s were moved onto fresh *elt-1/3* RNAi plates and scored for YFP expression in the late L4.

II.g. RT-qPCR

Synchronized and RNAi-treated animals were collected at the mid L4 stage; synchronized and heat-shocked animals were collected either at the L2/L3 molt stage or at the first day of adulthood, one hour after the heat shock. Each condition was assessed by two-step RT-qPCR in three independent biological replicates. In all cases, samples consisted of pellets of 50-100 µl of worms which were washed multiple times and resuspended (~600 µl) in DEPC water. Worms were homogenized with gentleMAC dissociator and used for RNA preparations with commercial kit Quick-RNA™ MiniPrep (Zymo Research). Total RNA was reverse transcribed with a blend of oligo(dT) and random primers provided by iScript cDNA synthesis kit (BioRad). Real-time PCRs were performed with exon-exon spanning primers (Table 2) and the iQ™ SYBR® Green Supermix system (BioRad). All Ct values were normalized to housekeeping gene *gpd-2* and data was analyzed by the 2(delta-delta-Ct method) (Livak & Schmittgen, 2001).

Table 2. Primers used for quantitative PCR.

Primer name	Description	Sequence
OPA28	*col-38* qPCR Fwd	GGAGTCCATGACATGAAGGTG
OPA29	*col-38* qPCR Rev	CCTTGAGAGTTGGCATCACA
OPA30	*col-49* qPCR Fwd	TCATTTCGTTTGAGCATTCG
OPA31	*col-49* qPCR Rev	ATCCCTTCTCTCCTGGTGGT
OPA07	*col-63* qPCR Fwd	CTATTGTTCCAGCTATTTTTGCC
OPA08	*col-63* qPCR Rev	GCATCTCCATATCCTCTTCTGA

OPA11	*col-138* qPCR Fwd	AGCAAGGACCAAAGGGAGAAG
OPA12	*col-138* qPCR Rev	ATATCCTGGAGCAGTTCTTGGT
col-54_RTFPnew	*col-54* qPCR Fwd	GGAACACCAGGAGCCGAC
*col-54*_RTRPnew	*col-54* qPCR Rev	TCCTGGCTGTCCATCTGGT
lin-29FP	*lin-29* qPCR Fwd	CATTAGTTCTCTTACAGACCAGTTTG
lin-29RP	*lin-29* qPCR Rev	TTATCATGTTGAACCCAGGAG
nhr-25FP	*nhr-25* qPCR Fwd	AATGGTGATGAAGTTCCCGT
nhr-25RP	*nhr-25* qPCR Rev	AGTAGAGACTGCTGGGAAAGAGT
gpd-2_F	*gpd-2* qPCR Fwd	CCTCTGGAGCCGACTATGTC
gpd-2_R	*gpd-2* qPCR Rev	TGGCATGATCGTACTTCTCG

II.h. Electrophoretic mobility shift assays

LIN-29 DNA-binding domain protein was made with the *in vitro* Protein Synthesis kit (PURExpress) using plasmid pTH9033 as a template, which was a gift from K. Narasimhan and T. Hughes (Narasimhan et al., 2015). Double stranded DNA probes were made by PCR amplification using 5' biotinylated primers (Eurofins Operon). Probe and competitor oligonucleotide sequences are listed below. Electrophoretic mobility shift assays were done using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific). Per 20 µl of [0.2 mM EDTA/3 mM MgCl2/50uM Zinc Acetate/1mg/ml BSA] binding reaction we used 5 fmol of probe and 2 µl of PURExpress *in vitro* protein synthesis reaction. Binding reactions were incubated at room temperature for 20 minutes, then run on a 4% acrylamide gel for 70 minutes at 100 volts. Cold competitor oligonucleotide was included in the

binding reaction at 1000-fold molar excess. Samples were then transferred to a nylon membrane (100 volts for 40 minutes) and DNA was crosslinked by UV exposure. Detection was done by chemoluminescence and exposure on X-ray film following manufacturer's protocol.

EMSA probes and competitor DNA sequences are shown below—LIN-29 sites are underlined; and the number in name indicates first base relative to start codon:

-282 *col-49p*

ctagttttgagcagctacgacaaaaattcgttttggctttcagtacagattctaaactggtctattaaaaatgtgctacggtttcgtcagagatacagtgattat
ggctatcacgattcgggcgtgtcttgggaacttttttcaccgaacatttcctccgccgcttgttgtatccagttcatcaaagagaacaacgcaacgaacaac
gttttttctgcagttctaatagtgagtataaaaatgagatgaatatcttccgaaaatctcattcttcctcatcatc

-222 *col-63p*

tcttgaaattctaagccacatacacccatcggttattgctatgtgaaataacactcgataaagggaaggtgtatcgccgaaaaattgagaaaaaaagtagac
ggataggggatcctttccaagcgaaaaactggctattatgcaacacgcagatctacttttttgagctttttggatacgtatataatggcaacgatttttacga
atagacatcatccaacg

-262 *col-38p*

cactcaagaagattatccggcaatgaaaaatggacggaaaagcaggttttccttatcactaatggttttttgcaccaaaaacttagaaacaatttcccatac
ggtcctaacttttattcgattcttcgcgctgtcgaaaaaagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacg
tcttcactattaaaatgaatcgtttcctgatggaacactcagacatcgtctaaaa

-155 bp *col-38p* WT

taacttttattcgattcttcgcgctgtcgaaaaaagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtcttca
ctattaaaatgaatcgtttcctgatggaacactcagacatcgtctaaaa

-155 bp *col-38p* L4m

taactgggtattcgattcttcgcgctgtcgaaaaaagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtctt
cactattaaaatgaatcgtttcctgatggaacactcagacatcgtctaaaa

-155 bp *col-38p* L5m

taactttttattcgattcttcgcgctgtcgaggggagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtcttc
actattaaaatgaatcgtttcctgatggaacactcagacatcgtctaaaa

-155 bp *col-38p* L4L5m

taactgggtattcgattcttcgcgctgtcgaggggagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtctt
cactattaaaatgaatcgtttcctgatggaacactcagacatcgtctaaaa

-139 bp *col-38p* WT

tcttcgcgctgtcgaaaaaagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtcttcactattaaaatgaat
cgtttcctgatggaacactcagacatcgtctaaaa

-139 bp *col-38p* L5m

tcttcgcgctgtcgaggggagtggattcgcttccagtaccgtcacgcatctctctcaagttcacattatatctcaccacgaacgtcttcactattaaaatgaat
cgtttcctgatggaacactcagacatcgtctaaaa

L5 competitor

tcttcgcgctgtcgaaaaaagtggattcgcttcc

III. Results and Discussion

III.a. *in vivo* expression of *col* transcriptional reporters confirms developmental RNA-seq data

To validate the modENCODE data analysis showing stage-specific *col* expression, we built several transcriptional YFP reporters using entire *col* upstream promoter regions and integrated them in single copy into the genome. Reporters for *col* genes with peak expression in the embryo (*col-121*, *dpy-17*), L2 stage (*col-54*, *col-41*) and L4 stage (*col-49*, *col-63* and *col-38*) all began to show strong YFP expression at the expected time (Figure 1), indicating that members of this large gene family display specific temporal expression at defined points in the life cycle, and this temporal control is mediated by upstream genomic regions.

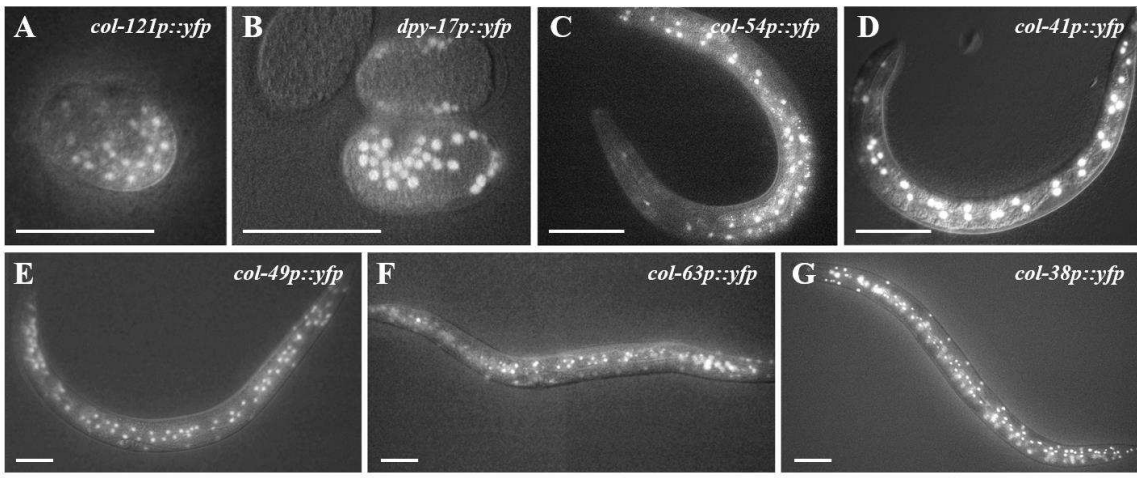


Figure 1. Stage-specific expression of cuticle collagen genes. Animals carrying the indicated YFP transcriptional reporters were imaged by epifluorescence and Nomarski microscopy at different stages: (A, B) embryo; (C, D) L2 stage; and (E-G) L4 stage. For each reporter, the figure shows the earliest timepoint when YFP was visible during development. YFP expression often perdured past this time. In the case of L2 and L4 stage *col* reporters, YFP was observed in hypodermal cells of the tail and head, hyp7, and seam cells. For details of *col-38p::yfp* developmental expression, see Methods. Scale bars indicate 50 um.

III.b. Upstream regions of less than 300 bp contain multiple LIN-29 sites and are sufficient to drive L4-specific *col* expression

To understand this temporal control of *col* gene expression, we investigated the regulation of three *col* genes with peak expression in the L4 stage. Promoter deletion analysis on these reporters narrowed the elements necessary for correct temporal expression to regions 262, 282 and 222 bp upstream of the start codons of *col-38*, *col-49*, and *col-63* respectively (Figure 2), remarkably small regions for *C. elegans* genes.

Previous work has identified four transcription factors regulating *col* gene expression. First, we showed that the L4 *col* genes *col-38*, *col-49*, and *col-71* are regulated by the Wnt pathway transcription factor BAR-1 (beta-catenin) (Gorrepati et al., 2015; Jackson et al., 2014; Van Der Bent et al., 2014), which binds to its target genes via interaction with the TCF transcription factor POP-1 (Jackson & Eisenmann, 2012). Second GATA factors ELT-1 and ELT-3 are required for proper expression of the *col* genes *dpy-7* (embryo peak),

col-41 (L2 peak), and *col-144* (no peak) (Budovskaya et al., 2008; Gilleard, Barry, & Johnstone, 1997; Gilleard & McGhee, 2001; Yin, Madaan, Park, Aftab, & Savage-Dunn, 2015). Finally, LIN-29, the terminal transcription factor in the heterochronic pathway, regulates expression of the adult *col* genes *col-7* and *col−19* and can bind large DNA fragments from the *col-19* promoter region (Liu et al., 1995; Rougvie & Ambros, 1995). Therefore, we looked for POP-1(TCF), GATA and LIN-29 binding sites in the minimal promoters of *col-38, col-49* and *col-63*.

Motif searching for a POP-1 binding motif ([T/C]TTTG[T/A][T/A]) (Jackson & Eisenmann, 2012) in the *col* gene minimal promoter regions showed a single site in *col−49p(-282)*(Figures 2 and 3), which could mediate the BAR-1 responsiveness of this gene, however there were no POP-1 sites in the minimal fragments *col−38p(-262)* and *col−63p(-222)*. This suggests that while two of these *col* genes may be responsive to BAR-1, their temporal pattern of expression is not likely to depend on BAR-1/POP-1 binding.

Unlike POP-1, there are putative LIN-29 and GATA sites in all three *col* minimal promoters. Narasimhan et al. showed that the LIN-29 DNA binding domain prefers sequences of 5As or 6As *in vitro* (Narasimhan et al., 2015). For each promoter, we identified five putative LIN-29 binding sites and called them L1 - L5 (Figures 2 and 3). In a like manner, we identified GATA factor binding sites (GATA[A/G]) and named them G1, G2, G3, etc (Figures 2 and 3). We evaluated the requirement of these putative LIN-29 and GATA binding sites for proper temporal expression of *col-38* by mutating or deleting them in our *col-38p(-262)::yfp* minimal promoter reporter and assessing YFP expression *in vivo*.

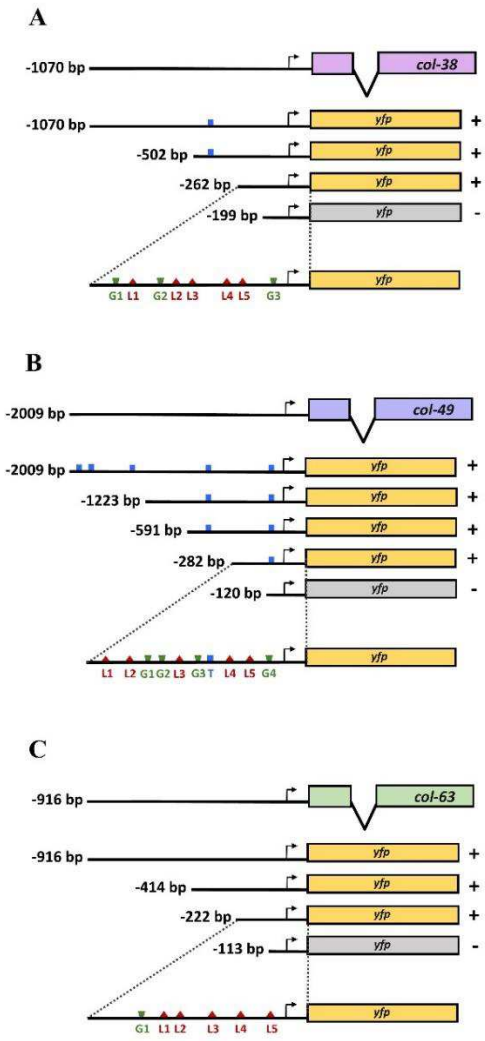


Figure 2. Identification of regulatory regions required for L4 expression in three L4 *col* YFP reporters. Promoter deletion analyses allowed the identification of minimal promoter regions of 262, 282 and 222 bp in (A) *col-38*, (B) *col-49* and (C) *col-63*, respectively. For each construct ≥22 animals were assessed in at least two independent lines. In all cases, YFP expression in the L4 stage was either present in ≥80% of the animals (+), or undetectable (-), in which case $P<0.001$ (Fisher's exact test) when compared to full length promoter. Locations of predicted binding motifs are shown for TCF/POP-1 (T, blue), GATA factors (G, green), and LIN-29 (L, red).

Figure 3. L4 *col* minimal promoters contain predicted regulatory sequences. LIN-29 (red), GATA (green) and POP-1 (blue) binding sites are found in promoters of *col-38*, *col-49* and *col-63* as indicated. Number in name refers to the first base relative to start codon

III.c. Dissection of *col-38* minimal promoter reporter reveals requirements of LIN-29 binding sequences

Removal of predicted LIN-29 binding site L1 or simultaneous mutation of sites L2 and L3 had no effect; however, mutation of L1, L2 and L3 together caused a slight reduction in

the number of animals showing YFP expression (L1L2L3m, Table 3). Interestingly, while neither the individual mutation of sites L4 or L5 showed an effect on *col-38p(-262)::yfp* expression, when L4 and L5 were simultaneously mutated YFP expression was completely abolished *in vivo* (L4L5m, Table 3). In addition, when either L4 or L5 was the only intact site, YFP expression was also absent (L1L2L3L4m; L1L2L3L5m, Table 3). Evolutionarily conserved, predicted LIN-29 binding sites in the position of L4 and L5 are found in the *col-38* gene promoters from four other *Caenorhabditis* species (Figure 4). These results indicate that: 1) putative LIN-29 binding sites are necessary for *col-38* expression *in vivo*, 2) these sites act redundantly; 3) sites L1, L2 and L3 enhance *col-38* expression but are not sufficient; 4) loss of both L4 and L5 prevents *col-38* expression; and 5) neither site L4 nor L5 alone is sufficient for activation of *col-38*.

We also tested the requirement of GATA sites for *col-38p(-262)::yfp* expression. While removal of G1 had no effect on YFP expression (ΔL1G1(-231), Table 3), deletion of the region containing G1 and G2 caused a loss of expression (-199; Figure 2C), suggesting these sites may be necessary. However, simultaneous mutation of G1 and G2 had no effect on YFP expression (G1G2m, Table 3), arguing that it is not the sites, but some other sequence in the region or the altered spacing in the -199 mutant that is important. Mutation of site G3 alone did completely abolish YFP reporter expression (G3m, Table 3) indicating this site is necessary for *col-38* expression.

Table 3. *in vivo* expression of *col-38* transcriptional reporters

***col-38p(-262)* reporter**	**% YFP expression (*n*)**
WT	100 (34)
Δ*L1G1 (-231)*	97 (26)
L2L3m	87 (39)
L1L2L3m	82* (46)
L4m	100 (38)
L5m	100 (48)
L4L5m	0** (29)
L1L2L3L4m	0** (39)
L1L2L3L5m	0** (25)
G1G2m	100 (28)
G1G2G3m	0** (10)
G3m	0** (28)

Transgenic animals carrying the indicated constructs were scored for YFP expression in the late L4 larval stage. Except for the -231 bp deletion (ΔL1G1), all reporter variants were mutagenized versions of the *col-38p* -262 bp reporter (*col-38p(262)::yfp*), and their names indicate which LIN-29 or GATA sites were mutated. Animals were scored as either positive or negative for YFP expression, since animals at the L4 stage all showed similar YFP intensity. *$P \leq 0.01$ and **$P < 0.0001$ (Fisher's exact test) compared to WT.

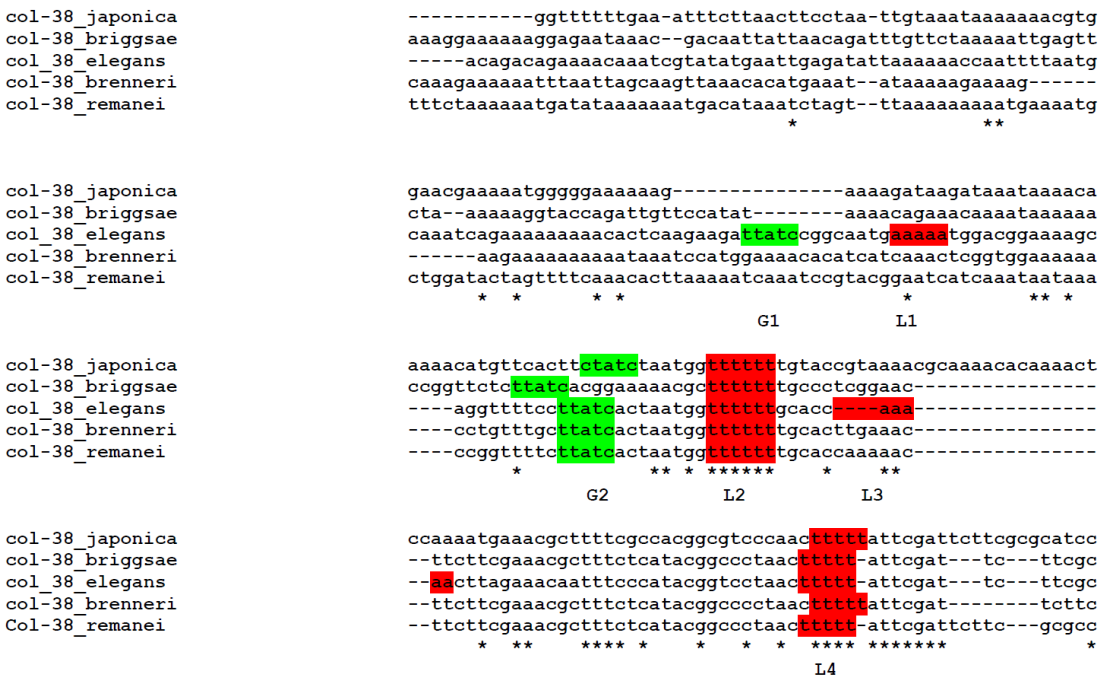


Figure 4. *col-38* regulatory sequences are conserved. Multiple sequence alignment of *col-38p* from five species of Caenorhabditis shows LIN-29 sites L2 and L4, as well as GATA site G2 as present in all five organisms (CLUSTAL O 1.2.4).

III.d. Transcription factor LIN-29 is required for L4 *col* expression

Since our analyses of *cis*-regulatory elements implicated GATA factors and the zinc finger transcription factor LIN-29 in L4-specific expression of *col-38,* we used RNAi to reduce function of these transcription factors. We examined *col-38* expression via *in vivo* observation of our *col-38p::yfp* reporter strain and also performed qPCR to assay endogenous expression of L4-specific *col* genes *col-38*, *col-49*, *col-63* and *col-138* in an RNAi-hypersensitive background (*rrf-3*(*pk1426*)).

ELT-1 and ELT-3 are hypodermal-specific GATA factors that are essential for hypodermal cell fate specification during embryogenesis (Chisholm & Hsiao, 2012) and that positively regulate expression of *col-41*, *col-144* and *dpy-7* (Budovskaya et al., 2008; Gilleard et al., 1997; Gilleard & McGhee, 2001; Yin et al., 2015). Neither *elt-1(RNAi)* nor *elt-3(RNAi)* caused a change in expression of *col-38p(-262)::YFP* or endogenous *col-38* (data not shown). However, since *elt-1* and *elt-3* may act redundantly (Gilleard & McGhee, 2001), we also tested combined *elt-1* + *elt*-3 RNAi. While we know this treatment was effective based on observation of expected phenotypes (see Methods), the *col-38p(-262)::yfp* reporter showed no observable change in expression when *elt-1/elt-3* combined RNAi was performed on L1 stage animals, and qPCR analysis of endogenous gene expression showed no effect except for a slightly higher expression of *col-63* in *elt-1/elt-3(RNAi)* animals (Table 4; Figure 5A). We repeated the double RNAi treatment on a strain containing the *col-38(-262G1G2m)::YFP* reporter in which only GATA site G3 is intact, and in an RNAi-hypersensitive background (*col-38(-262)::yfp; eri-1(ok2683)*), but still observed no change in YFP expression *in vivo* (Table 4). Finally, performing *elt-1/elt-3* combined RNAi on RNAi-hypersensitive mothers and observing *col-38p(-262)::yfp* reporter expression in the surviving progeny also caused no obvious change in penetrance or expressivity of YFP expression in the hypodermis (data not shown). These data suggest that either our RNAi treatment was strong enough to cause embryonic and larval somatic phenotypes but was not strong enough to compromise *col* gene expression, or that ELT-1 and ELT-3 do not play a major role in expression of this gene and some other protein may bind and function at or near the G3 site.

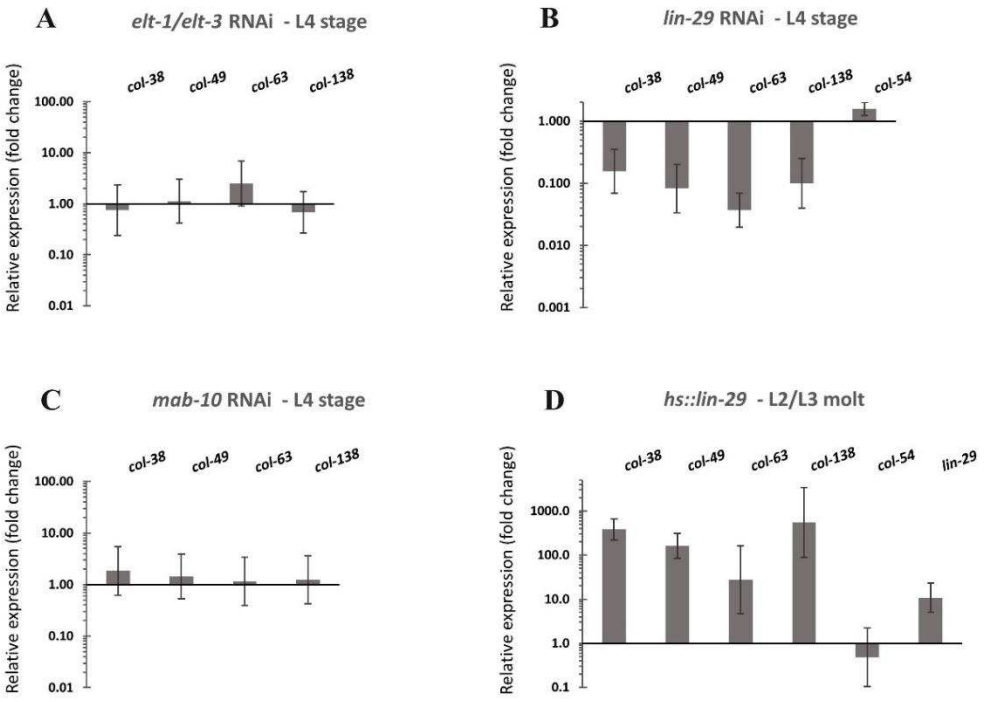
On the other hand, YFP expression from our single copy *col-38p(-262)::yfp* strain was completely abolished when treated with *lin-29* RNAi (Table 4). Similarly, transgenic strains carrying multicopy full length-promoter YFP reporters of the L4 *col* genes *bli-1* and *col-38* (Jackson et al., 2014) also showed strong YFP reduction in the L4 stage under the same conditions (Table 5). In addition to these *in vivo* YFP reporter observations, endogenous levels of the L4 *col* genes *col-38*, *col-49*, *col-63* and *col-138,* but not the L2 *col* gene *col-54*, were significantly reduced in the L4 when treated with *lin-29* RNAi (Figure 5B). Together these results strongly implicate LIN-29 in the regulation of five cuticle collagen genes showing a peak of expression in the L4 stage.

Harris & Horvitz showed that the transcription co-factor MAB-10 physically interacts with LIN-29 and together they promote seam cell differentiation and prevent extra molting events in males during the larva-to-adult switch (Harris & Horvitz, 2011). Therefore, we investigated whether MAB-10 is also required for LIN-29 regulation of L4-expressed *col* genes in hermaphrodites. We examined *col-38p(-262)::yfp* expression in *mab-10* RNAi treated animals and observed no effect (Table 4). Similarly, when we analyzed endogenous transcript levels of L4 *col* genes at the L4 stage under *mab-10* RNAi, we did not observe major changes in expression (Figure 5C). These results suggest that while MAB-10 may interact with LIN-29 to regulate several processes in males, it is not required for the regulation of *col* gene expression by LIN-29 in the L4 stage hermaphrodite.

Table 4. LIN-29 is required for L4 expression of *col-38*

		% YFP expression (*n*)		
Strain	**Treatment**	**Early L4**	**Mid L4**	**Late L4**
col-38p(-262)::yfp	control RNAi	0 (4)	84 (30)	100 (15)
col-38p(-262)::yfp	*lin-29* RNAi	0 (5)	0* (13)	0* (24)
col-38p(-262)::yfp	*mab-10* RNAi	0 (5)	92 (26)	100 (31)
col-38p(-262)::yfp	*elt-1*/*elt-3* RNAi	n.d.	78 (32)	100 (25)
col-38p(-262)::yfp; eri-1(ok2683)	control RNAi	0 (4)	77 (13)	100 (10)
col-38p(-262)::yfp; eri-1(ok2683)	*elt-1*/*elt-3* RNAi	0 (10)	81 (31)	100 (25)
col-38p(-262G1G2m)::yfp	control RNAi	0 (13)	83 (23)	97 (29)
col-38p(-262G1G2m)::yfp	*elt-1*/*elt-3* RNAi	0 (9)	81 (42)	96 (24)

Transcription factor requirements for the regulation of *col-38* expression were evaluated by *in vivo* imaging of YFP transcriptional reporters under different RNAi treatments. In all cases, RNAi was by the 'L1 feeding' method, and the efficiency of each RNAi treatment was corroborated by observation of known phenotypes or by the effect on the expression of known downstream target genes (see Methods). Hypodermal GATA requirements were tested by feeding animals with a mix of bacteria expressing *elt-1* RNAi combined with bacteria expressing *elt-3* RNAi constructs. This experiment was also conducted in an RNAi-sensitive background (*eri-1(ok2683)*; Kennedy et al., 2004), and in animals carrying the *col-38p* reporter with G1 and G2 sites mutated, so G3 is the only functional GATA site (*col-38p(-262G1G2m))*. Animals were scored as either positive or negative for YFP expression, since animals at the same L4 stage all showed similar YFP intensity. **P*<0.001 (Fisher's exact test) compared to the corresponding control.



E

	% YFP expression 60 minutes after heat shock (*n*)			
Strain	**Embryo (eggs)**	**L1 (3 hpf)**	**L2/L3 Molt (24 hpf)**	**Adult (72 hpf)**
col-38p(262)::yfp	**0 (20)**	**0 (20)**	**0 (46)**	**0 (36)**
col-38p(262)::yfp; hs::lin-29	**0 (20)**	**0 (20)**	**97* (115)**	**92* (52)**

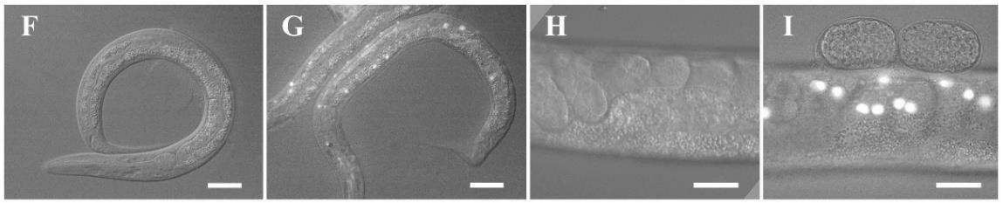


Figure 5. LIN-29 is necessary and sufficient for L4 *col* expression. Endogenous *col* gene expression in the L4 stage was assessed by RT-qPCR after different RNAi treatments: (A) combined GATA factors *elt-1/elt-3* RNAi, (B) *lin-29* RNAi and (C) *mab-10* RNAi. Quantification was relative to expression in animals treated with empty vector RNAi control. (D) L4 *col* gene expression was evaluated one hour after inducing LIN-29 at the L2/L3 molt in a *hs::lin-29* background. Quantification was relative to expression in *hs::control* animals. *col-54* peaks in the L2 stage when *lin-29* is not normally expressed and served as a control.

Error bars represent standard errors of the mean. (E) Expression of *col-38p(-262)::yfp* reporter was assessed after ectopic induction of LIN-29 in the embryo, in the L1, at the L2/L3 molt, and in the adult. (F-I) Epifluorescence and Nomarski microscopy of worms carrying either *col-38p(-262)::yfp* alone (F,H) or in a *hs::lin-29* background (G, I) examined at the L2/L3 molt (F, G) and in the adult (H, I). Scale bars are 25 um. **P*<0.001 (Fisher's exact test) when compared to strains carrying *col-38p(262)::yfp* alone.

Table 5. LIN-29 is required for the L4 expression of *bli-1* and *col-38*

			% YFP expression (*n*)	
Strain	**Treatment**	**Temperature**	**Mid L4**	**Late L4**
bli-1p::yfp	control RNAi	20°C	100 (10)	100 (20)
bli-1p::yfp	*lin-29* RNAi	20°C	80 (15)	54* (26)
col-38p::yfp	control RNAi	20°C	90 (29)	78 (55)
col-38p::yfp	*lin-29* RNAi	20°C	88 (25)	37* (24)
col-38p::yfp	control RNAi	25°C	100 (35)	100 (45)
col-38p::yfp	*lin-29* RNAi	25°C	53* (49)	4* (49)

LIN-29 requirements for the regulation of *bli-1* and *col-38* were assessed by the *in vivo* imaging of strains carrying multicopy full length-promoter YFP transcriptional reporters following *lin-29* RNAi treatment. The strains used in these experiments were from Jackson et al. (2014). The efficiency of each RNAi treatment was corroborated by observation of known *lin-29(lf)* phenotypes (see Methods). *lin-29* RNAi treatment at 25°C on *col-38* reporter showed a stronger effect than that at 20°C. Animals were scored as either positive

or negative for YFP expression, since L4 animals all showed the same YFP intensity. *P<0.001 (Fisher's exact test) compared to the corresponding control.

III.e. Misexpression of lin-29 is sufficient to activate expression of L4 cols in the L2/L3 molt and in the adult stage

Since LIN-29 alone was strongly required for the regulation of L4 *col* expression, we investigated whether ectopically induced LIN-29 is sufficient for the misexpression of the L4 *col* genes. We generated transgenic lines carrying a single copy of either *hs::lin-29* or *hs::control* (see Methods) and assessed endogenous L4 *col* gene expression after ectopic induction at the L2/L3 molt and in the adult. Unlike the L2 gene *col-54*, the L4 *cols* examined showed a strong increase when *lin-29* was induced at the L2/L3 molt (Figure 5D), and a modest increase when induction was done in the adult (data not shown). Consistent with these results, we heat-shocked *col-38p::yfp; hs::lin-29* animals and found that LIN-29 was sufficient to induce YFP expression at the L2/L3 molt and in the adult, but not in earlier stages (Figures 5E-I). This result shows that the provision of LIN-29 at a time it is not normally present is sufficient to induce expression of L4 *col* genes, suggesting it may be the major regulator of their expression at the normal L4 stage. Curiously, LIN-29 was not able to induce L4 *col* expression when provided in the embryo and L1 stage (Figure 5E), suggesting that at the earlier times either 1) some additional positive acting factor may be missing, 2) a repressor may be present, or 3) a non-permissive chromatin state exists at these genes.

III.f. LIN-29 specifically binds to LIN-29 binding sequences of L4 *col* minimal promoters *in vitro*

To determine whether LIN-29 may regulate L4 *col* gene directly, we tested LIN-29 binding to *col* promoter sequences *in vitro*. Notably, gel shift experiments showed that the LIN-29 DNA binding domain successfully binds the *col-38, col-49,* and *col-63* minimal promoter pieces *in vitro* (Figure 6A), producing multiple shifted species, consistent with multiple predicted LIN-29 binding sites in these sequences. In all cases, the LIN-29-promoter interactions were competed by a 34-bp oligo containing the single LIN-29 site L5 from *col-38p* and its flanking sequence (Figure 6A). A smaller 155 bp fragment of *col-38p* containing only L4 and L5, the two sites necessary for expression *in vivo*, also showed direct interaction with LIN-29 *in vitro.* This interaction was reduced when the L5 site was mutated, and almost abolished when both L4 and L5 sites were mutated (Figure 6B). Lastly, an even smaller 139 bp *col-38p* fragment containing only the L5 site also bound LIN-29 protein. However, when we mutated only the L5 site in the probe, the interaction was abolished (Figure 6B).

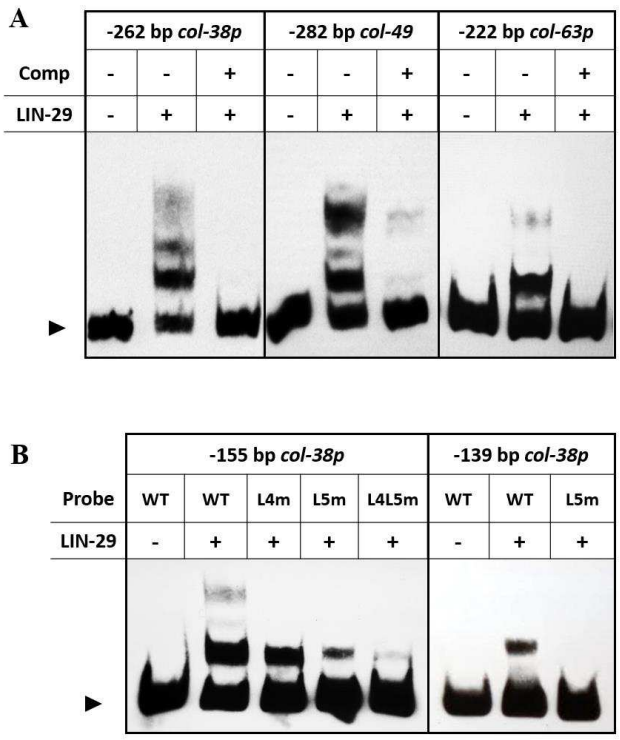


Figure 6. LIN-29 binds predicted DNA motifs in L4 *col* promoters *in vitro*. A) Electrophoretic mobility shift assays done with LIN-29 DNA binding domain-GST fusion protein and the minimal promoters of *col-38*, *col-49* and *col-63* as probes. In each case, the binding was competed away by a 34 bp oligo consisting of the single *col-38* LIN-29 site L5 and its flanking sequence (Comp). B) Electrophoretic mobility shift assays with LIN-29 DNA binding domain-GST fusion protein and either a -155 bp region of *col-38p* containing predicted LIN-29 sites L4 and L5 (left) or a -139 bp fragment of *col-38p* containing only site L5 (right). Binding of LIN-29 to these fragments was reduced or abolished when sites L4 and L5 were mutated individually (L4m, L5m) or together (L4L5m). Arrowhead indicates free probe.

III.g. Conclusion

In summary, in this work, we have 1) identified the heterochronic protein LIN-29 as a major regulator of the temporal expression of several cuticle collagen genes expressed during the larva-to-adult transition (*col-38, col-49, col-63,* and likely *col-138* and *bli-1*), 2) validated the predicted LIN-29 binding motif derived by Narasimhan et al. (2015), and 3) in the case of *col-38*, have shown direct binding of LIN-29 to two sites *in vitro*, and the requirement for those sites for *in vivo* expression in the L4 stage. Although the demonstration of direct regulation of these genes will require proof of LIN-29 *in vivo* binding site occupancy in L4 animals, we believe this is the first identification of specific binding sites for LIN-29 necessary for *in vivo* target gene expression. These results should aid with future efforts to understand temporal regulation of gene expression by this heterochronic protein at the larval-to-adult transition.

Acknowledgements

We thank K. Narasimhan and T. Hughes, C. Frøkjær-Jensen and E. Jorgensen, A. Fire, A. Golden, and K. O'Connell for providing plasmids, and A. Rougvie for providing plasmids and sharing unpublished results. We thank M. Negesse and D. Stonko for creation and sharing of reporter constructs. We thank members of the Eisenmann lab for advice and support. Some strains were provided by the Caenorhabditis Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Funding for this work was provided by the University of Maryland Baltimore County.

References

Baugh, L. R., Wen, J. C., Hill, A. A., Slonim, D. K., Brown, E. L., & Hunter, C. P. (2005). Synthetic lethal analysis of Caenorhabditis elegans posterior embryonic patterning genes identifies conserved genetic interactions. *Genome Biology*, *6*(5), R45. https://doi.org/10.1186/gb-2005-6-5-r45

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, *77*(1), 71–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4366476

Budovskaya, Y. V., Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., & Kim, S. K. (2008). An elt-3/elt-5/elt-6 GATA transcription circuit guides aging in C. elegans. *Cell*, *134*(2), 291–303. https://doi.org/10.1016/j.cell.2008.05.044

Chisholm, A. D., & Hsiao, T. I. (2012). The C. elegans epidermis as a model skin. I: development, patterning, and growth. *Wiley Interdisciplinary Reviews. Developmental Biology*, *1*(6), 861–878. https://doi.org/10.1002/wdev.79

Cox, G. N. (1992). Molecular and biochemical aspects of nematode collagens. *Journal of Parasitology*, *78*, 1–15. Retrieved from http://www.jstor.org/stable/3283678

Frøkjær-Jensen, C. (2015). Transposon-Assisted Genetic Engineering with Mos1-Mediated Single-Copy Insertion (MosSCI) (pp. 49–58). Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-4939-2842-2_5

Frøkjær-Jensen, C., Davis, M. W., Ailion, M., & Jorgensen, E. M. (2012). Improved Mos1-mediated transgenesis in C. elegans. *Nature Methods*, *9*(2), 117–118. https://doi.org/10.1038/nmeth.1865

Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., … Trifonov, T. (2010). Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. *Science*, *330*(6012), 1775 LP-1787. https://doi.org/10.2436/20.7010.01.53

Gilleard, J. S., Barry, J. D., & Johnstone, I. L. (1997). cis Regulatory Requirements for Hypodermal Cell-Specific Expression of the Caenorhabditis elegans Cuticle Collagen Gene dpy-7, *17*(4), 2301–2311. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC232079/pdf/172301.pdf

Gilleard, J. S., & McGhee, J. D. (2001). Activation of hypodermal differentiation in the Caenorhabditis elegans embryo by GATA transcription factors ELT-1 and ELT-3. *Molecular and Cellular Biology*, *21*(7), 2533–44. https://doi.org/10.1128/MCB.21.7.2533-2544.2001

Gorrepati, L., Krause, M. W., Chen, W., Brodigan, T. M., Correa-Mendez, M., & Eisenmann, D. M. (2015). Identification of Wnt Pathway Target Genes Regulating the Division and Differentiation of Larval Seam Cells and Vulval Precursor Cells in *Caenorhabditis elegans*. *G3: Genes|Genomes|Genetics*, *5*(8), 1551–1566. https://doi.org/10.1534/g3.115.017715

Harris, D. T., & Horvitz, H. R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. *Development*, *138*(18), 4051–4062. https://doi.org/10.1242/dev.065417

Jackson, B. M., Abete-Luzi, P., Krause, M. W., & Eisenmann, D. M. (2014). Use of an Activated Beta-Catenin to Identify Wnt Pathway Target Genes in Caenorhabditis

elegans , Including a Subset of Collagen Genes Expressed in Late Larval Development. *G3: Genes|Genomes|Genetics*, *4*(4), 733–747. https://doi.org/10.1534/g3.113.009522

Jackson, B. M., & Eisenmann, D. M. (2012). β-catenin-dependent Wnt signaling in C. elegans: teaching an old dog a new trick. *Cold Spring Harbor Perspectives in Biology*, *4*(8), a007948–a007948. https://doi.org/10.1101/cshperspect.a007948

Johnstone, I. L. (2000). Cuticle collagen genes. Expression in Caenorhabditis elegans. *Trends in Genetics : TIG*, *16*(1), 21–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10637627

Johnstone, I. L., & Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development. *The EMBO Journal*, *15*(14), 3633–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8670866

Kamath, R. S., & Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. *Methods (San Diego, Calif.)*, *30*(4), 313–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12828945

Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., & Ahringer, J. (2000). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. *Genome Biology*, *2*(1), research0002.1. https://doi.org/10.1186/gb-2000-2-1-research0002

Liu, Z., Kirch, S., & Ambros, V. (1995). The Caenorhabditis elegans heterochronic gene pathway controls stage-specific transcription of collagen genes. *Development (Cambridge, England)*, *121*(8), 2471–2478.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Mello, C., & Fire, A. (1995). DNA transformation. *Methods in Cell Biology*, *48*, 451–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8531738

Moss, E. G. (2007). Heterochronic Genes and the Nature of Developmental Time. *Current Biology*, *17*(11), 425–434. https://doi.org/10.1016/j.cub.2007.03.043

Narasimhan, K., Lambert, S. A., Yang, A. W. H., Riddell, J., Mnaimneh, S., Zheng, H., … Hughes, T. R. (2015). Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. *ELife*, *2015*(4), 1–53. https://doi.org/10.7554/eLife.06967

Page, A., & Johnstone, I. L. (2007). The cuticle. *WormBook*, 1–15. https://doi.org/10.1895/wormbook.1.138.1

Page, B. D., Zhang, W., Steward, K., Blumenthal, T., & Priess, J. R. (1997). ELT-1, a GATA-like transcription factor, is required for epidermal cell fates in Caenorhabditis elegans embryos. *Genes & Development*, *11*(13), 1651–1661. https://doi.org/10.1101/gad.11.13.1651

Rougvie, A. E., & Ambros, V. (1995). The Heterochronic Gene Lin-29 Encodes a Zinc-Finger Protein that Controls a Terminal Differentiation Event in Caenorhabditis-Elegans. *Development*, *121*(8), 2491–2500. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7671813

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

Smith, J. A., McGarr, P., & Gilleard, J. S. (2005). The Caenorhabditis elegans GATA factor elt-1 is essential for differentiation and maintenance of hypodermal seam cells and for normal locomotion. *Journal of Cell Science*, *118*(24), 5709–5719. https://doi.org/10.1242/jcs.02678

Trent, C., Tsuing, N., & Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode Caenorhabditis elegans. *Genetics*, *104*(4), 619–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11813735

Van Der Bent, M. L., Sterken, M. G., Volkers, R. J. M., Riksen, J. A. G., Schmid, T., Hajnal, A., … Snoek, L. B. (2014). Loss-of-function of b-catenin bar-1 slows development and activates the Wnt pathway in Caenorhabditis elegans. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep04926

Yin, J., Madaan, U., Park, A., Aftab, N., & Savage-Dunn, C. (2015). Multiple cis elements and GATA factors regulate a cuticle collagen gene in Caenorhabditis elegans. *Genesis*, *53*(3–4), 278–284. https://doi.org/10.1002/dvg.22847

Chapter 3: Identification of LIN-29 target genes suggest a novel role in the regulation of lipid metabolism

The work described in this chapter will be submitted for publication with the following authors: Patricia Abete-Luzi, Tetsunari Fukushige, Sijung Yun, Michael W. Krause and David M. Eisenmann.

While most of the work was done by me, Patricia Abete-Luzi, the RNA-Seq library preparation and RNA sequencing was done by Dr. Tetsunari Fukushige, and the differential expression analysis pipeline was done by Dr. Sijung Yun, as part of our collaboration with the laboratory of Dr. Michael Krause at the NIH/NIDDK.

I. Introduction

I.a. Temporal regulation of *C. elegans* development

Probably one of the most crucial aspects of successful animal development is that cell division, cell specification and differentiation, cell migration, cell fusion and even cell death, must all occur in the right place and at the right time. The coordination and execution of these events is controlled by molecular mechanisms such as the activation of signaling pathways and the regulation of gene expression by key transcription factors. While much is known about spatial organization of these processes during metazoan embryonic development, much less is known about the equally-important temporal control of such matters. Here we examine the temporal control of gene expression during the last phase of development in the nematode *C. elegans*.

The life cycle of this organism takes four days and consists of 6 stages. *C. elegans* development begins with embryogenesis inside an eggshell, after which the animal hatches as a small, worm-shaped L1 stage larva which begins to feed and grow. The ecdyosozoan worm goes through a total of four larval stages (L1 – L4), molting its outer, collagen-rich cuticle covering between stages, before becoming an adult that is capable of laying eggs (Altun & Hall, 2009). Genetic and molecular analyses uncovered the heterochronic pathway as the main regulator of developmental timing in *C. elegans* (for review, see Chapter 1 section IV). This pathway consists of a network of proteins and microRNAs that interact to control the expression and stability of key transcription factors that regulate

developmental events in a stage-specific manner (Moss & Romer-Seibert, 2014; Nimmo & Slack, 2009; Rougvie & Moss, 2013). Mutations in components of this pathway lead to either the precocious or retarded occurrence of stage-specific events. Several of the members of the heterochronic pathway are conserved in vertebrates and control developmental timing and stem cell fate in those organisms as well (Moss, 2007).

I.b. Known targets of the heterochronic protein LIN-29

The most downstream heterochronic pathway regulator is LIN-29, a Kruppel-family zinc finger transcription factor with limited homology to mammalian EGR (early growth response) proteins (Harris & Horvitz, 2011). *lin-29* function is required for a number of developmental events that take place in the L4 stage in coordination with the worm's transition from larval to adult life. Some of these LIN-29-regulated events include the formation of the adult cuticle, the terminal differentiation and fusion of the lateral hypodermal cells (also called seam cells), the cessation of the molting cycle, the migration of the developing gonad, and the formation of various reproductive structures in both hermaphrodites and males (Bettinger, Euling, & Rougvie, 1997; Bettinger, Lee, & Rougvie, 1996; Fielenbach et al., 2007; Gupta, Hanna-Rose, & Sternberg, 2012; Inoue, Wang, Ririe, Fernandes, & Sternberg, 2005; Newman, Inoue, Wang, & Sternberg, 2000; Ririe, Fernandes, & Sternberg, 2008; Rougvie & Ambros, 1995). According to whole-body RNA-Seq data, *lin-29* transcript levels peak in the L3 stage, whereas immuno-staining and reporter fusions show that a major accumulation of LIN-29 protein takes place in hypodermal cells starting in the L4 (Aeschimann et al., 2017; Bettinger et al., 1996;

Gerstein et al., 2010; Harris & Horvitz, 2011). In these cells, *lin-29* expression is negatively regulated before the L4 stage by upstream heterochronic proteins: HBL-1/Hunchback, which is likely to repress *lin-29* transcription in the L2, and LIN-41/Trim, which binds the 5' end of the *lin-29a* transcript and blocks its translation in the L3 stage (Aeschimann et al., 2017; Lin et al., 2003a; Slack et al., 2000). Negative regulation of HBL-1 and LIN-41 by members of the *let-7* miRNA family subsequently allows LIN-29 accumulation to occur (Abbott et al., 2005; Juan E Abrahante et al., 2003; Lin et al., 2003a; Slack et al., 2000). Nevertheless, it should be noted that LIN-29 is also present in other tissues at earlier times (i.e. the anchor cell, sex myoblasts and distal tip cells in the L3 stage, and in cells of the pharynx starting in the L1 and persisting through adulthood), and that in some cases this expression is independent of upstream heterochronic regulation (Bettinger et al., 1996; Harris & Horvitz, 2011).

A number of target genes regulated by LIN-29 that function in stage-specific developmental events have been identified. In hypodermal seam cells, LIN-29 positively regulates *cki-1*, a cyclin-dependent kinase inhibitor, and negatively regulates *cdk-1*, a cyclin dependent kinase, leading to cell division arrest (Hong, Roy, & Ambros, 1998; Rausch, Ecsedi, Bartake, Müllner, & Großhans, 2015). In these cells *lin-29* is also necessary for the expression of *aff-1*, which encodes a fusogen that mediates cell-cell fusion (Friedlander-shani & Podbilewicz, 2011), and for the expression of the adult-specific cuticle collagen (*col*) genes *col-7* and *col-19* (Liu, Kirch, & Ambros, 1995; Rougvie & Ambros, 1995). In recent work, we showed that LIN-29 also regulates the L4-specific *col* genes *col-38*, *col-49*, *col-63* and *col-138* (Abete-Luzi & Eisenmann, 2018).

The nuclear hormone receptors NHR-23 and NHR-25, both positive regulators of the molting cycle, appear to be repressed by LIN-29 (Harris & Horvitz, 2011). *lin-29* is also partially required for the expression of certain genes (*ceh-2, egl-17, lin-11,* and *zmp-1*) in specific subsets of vulval cells in the L4 stage and adult (Inoue et al., 2005; Ririe et al., 2008). In the anchor cell of the somatic gonad, LIN-29 activates expression of *lag-2*, a Notch ligand that promotes uterine cell differentiation and the formation of the uterine-seam cell connection, or utse (Newman et al., 2000). In the distal tip cells of the somatic gonad, LIN-29 negatively regulates expression of the transcription factor BLMP-1 to regulate expression of the guidance receptor *unc-5*, to control turning of the migrating gonad arms (Fielenbach et al., 2007; T. F. Huang et al., 2014). Finally, LIN-29 was recently shown to be necessary for expression of vitellogenin genes *vit-1*, *vit-2*, *vit-3* and *vit-6* in the intestine, promoting an adult-specific event required for fertility (Dowen, Breen, Tullius, Conery, & Ruvkun, 2016).

For most of the genes listed above, it is unknown whether LIN-29 regulation is direct or indirect. However, we recently showed that LIN-29 binds to the promoters of *col-38, col-49* and *col-63 in vitro*, and in the case of *col-38,* mutation of specific LIN-29 binding sites abolished expression of a *col-38* reporter transgene *in vivo* in the L4 hypodermis (Abete-Luzi & Eisenmann, 2018). This data, along with the previous demonstration that LIN-29 binds DNA from the *col-19* genomic region (Rougvie & Ambros, 1995), suggests that at least some cuticle collagen genes are likely to be direct targets of LIN-29 in the hypodermis at the larval-to-adult transition. On the other hand, intestinal *vit* gene expression requires *lin-29* activity in the hypodermis as well as the activation of an mTORC2 signaling

pathway in intestinal cells, suggesting that a LIN-29-dependent signal acts between these two tissues to coordinate fat accumulation in the germline (Dowen et al., 2016).

I.c. Our goal and our approach

The transition to adulthood is a fundamental life history event for all animals, and it involves at least three major changes: the conclusion of a period of rapid somatic growth and differentiation, the acquisition of reproductive capabilities (e.g. sexual organogenesis), and the associated metabolic adjustment underlying a switch in energy investment from somatic to germinal functions. LIN-29 has previously been shown to contribute to all three aspects of the worm's puberty to some degree. Yet, we wondered whether other relevant but less obvious functions and targets of LIN-29 are still to be discovered. To further explore the regulatory network coordinated by LIN-29, and to possibly uncover potential new roles for this transcription factor, we searched for its downstream target genes at the genomic level. To do this, we chose not to modify LIN-29 expression via loss-of-function mutations in upstream heterochronic pathway components, but rather to temporally misexpress LIN-29 directly and examine changes in development and in gene expression. Using the same strains generated for the work described in Chapter 1 (*hs::control* and *hs::lin-29*), we carried out early overexpression of LIN-29 and assessed worm transcript levels by RNA-Seq. This approach allowed us to identify 1,101 target genes which we further evaluated for functional enrichment and spatiotemporal expression patterns. In this chapter, I will present our findings and how they support the idea of a role for LIN-29 that

was previously ignored: the regulation of fatty acid metabolism during the larva-to-adult transition.

II. Methods

II.a. *C. elegans* growth and strains used

In all cases, *C. elegans* animals were cultured using standard methods (Brenner, 1974). Worms were grown on NGM plates and fed with *E. coli* OP50, or HT115 in the case of RNAi experiments. Experiments were performed at 20°C unless indicated otherwise. The following strains were used in this work:

- NL2099: *rrf-3*(*pk1426*) *II*
- EG669: *ttTi5605 II; unc-119(ed3) III*
- SV1009: *heIs63* [*wrt-2p::gfp::ph + wrt-2p::gfp::H2B + lin-48p::mCherry] V*

II.b. Molecular cloning and generation of transgenic strains

To express heat-shock inducible LIN-29 (*hs::lin-29)*, we first used PCR to remove a fragment from heat shock vector pPD48.79 (a gift from Andrew Fire, Addgene plasmid # 1447) containing [*hsp-16.2p:multicloning site:unc-54-3'UTR*] and cloned it into MosSCI (*ttTi5605*) targeting vector pCFJ350 (Frøkjær-Jensen, Davis, Ailion, & Jorgensen, 2012) to create plasmid pPA4 (Abete-Luzi & Eisenmann, 2018). We then introduced a *lin-29a* cDNA (gBlocks®, IDT) into pPA4 via Gibson Assembly® (NEB) to generate pPA5 (*hsp-16.2p::lin-29::unc-54*-3'UTR). Inducible LIN-29 (*hsp-16.2p::lin-29::unc-54*-3'UTR) and control (*hsp-16.2p::unc-54*-3'UTR) strains were generated by microinjection of pPA5 and pPA4 MosSCI targeting vectors specific for the *ttTi5605* site, into strain EG6699 (contains a single Mos insertion site on LGII) following standard protocols for injection (Mello & Fire, 1995) and selection (Frøkjær-Jensen, 2015). Both strains were previously used in Abete-luzi & Eisenmann (2018).

II.c. Ectopic induction of LIN-29 via heat shock

In all cases, populations were synchronized at the L1 stage via bleaching. Strains carrying either *hsp-16.2p::lin-29::unc-54*-3'UTR or *hsp-16.2p::unc-54*-3'UTR (control) were grown at 20°C (or at 25°C when indicated) and induced by heat shock exposure of 30 minutes at 37°C and returned to growing temperature until scoring, imaging, or collection for RNA preparations. Specific developmental stages were determined by developmental

time post feeding (hpf) and verified by the extent of gonad migration and/or vulval cell division/morphology.

- **Induction protocols for analyses of body morphology and vulva phenotypes:**

In all cases animals were grown at 20°C. Heat shocks corresponding to the L2 (late L2), L3 (late L3), and L4 (mid L4) stages were done at 23, 33 and 43 hours-post-feeding (hpf), respectively.

- **Induction protocols for analyses of precocious seam cell fusion:**

Strains also carried the *heIs63* array. The late L2 induction was done in animals grown at 20°C, heat-shocked 23 hpf and scored 28 hpf. Both single and double L3 inductions were done in animals grown at 25°C. Early L3 heat shocks were done at 22 hpf; mid L3 scoring or a second heat shock was done 25 hpf. Late L3 scoring was done 27 hpf.

- **Induction protocols for analyses of precocious alae and gonad migration defects:**

Worms were grown at 25°C and double heat-shocked in the early and mid L3, 22 and 25 hpf, respectively., then scored in the early to mid L4 stage (29 to 32 hpf).

- **Induction protocols for assessment of LIN-29 target gene expression:**

In all cases animals were grown at 20°C. Induction for RNA-Seq analysis was done in early L3 staged animals, 28 hpf. Adult induction for RT-qPCR assessment of intestinal targets was carried out in gravid adults 66 hpf.

II.d. Imaging

Animals were mounted in 2.5% agarose pads and suspended in anesthetic solution (5mM levamisole). Nomarski (DIC) and epifluorescence microscopy was done on a Zeiss Axioplan 2 and recorded with a Lumenera Infinity 3 camera and Infinity Analyze software.

II.e. RNA-Seq

hs::lin-29 and *hs::control* worms grown at 20°C were induced in the early L3 (28 hpf), and after one-hour recovery at 20°C, samples were collected and frozen at -80°C for a minimum of 15 minutes. Samples consisted of pellets of 50-100 µl of worms which were washed multiple times and resuspended (~600 µl) in DEPC water before being homogenized with a gentleMAC dissociator (miltenyibiotec). RNA preparations were performed with Quick-RNA™ MiniPrep kit (Zymo Research), and a total of 6 samples (triplicates for *hs::lin-29* and *hs::control*) were sequenced with single-end 50 base reads on Illumina HiSeq 2500. Bioinformatics quality controls were done using FastQC, version 0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). ce10 reference genome was aligned using STAR, version 2.5.1b. The number of reads mapped to genes were counted using htseq, version 0.6.1p1. Differentially expressed genes were determined using DESeq2, version 1.12.3 with the cutoff of 0.05 on False Discovery Rate (FDR). Transcriptomic data from this work has been deposited in the GEO archive under accession number GSE118433.

II.f. RNAi treatments

Synchronized L1-stage *rrf-3*(*pk1426*) animals were incubated at 20°C (or at 25°C when indicated) and RNAi treated by feeding as described (Kamath, Martinez-Campos, Zipperlen, Fraser, & Ahringer, 2000). The RNAi control was empty 'feeding' vector L4440, a gift from Andrew Fire (Addgene plasmid # 1654). RNAi clones used in this work for *lin-29*, *wrt-6* and *ins-37*, were obtained from the Ahringer RNAi library (Kamath & Ahringer, 2003).

II.g. RNA isolation and RT-qPCR

Samples of late L4 (45 hpf) RNAi-treated and adult heat-shocked animals were collected along with their respective controls and stored at -80°C for a minimum of 15 minutes. For each experiment, relative transcript levels were assessed by two-step RT-qPCR in three-to-four independent biological replicates. In all cases, samples consisted of pellets of 50-100 µl of worms which were washed multiple times and resuspended (~600 µl) in DEPC water. Worms were homogenized with gentleMAC dissociator (miltenyibiotec) and used for RNA preparations with Quick-RNA™ MiniPrep kit (Zymo Research). Total RNA was reverse transcribed with a blend of oligo(dT) and random primers provided by iScript cDNA synthesis kit (BioRad). Real-time PCRs were performed with exon-exon spanning primers (Table 1) and the iTaq™ Universal SYBR® Green Supermix system (BioRad). All Ct values were normalized to housekeeping gene *gpd-2* and data was analyzed by the 2(delta-delta-Ct) method (Livak & Schmittgen, 2001).

Table 1. Primers used for quantitative PCR.

Primer description	Sequence
***hacd-1* qPCR Fwd**	**GATTCATTGTCAACCGGCTATT**
***hacd-1* qPCR Rev**	**TCGCGGAAGATTTTCAAAGTG**
***acs-7* qPCR Fwd**	**GAAGCATATTAATCAGAAGCTAGCCA**
***acs-7* qPCR Rev**	**CATCCAAAAACGCTTTCTTCA**
***dhs-18* qPCR Fwd**	**ATTGAGAAGGCAGGAGGTCAT**
***dhs-18* qPCR Rev**	**GGTAAGAGAGATGGCTGAAGCA**
***nhr-80* qPCR Fwd**	**CAGAATAAGCATTGTCCCTACAAA**
***nhr-80* qPCR Rev**	**ATTAATACAGCTTTCCGGAGCA**
***vit-1* qPCR Fwd** [§]	**GAGGTTCGCTTTGACGGATA**
***vit-1* qPCR Rev** [§]	**GGCTTCACATTCCTCGTTCT**
***vit-3* qPCR Fwd ***	**CATGTGCACCATCGAAGAACTC**
***vit-3* qPCR Rev ***	**CCAATGTGGTTTCAATGACAAGTTG**
***vit-6* qPCR Fwd ***	**TTCACCCAGAAGCCAGTTC**
***vit-6* qPCR Rev ***	**AGGATGGGAGGCAGTAGAC**
***col-38* qPCR Fwd**	**GGAGTCCATGACATGAAGGTG**
***col-38* qPCR Rev**	**CCTTGAGAGTTGGCATCACA**
***gpd-2* qPCR Fwd**	**CCTCTGGAGCCGACTATGTC**
***gpd-2* qPCR Rev**	**TGGCATGATCGTACTTCTCG**

*From Dowen et al. (2016); [§]From Ding & Grosshans (2009).

II.h. Motif search

MEME-ChIP, version 4.11.2 (http://meme-suite.org/tools/meme-chip; Bailey et al., 2009) was used with default parameters to find motifs in promoter regions of HS up-regulated genes using 2 kb upstream of gene start sites.

II.i. Protein category (GO term) and tissue enrichment analyses

Analysis of target gene lists for protein function was performed using Uniprot Knowledgebase (www.uniprot.orgl); GO term enrichment was performed using DAVID (https://david.ncifcrf.gov/; D. W. Huang, Sherman, & Lempicki, 2009) and AmiGO 2/PANTHER (http://amigo.geneontology.org; Carbon et al., 2009; Munoz-Torres & Carbon, 2017); and tissue enrichment analysis was performed using the Wormbase Enrichment Analysis tool (www.wormbase.org; Angeles-Albores, N. Lee, Chan, & Sternberg, 2016). Enrichment analyses were done using default parameters. Published data on target genes (IDs; RNAi phenotypes, sites of expression, etc.) from Supplemental Table 1 was retrieved using the Wormbase Simplemine tool (www.wormbase.org; Lee et al., 2018).

II.j. Survival analysis

hs::lin-29 and *hs::control* worms were synchronized at the L1 stage and grown under standard conditions (20°C) until the first day of adulthood when they were transferred to FUDR solid media (to induce complete sterility) at 25°C. Animals were fed with either live bacteria (*E. coli* OP50) or dead bacteria (UV-killed *E. coli* OP50 on 50µg/ml carbenicillin NGM plates); and heat shock inductions of LIN-29 were done either daily or every other day. All cohorts were followed until 100% mortality and survival curves were estimated with OASIS 2 (https://sbi.postech.ac.kr/oasis2/ Oncotarget 11269; Han et al., 2016) using the Kaplan-Meier method and statistically analyzed with the log rank test.

II.k. Protein-protein association networks

Both *C. elegans* and *H. sapiens* gene sets (selected LIN-29 metabolic targets and their human homologs; see Figure 11) were analyzed with STRINGv10 (Szklarczyk et al., 2015) in which interactions were selected from 'experiments', 'data bases', and 'co-expression' sources using the highest confidence score and the first shell of ten interactions. Graphical visualizations were made with Cytoscape software (Shannon et al., 2003).

III. Results

III.a. *hs::lin-29* induction before the L4 stage leads to body morphology and vulval development phenotypes

Loss-of-function mutations in upstream heterochronic pathway regulators precociously express LIN-29 earlier in development (Aeschimann et al., 2017; Slack et al., 2000), however the consequences of direct misexpression of LIN-29 have not previously been assayed**.** To that end we created a strain containing a single-copy, integrated transgene containing a full length *lin-29a* cDNA downstream of a heat shock promoter (referred to as *hs::lin-29*) and a control strain with the identical heat shock promoter and no insert (*hs::control*). Before using this strain for target gene identification, we tested whether temporal expression of LIN-29 during development was sufficient to cause phenotypes in processes associated with *lin-29*.

We tested induction of *hs::lin-29* at different larval times and looked for morphological defects in young adults. We heat-shocked animals either once in the L2, L3, or L4 stage, or twice; in the L2 and L3 stages, or in the L2 and L4 stages. Adults that were subject to early temporal overexpression of LIN-29 displayed three morphological phenotypes; whereas a few animals developed dumpy bodies (Dpy), many were egg-laying defective (Egl) or showed a substantial decrease in body size (Small), or both (Table 2 and Figure 1). The Small and Egl phenotypes, alone or combined, were predominant in most induction protocols. No morphological phenotypes were observed when LIN-29 expression was

induced only in the L4 stage when LIN-29 is normally present, however the penetrances for the Egl and Small phenotypes in animals subjected to a heat shock in the L2 were increased with an additional L4 induction, suggesting that excess LIN-29 in the L4 stage can contribute to these phenotypes. The morphological phenotypes we observed in *lin-29* gain-of-function conditions are similar to those seen in *lin-41* loss-of-function mutants, in which there is early accumulation of LIN-29 in the L3 stage: these animals also show Dpy, Small and slightly Egl phenotypes (Slack et al., 2000; Tocchini et al., 2014).

Table 2. Adult phenotypes following LIN-29 induction at different times of larval life

Time of heat shock	Strain	% WT	% Dpy	% Egl	% Small	% Small-Egl	*N*
No heat shock	*hs::lin-29*	98	0	0	2	0	130
L2	*hs::control*	99	0	0	1	0	73
	hs::lin-29	55	7	1	16*	21*	73
L3	*hs::control*	94	0	2	2	0	48
	hs::lin-29	22	8	38*	8	23*	86
L4	*hs::control*	100	0	0	0	0	72
	hs::lin-29	100	0	0	0	0	56
L2 + L3	*hs::control*	98	0	0	2	0	111
	hs::lin-29	12	5	55*	8	20*	145
L2 + L4	*hs::control*	99	0	0	1	0	224
	hs::lin-29	17	11*	7	28*	43*	161

Strains carrying *hs::lin-29* or *hs::control* were submitted to different protocols of heat shock (column 1) to test the effect of LIN-29 induction at different timepoints during development. Animals were assessed in the adult stage (72-76 hpf) for body morphology phenotypes by direct observation. * $P<0.001$ (Fisher's exact test) compared to the corresponding control.

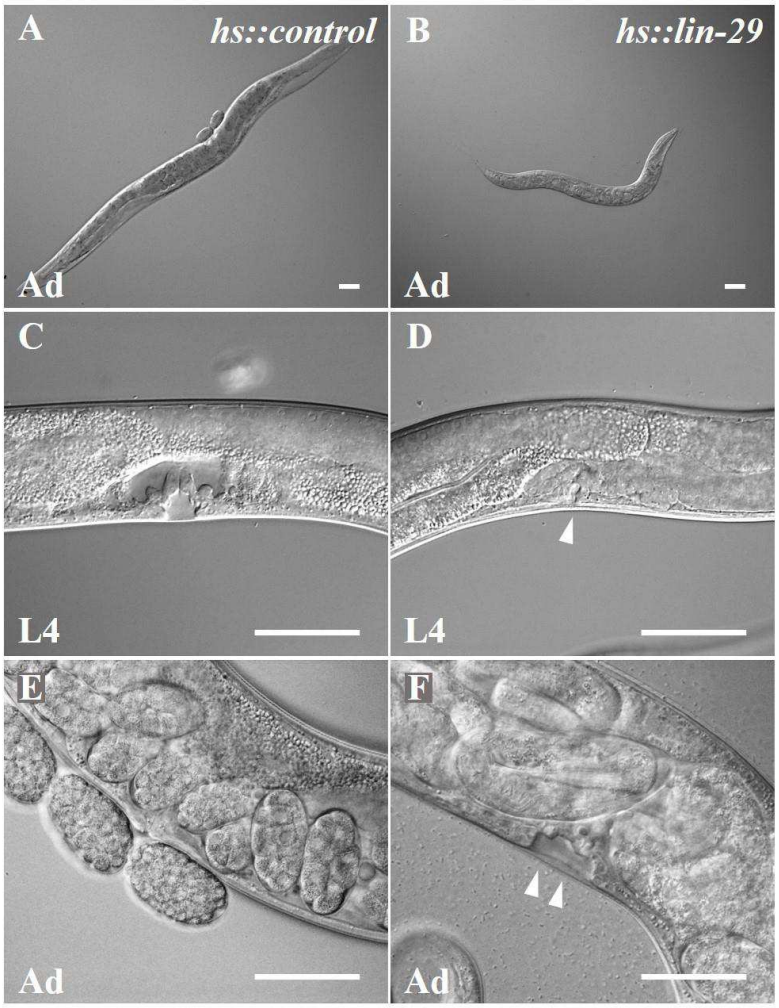


Figure 1. Early expression of LIN-29 results in body morphology and vulval defects. Nomarski images of synchronized *hs::control* (A, C, E) or *hs::lin-29* (B, D, F) animals that were subjected to heat shock early in development (see Table Phenotypes and Table Vulva; see Methods). (A, B) Adult animals that were given a single heat shock in both the L2 and L4 stages; note the Small/Egl phenotype in *hs::lin-29*. (C - E) Animals were given a single heat shock in both the L2 and L3 stages, and scored in the L4 (C, D) or the adult (E, F) stage. Single arrowhead in (D) indicates an underinduced vulva in the L4 stage. Double arrowhead in (F) indicates an L4 stage vulva (compare to *hs::control* L4 (C)) in an adult *hs::lin-29* animal (note the presence of unlaid, late stage embryos in the uterus). Scale bar: 50μm.

Table 3. Vulva defects induced by early expression of *lin-29*

		Vulva phenotypes				
Stage	**Strain**	**% WT**	**% abnormal**	**% under induced**	**% L4-like lumen**	**N**
L4	*hs::control*	100	0	0	n.d.	73
L4	*hs::lin-29*	66	18	15	n.d.	110
Ad	*hs::control*	100	0	n.d.	0	30
Ad	*hs::lin-29*	21	46	n.d.	33	67

Vulva developmental defects were assessed in the indicated strains after early induction of LIN-29 by heat-shocking once in the L2 stage and then again in the L3 stage. Phenotypes were scored by *in vivo* imaging (DIC) first in the L4 stage, then in adults of the same cohorts. In all cases $p < 0.001$ (Fisher's exact test) compared to the corresponding control.

To determine the basis for the Egl phenotype observed upon misexpression of LIN-29, we investigated the L2+L3 heat shocked-animals for vulval phenotypes at the L4 and young adult stages; we saw vulval abnormalities with significant penetrance at both times (see Table 3 and Figure 1). In the L4 stage, heat shock treated animals showed too few cells adopting vulval fates ('underinduced') or vulval inductions that were abnormal in morphology ('abnormal'). *lin-29* is known to be required for development of the egg-laying apparatus: *lin-29* mutants were first identified based on their loss-of-function Egl and protruding vulva (Pvl) phenotypes, and *lin-29* was later shown to be required for the formation of the connection between the uterus and vulva and for expression of certain genes in vulval cells (Ambros & Horvitz, 1984; Bettinger et al., 1997, 1996; Inoue et al., 2005; Newman et al., 2000). However, one vulval phenotype we observed deserves

comment: when examined as gravid adults, one third of heat-shocked animals showed a vulval morphology that resembled that found in mid- to late-L4 stage animals (Table 3; Figure 1F). Thus, it appears that in some cases vulva induction and vulval development proceeded normally until the mid L4, when further vulval development arrested even though these animals went on to be reproductively capable (Figure 1F). Previously it was shown that when L3 stage animals are removed from food, vulval development arrests at the early L4 stage of morphology, but animals with a mid/late L4 vulval morphology were not seen and these animals were not reported to advance to adulthood (Schindler, Baugh, & Sherwood, 2014). To our knowledge, this type of 'arrested L4 developing vulva in an adult' phenotype has not been observed before in other heterochronic or vulval mutants.

We observed that overexpression of LIN-29 in the L3 stage was also sufficient to cause a mild gonad migration phenotype. Although reduction of *lin-29* function does not cause a gonad migration defect on its own, genetic and molecular analyses indicate that *lin-29* acts on the migrating distal tip cell to control the timing of its turning event (Fielenbach et al., 2007; T. F. Huang et al., 2014). We found that 13% of *hs::lin-29* animals given two heat shocks in the L3 stage had defects in gonad migration (n = 48; compared to 0% for *hs::control* animals n=32). In these animals the elongating gonad arms turned dorsally correctly, but then one arm migrated in the incorrect direction along the anterior-posterior axis (data not shown).

III.b. Early *hs::lin-29* induction is sufficient to promote precocious seam cell fusion, but not precocious alae formation

LIN-29 is also known to regulate the fusion of the hypodermal seam cells with each other at the end of L4 stage via expression of the fusogen *aff-1* (Friedlander-shani & Podbilewicz, 2011). Therefore, we looked at seam cell behavior when precociously inducing LIN-29 in the late L2 and in the L3. We used the *heIs63* transgene, which expresses nuclear- and membrane-localized GFP from a seam cell-specific promoter (Wildwater, Sander, de Vreede, & van den Heuvel, 2011), to examine seam cell morphology in *hs::lin-29* and *hs::control* animals. We found that a single heat shock is enough to induce precocious seam cell fusion at high penetrance (Table 4 and Figure 2). We noted that the timing of the heat shock relative to the timing of the seam cell division affected the penetrance of the phenotype: cells that were newly divided and had not yet restored cell-cell contact did not display a precocious fusion phenotype, while single seam cells in contact with neighboring cells usually did show precocious fusion when LIN-29 was induced early. Our results indicate that LIN-29 is not only required but also sufficient for seam cell fusion, at least in the L3 stage.

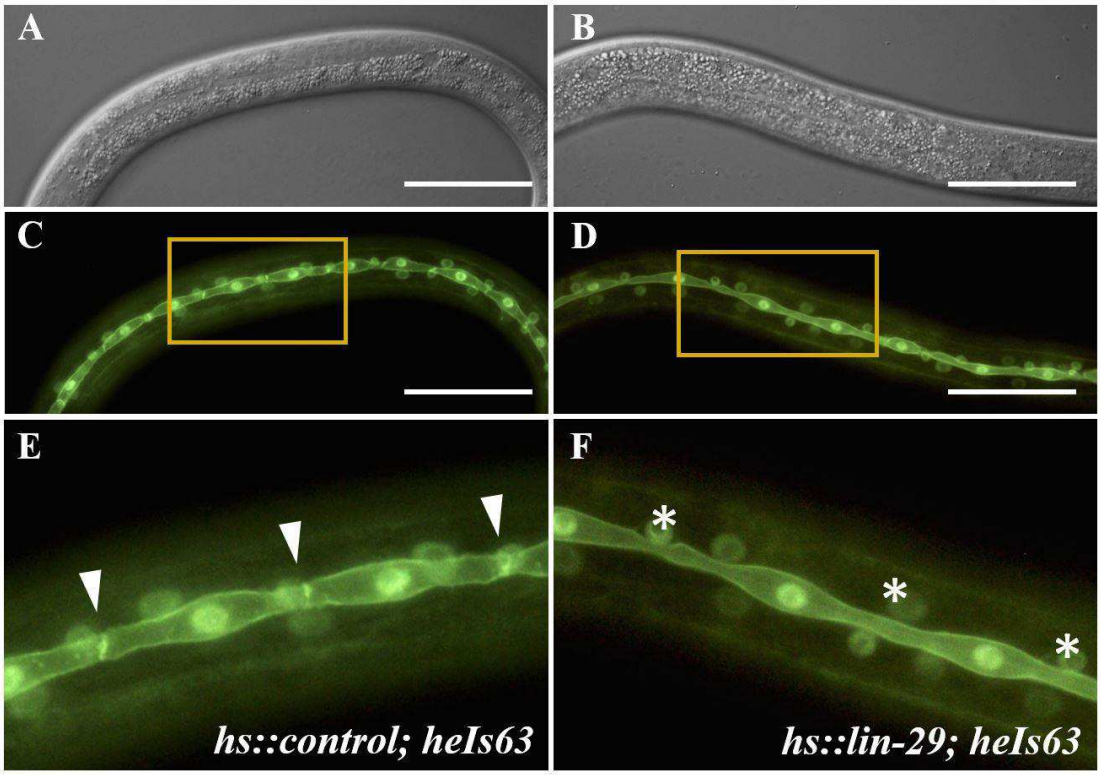


Figure 2. Early expression of LIN-29 is sufficient to cause precocious seam cell fusion. Shown here are synchronized L3-stage animals expressing nuclear and plasma membrane localized GFP in the hypodermal seam cells (from array *heIs63*) and carrying either *hs::control* (A, C, E) or *hs::lin-29* (B, D, F). Populations of animals were given a heat shock in the late L2 (see Table seam cell; Methods). Nomarski (A, B) and epifluorescence (C, D) microscopy of larvae 5 hours after induction. Precocious seam cell fusion is observed, as seen in the magnified view (E, F; from insets shown in C, D): cell junctions between seam cells are still present in the *hs::control* strain (E; arrow heads) but are absent in animals carrying *hs::lin-29* (F; asterisks). Scale bar: 50µm

Table 4. LIN-29-induced seam cell fusion in earlier developmental stages

Time of heat shock	Time observed	Strain	% Multiple or all seam cells fused	% Only one fusion between 2 cells	N
late L2	early L3	*hs::control; heIs63*	0	7	42
		hs::lin-29; heIs63	45	0	42
early L3	mid L3	*hs::control; heIs63*	0	0	30
		hs::lin-29; heIs63	82	3	60
early L3 + mid L3	late L3	*hs::control; heIs63*	0	0	18
		hs::lin-29; heIs63	100	0	30

Animals expressing GFP in the nucleus and at the plasma membrane of the seam cells (from *heIs63*; see Methods) and carrying either *hs::lin-29* or *hs::control* were heat-shocked and observed for precocious seam cell fusion at the indicated time. Multiple and single fusion events were scored with epifluorescence microscopy. In all cases $p \leq 0.0001$ (Fisher's exact test) compared to the corresponding control.

Unlike seam cell fusion and gonad migration, we found that early overexpression of LIN-29 was not sufficient to induce adult alae formation. It has long been known that *lin-29* mutants lack adult alae, indicating *lin-29* is necessary for secretion of these cuticular structures in the L4 stage (Ambros & Horvitz, 1984). We gave *hs::lin-29* animals two heat shock treatments in the L3 stage and observed them from 3 – 5 hours after the second heat shock period. Although we observed short, disorganized striations in rare animals, in no case did we observed the presence of unambiguous adult alae, even in small amount. Curiously, precocious adult alae have been observed at low levels at the L3 molt in *lin-41(lf)* mutants and in *lin-41* or *hbl-1* RNAi-treated animals (Fielenbach et al., 2007; Lin et

al., 2003a; Slack et al., 2000) in which LIN-29 was likely to have accumulated early. These results suggest that in addition to expression of LIN-29, precocious secretion of adult alae may also require a context wherein heterochronic regulators HBL-1 and/or LIN-41 are repressed or other target genes of these factors are expressed.

III.c. Identification of genes regulated after LIN-29 temporal misexpression

To identify target genes regulated by LIN-29 we used a *gain-of-function* approach in which we examine global changes in gene expression following overexpression of wild type LIN-29 using the heat shock promoter, as we did recently for the transcription factor BAR-1 (Gorrepati et al., 2015; Jackson, Abete-Luzi, Krause, & Eisenmann, 2014). To our knowledge, most work previously done on LIN-29 has been done using *lin-29* loss-of-function mutation or RNAi-treated strains. For example, while investigating targets regulated by the miRNA *let-7*, Hunter et al. conducted microarray analyses of L4-staged *lin-29*(*n333*) mutant animals versus wild type (Hunter et al., 2013). However, one of the caveats of looking for target genes with a loss-of-function approach is that observed changes in gene expression or phenotype may be an indirect, downstream consequence of changes in cell fate or other defects caused by the loss of a regulatory factor during development. By expressing LIN-29 at a discrete time in otherwise normally-developed animals and then examining changes in gene expression a short time later, we are more likely to avoid such secondary effects and identify direct targets of LIN-29. Since overexpression of LIN-29 earlier in development is sufficient to cause phenotypes, we used

our *hs::lin-29* and *hs::control* strains to examine global changes in gene expression following LIN-29 temporal misexpression. Although there are caveats to this approach as well (see Discussion), we believe genes showing altered regulation upon overexpression of LIN-29 at an earlier time in development are likely to identify targets of LIN-29 during its normal role in the L4 stage.

Strains carrying either *hs::lin-29* or *hs::control* were given a single heat shock in the early L3 stage and RNA-Seq transcriptomic analysis was performed on triplicate samples collected one hour after the end of the heat shock period. We chose this time since it is close to the normal peak of LIN-29 gene expression, so other aspects necessary for LIN-29 function such as the presence of other transcription factors or a permissible chromatin state, may be present. We found 1,101 genes that were differently expressed ($p<0.05$) between the two heat shocked strains. As a first way to evaluate the validity of these results, we used MEME analysis (MEME Suite; Bailey et al., 2009) to search for common regulatory sequences in the 2kb upstream regions of the 410 significantly upregulated genes ($p <0.05$). Interestingly, one of the two significantly overrepresented motifs identified contains several consecutive A/T bases, a sequence that resembles the recently identified LIN-29 binding site (Figure 3; Abete-luzi & Eisenmann, 2018; Narasimhan et al., 2015).

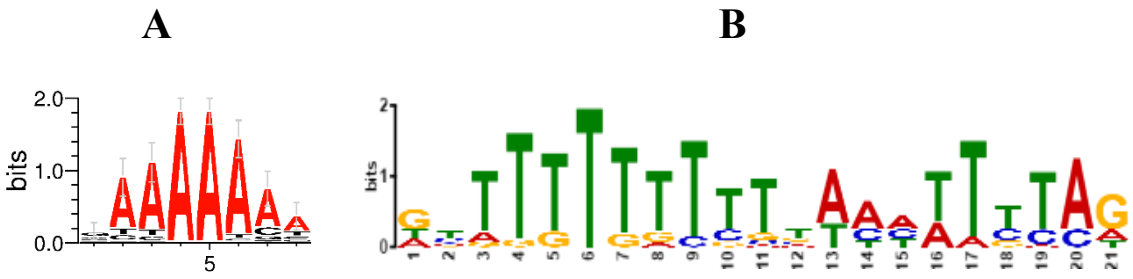


Figure 3. The LIN-29 DNA binding motif is present in promoter regions of LIN-29 upregulated targets. A) consensus LIN-29 binding motif determined from *in vitro* protein binding assays performed with LIN-29 DNA-binding domain (Narasimhan et al., 2015); logo made with WebLogo 3.6.0 (Crooks, Hon, Chandonia, & Brenner, 2004). B) motif obtained by MEME analysis (MEME Suite; Bailey et al., 2009) using 2 kb upstream regions from 410 genes significantly upregulated in *hs::lin-29* animals vs. *hs::control* ($p < 0.05$).

III.d. Genes encoding cuticle constituents as well as enzymes involved in fatty acid metabolism are enriched among LIN-29 targets

By using a cutoff of 1.7-fold or greater, we narrowed our target list and defined 230 and 350 genes that were upregulated and downregulated respectively, upon early overexpression of LIN-29 (Supplemental table). To gain insight into the role of these differentially regulated genes we characterized them based on their molecular function (Table 5). The major categories of upregulated genes included those encoding: 1) proteins of unknown function (45%; n=104), cuticle collagens (14%; n=33), seven transmembrane receptors (5%; n=12), transcription factors (5%; n=12) zinc metalloproteases (4%; n=10)

and C-type lectins (4%; n=9). Among the cuticle collagens we found *col-19, col-38*, *col-49*, *col-63*, *col-138* and *bli-1*, which were previously shown to be regulated by LIN-29 (Abete-Luzi & Eisenmann, 2018; J E Abrahante, Miller, & Rougvie, 1998; Liu et al., 1995). For the downregulated genes, the major categories were: 1) proteins of unknown function (51%; n=180), enzymes functioning in fatty acid metabolism (7%; n=25), F-box proteins (5%; n=18) and transcription factors (5%; n=17). In addition to these gene classes, other types of genes that were found in large numbers in the set of all 1,101 LIN-29 regulated genes were those encoding nuclear hormone receptor transcription factors (25 genes), cytochrome P450 enzymes (11 genes), UDP glycosyl transferase genes (11 genes), and proteins with transthyretin domains (11 genes) or prion-like (polyQ/N) domains (10 genes).

We performed Gene Ontology analyses using DAVID (D. W. Huang et al., 2009) and PANTHER (Carbon et al., 2009; Munoz-Torres & Carbon, 2017) classification systems to search for overrepresented categories of genes in three data sets: all our significant target genes (1,101) and both the up- and downregulated genes (230 and 350, respectively) (Table 6). The single significantly enriched category among the upregulated genes was '*structural constituent of cuticle'*. This group of 33 genes includes the six cuticle *col* genes previously shown to be regulated by LIN-29, but also 27 other cuticle collagen genes. The fact that 33 of the 187 *col* genes in *C. elegans* were found to be upregulated in our analysis suggests they are a major target of regulation by LIN-29.

Table 5. Main categories of gene products among LIN-29 regulated genes

	Genes with significant differential expression		1.7-fold upregulated genes		1.7-fold downregulated genes	
Category	**Count**	**%**	**Count**	**%**	**Count**	**%**
unknown	502	45.6%	104	45.2%	180	51.4%
lipid metabolism (Zhang)	65	5.9%	2	0.9%	25	7.1%
cuticle collagen	46	4.2%	33	14.3%	3	0.9%
F box protein	26	2.4%	3	1.3%	18	5.1%
nuclear hormone receptor	25	2.3%	4	1.7%	11	3.1%
other transcription factor	21	1.9%	8	3.5%	6	1.7%
C-lectin	14	1.3%	9	3.9%	1	0.3%
7TM receptor	13	1.2%	12	5.2%	1	0.3%
cytochrome P450	11	1.0%	3	1.3%	6	1.7%
transthyretin-domain	11	1.0%	0	0.0%	6	1.7%
UDP-glucuronosyltransferase	10	0.9%	3	1.3%	3	0.9%
prion-like-(Q/N-rich)-domain	10	0.9%	3	1.3%	2	0.6%
ncRNA	8	0.7%	2	0.9%	0	0.0%
O-acyltransferase	8	0.7%	2	0.9%	5	1.4%
nematode specific protein B	6	0.5%	5	2.2%	0	0.0%
zinc metalloprotease	6	0.5%	10	4.3%	2	0.6%
protein phosphatase	4	0.4%	4	1.7%	2	0.6%
extracellular signaling protein	4	0.4%	4	1.7%	0	0.0%
neuropeptide like protein	4	0.4%	0	0.0%	2	0.6%
peroxisomal assembly factor	3	0.3%	0	0.0%	3	0.9%
solute carrier protein	3	0.3%	0	0.0%	3	0.9%

'Genes with significant differential expression' are those with a differential change between *hs::lin*-29 and *hs::*control strains; $P<0.05$ (n=1,101). The other gene sets are those genes showing

1.7-fold or greater upregulation (n=230) or downregulation (n=350) in *hs::lin-29* compared to *hs::control* and $P < 0.05$.

The most highly enriched category for the downregulated target genes was '*peroxisomal membrane*': this category contained three genes encoding peroxisomal assembly factors (*prx-1, prx-5, prx-11)*, and two genes encoding enzymes acting within the peroxisome (*ndx-8, maoc-1)* (Zhang et al., 2010; WormBase). Surprisingly, the second highest overenriched category among the downregulated genes was '*fatty acid metabolism'*. To further explore the potential relevance of this result, we compared our gene target sets to a recent compendium of 471 *C. elegans* genes known to be involved in lipid metabolism (Y. Zhang et al., 2013). We found 25 of our 350 downregulated genes on this list, a number significantly higher than that expected by random sampling (hypergeometric $P = 4.61E-7$; Table 6; Supplemental Table). Likewise, when we searched the list of all 1,101 LIN-29 regulated genes against the lipid metabolism gene list, we found 65 genes in total, 57 of which were downregulated to some extent upon LIN-29 misexpression ($P = 1.72E-12$; Table 6; Supplemental Table). A potential link between the two enriched downregulated gene categories is the fact that many of the downregulated metabolic genes act in fatty acid beta-oxidation, a process which occurs in the peroxisome (and mitochondria). Indeed, three lipid metabolic genes that are downregulated 30-50% by overexpression of LIN-29 function in peroxisomal beta-oxidation (*maoc-1, dhs-28,* and *daf-22*), and reduction of their function by RNAi has been shown to cause an increase in lipid droplet size (S. O. Zhang et al., 2010). Lipid droplets are evolutionarily-conserved, triacylglycerol containing fat storage sub-cellular organelles found in the intestine and in the hypodermis of *C. elegans*

(Mak, 2012). Similarly, reduction of function of the three peroxisomal assembly factor genes we identified (*prx-1, prx-5, prx-11*) also leads to an increase in lipid droplet size (Zhu, Liu, Zhang, & Liu, 2018), suggesting that both of these enriched categories of downregulated genes may impinge on lipid storage and utilization. Also of note is the identification of nuclear hormone receptor gene *nhr-80* as the gene most downregulated upon *lin-29* induction (5.5-fold decrease): NHR-80 is known to physically interact with NHR-49 to regulate genes involved in fatty acid metabolism, including *fat-5* which was downregulated almost 2-fold upon LIN-29 overexpression (Brock, Browse, & Watts, 2006; Pathare, Lin, Bornfeldt, Taubert, & Van Gilst, 2012a; Van Gilst, Hadjivassiliou, Jolly, & Yamamoto, 2005).

Overall, two of the overrepresented categories among our LIN-29 targets are cuticle constituents (i.e. cuticle collagens), and metabolic enzymes (especially those involved in fatty acid metabolism). While the first category confirms—and expands—a previously known role for LIN-29 in the regulation of cuticle collagen gene expression, the other category supports the exciting possibility that LIN-29 acts in a regulatory network that controls developmentally-linked metabolic changes that are part of the larva-to-adult transition (Abete-Luzi & Eisenmann, 2018; Dowen et al., 2016; Liu et al., 1995).

Table 6. Enrichment analysis of LIN-29 target genes

All LIN-29-regulated targets (n=1,101)		Gene count				
Category/GO term		in genome	found	expected	Fold	*P* value
GO biological process	peroxisome organization	18	7	0.97	7.23	1.89E-04
	fatty acid metabolic process	105	21	5.65	3.72	1.56E-06
	defense response to Gram-positive bacterium	60	12	3.23	3.72	2.65E-04
	innate immune response	345	57	18.56	3.07	2.55E-12
	oxidation-reduction process	618	60	33.25	1.8	3.76E-05
	cellular response to chemical stimulus	612	57	32.92	1.73	1.57E-04
GO molecular function	structural constituent of cuticle	168	48	9.04	5.31	2.90E-18
	iron ion binding	118	20	6.35	3.15	2.38E-05
	oxidoreductase activity	539	56	29	1.93	1.08E-05
GO cellular component	peroxisomal membrane	18	8	0.97	8.26	3.08E-05
	membrane raft	80	14	4.3	3.25	2.87E-04
	extracellular space	294	37	15.82	2.34	7.31E-06
lipid metabolism*	n/a	471	65	25.33	2.56	1.72E-12

1.7-fold upregulated targets (n=230)		Gene count				
Category/GO term		in genome	found	expected	Fold	*P* value
GO molecular function	structural constituent of cuticle	168	33	1.9	17.36	5.46E-29
lipid metabolism*	n/a	471	2	5.29	0.94	6.84E-02

1.7-fold downregulated targets (n=350)		Gene count				
Category/GO term		in genome	found	expected	Fold	*P* value
GO biological process	fatty acid metabolic process	105	11	1.79	6.16	3.99E-06
	innate immune response	345	18	5.87	3.07	4.40E-05
GO cellular component	peroxisomal membrane	18	5	0.31	16.34	3.33E-05
lipid metabolism*	n/a	471	25	8.05	3.11	4.61E-07

Enrichment analyses were done for all LIN-29 significant targets and for both 1.7-fold LIN-29 up- and downregulated subsets using Gene Ontology Consortium (see Methods). *Genes were also compared to the list of *C. elegans* metabolic genes from Zhang et al. (2013).

III.e. The intersection of *gain-of-function* and *loss-of-function* transcriptomic data identifies a set of high confidence LIN-29 target genes

We compared our list of target genes differentially regulated by overexpression of wild type LIN-29 in the L3 stage to data from Hunter et al. which examined gene expression in *lin-29(n333)*(*lf*) animals versus wild type in the L4 stage (Hunter et al., 2013). Although *n333* is a known loss-of-function allele, the size and levels of *lin-29* transcripts are not altered in *n333* mutant animals, and the molecular nature of this mutation is unknown (Rougvie & Ambros, 1995; WormBase). The intersection of these sets of data gives a list of 21 strong candidates for LIN-29 activated genes: genes that went up in our gain-of-function/temporal misexpression approach and down in the Hunter et al. loss-of-function data (hypergeometric $P = 5.40E-06$), and a list of 35 genes likely to be directly or indirectly repressed by LIN-29 (genes that went down in our gain-of-function approach and up in the Hunter et al. loss-of-function data; hypergeometric $P= 2.34E-20$). . On average, the LIN-29 'activated 'genes are upregulated to a greater extent than the 230 upregulated genes as a whole (3.9-fold vs. 3.1-fold, respectively), while the LIN-29 'repressed' genes are affected only slightly more than the downregulated genes taken together (2.4-fold vs. 2.1-fold). We refer to the genes in common between these two transcriptomic data sets as '*bona fide'* LIN-29 target genes and consider that these genes represent some of the best candidates for true LIN-29 target genes (Supplemental Table).

We were surprised by the small number of genes in this overlap, particular the low number of cuticle collagen genes among the *bona fide* activated genes (four genes). We previously

showed that RNA interference targeting *lin-29* in wild type animals significantly reduced expression of the L4 cuticle *col* genes *col-38, col-49, col-63* and *col-138* in the L4 stage (Abete-Luzi & Eisenmann, 2018), and *col-49* was the most highly upregulated gene upon LIN-29 overexpression (48 fold; Supplemental Table). Yet of these four genes, only *col-38* was found as a *bona fide* activated target (Supplemental Table). Examination of the *lin-29(n333)* data shows that *col-49, col-63* and *col-138* were downregulated 3-to 10-fold in the *lin-29(n333)* mutant, however the data for these genes was slightly above the $P < 0.05$ cutoff. This suggests one (statistical) reason for the small overlap between the data sets. Likewise, we were surprised that very few of our 'metabolic' downregulated targets were in common with the *lin-29(n333)* data set for significantly upregulated genes. Given the unknown nature of the *lin-29*(*n333*) allele (Rougvie & Ambros, 1995), we decided to examine whether reduction of *lin-29* function by RNA interference caused increased expression of these metabolic downregulated genes in the L4 stage when LIN-29 protein normally accumulates. We chose three downregulated genes encoding enzymes that function in lipid metabolism (*acs-7*, *dhs-18* and *hacd-1*) and examined their expression in the L4 stage following *lin-29* RNAi treatment. Consistent with our *hs::lin-29* results, all three genes showed increased transcript levels in *lin-29* RNAi treated animals (Figure 4), suggesting that at least these three genes are indeed repressed by LIN-29 activity at the L4 stage, yet for unknown reasons, the *lin-29*(*n333*) allele failed to derepress them.

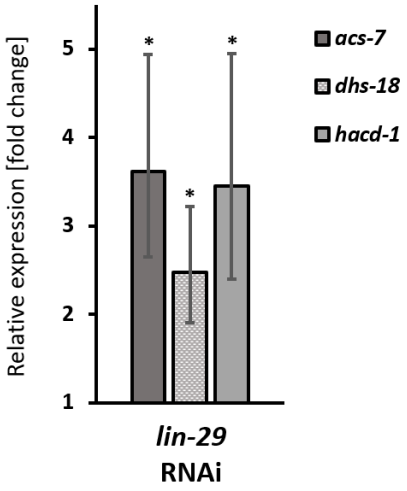


Figure 4. LIN-29 represses intestinal genes involved in fatty acid beta-oxidation in the L4 stage. Endogenous expression of intestinal genes *acs-7*, *dhs-18* and *hacd-1* was assessed in the L4 stage in *lin-29*(*RNAi*) animals by RT-qPCR. Quantification was relative to expression in animals treated with empty vector RNAi control. Error bars represent standard errors of the mean. **P*<0.05 (unpaired *t* test).

III.f. Temporal expression patterns of LIN-29-regulated genes

If the upregulated genes we identified upon early misexpression of LIN-29 are actual targets of LIN-29 regulation during the L4 stage, we would predict that these genes may show an increase in expression in the L4 stage during normal development. Conversely, genes in our downregulated gene set would be predicted to decrease in expression at that time. To determine the pattern of temporal expression for our set of LIN-29 targets, we examined modENCODE developmental expression data for these genes (Gerstein et al., 2010), and categorized them based on whether they show a peak of expression in any particular stage of the worm life cycle (see Jackson et al., 2014). We then compared the pattern of temporal expression of our LIN-29 target genes to the pattern for all 16,183 *C. elegans* genes in the modENCODE data sets (Figure 5).

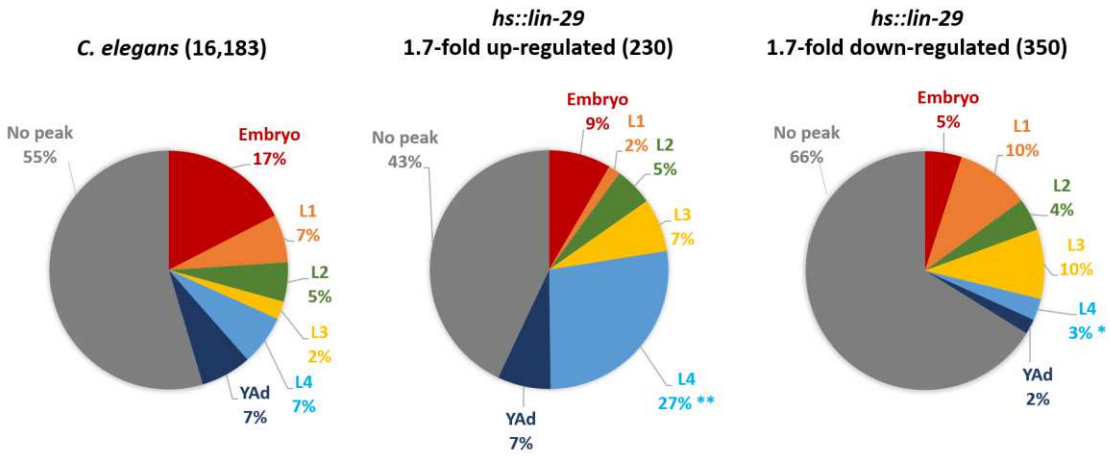


Figure 5. Genes which normally peak in the L4 stage are overrepresented among LIN-29 upregulated targets. Temporal expression peaks were assessed for the indicated gene sets based on modENCODE RNA-Seq data (Gerstein et al., 2010) using criteria from Jackson et al (2014): genes showing 50% or more of their total developmental expression in one stage were identified as having a 'peak' in that stage, the remainder are 'no peak'. Distributions for all genes in each set were calculated and are displayed as percentages. Compared to the genomic distribution (left), genes that peak in the L4 stage are overrepresented and underrepresented in the sets of LIN-29 target genes that were upregulated and downregulated, respectively; ** P <0.0001 and * P <0.01 (Chi-square with Yates correction).

Notably, the proportion of our upregulated genes that normally show a peak of expression in the L4 stage is significantly larger than the percentage of L4-peak genes in the genome as a whole (27% versus 7%; Chi-square with Yates correction P <0.0001; Figure 5). For example, 24 of the 33 upregulated cuticle collagen genes show a single peak of expression in either the L4 or young adult stage, while an additional 6 of 33 *col* genes show peaks in the L2 and L4 stages (Figure 6). We also looked at the distribution of temporal expression

patterns for the smaller '*bona fide* LIN-29 activated' subset and observed that genes that peak in the L4 stage are even more overrepresented (43% versus 7%; Chi-square with Yates correction $P < 0.0001$).

On the other hand, in the case of our LIN-29 downregulated targets, the proportion of genes that normally peak in the L4 stage was significantly smaller than that expected based on the known genomic distribution (3% versus 7%; Chi-square with Yates correction $P = 0.0064$; Figure 5). When we examined the temporal expression patterns of the *bona fide* repressed subset of genes we noted that the proportion of genes with a peak of expression in stages before the L4 (63%) was much higher than that observed in both the *bona fide* activated (28%) and the total genomic sets (31%). Additionally, we examined postembryonic expression profiles generated from the modENCODE data for the 25 lipid metabolism genes downregulated 1.7-fold or more by LIN-29 expression and found that 20/25 showed either a permanent (e.g., *dhs-18*, Figure GExplore) or temporary (e.g. *elo-5*, Figure 6) down-regulation in the L4 stage.

Together these trends are consistent with the hypothesis that many of the genes we identified as upregulated upon misexpression of LIN-29 in the early L3 are normally upregulated by the peak of LIN-29 protein in the L4 stage during development, and that many of the genes we identified as downregulated upon LIN-29 temporal misexpression may be expressed earlier in larval life and are normally downregulated in the L4 stage when LIN-29 levels peak**.**

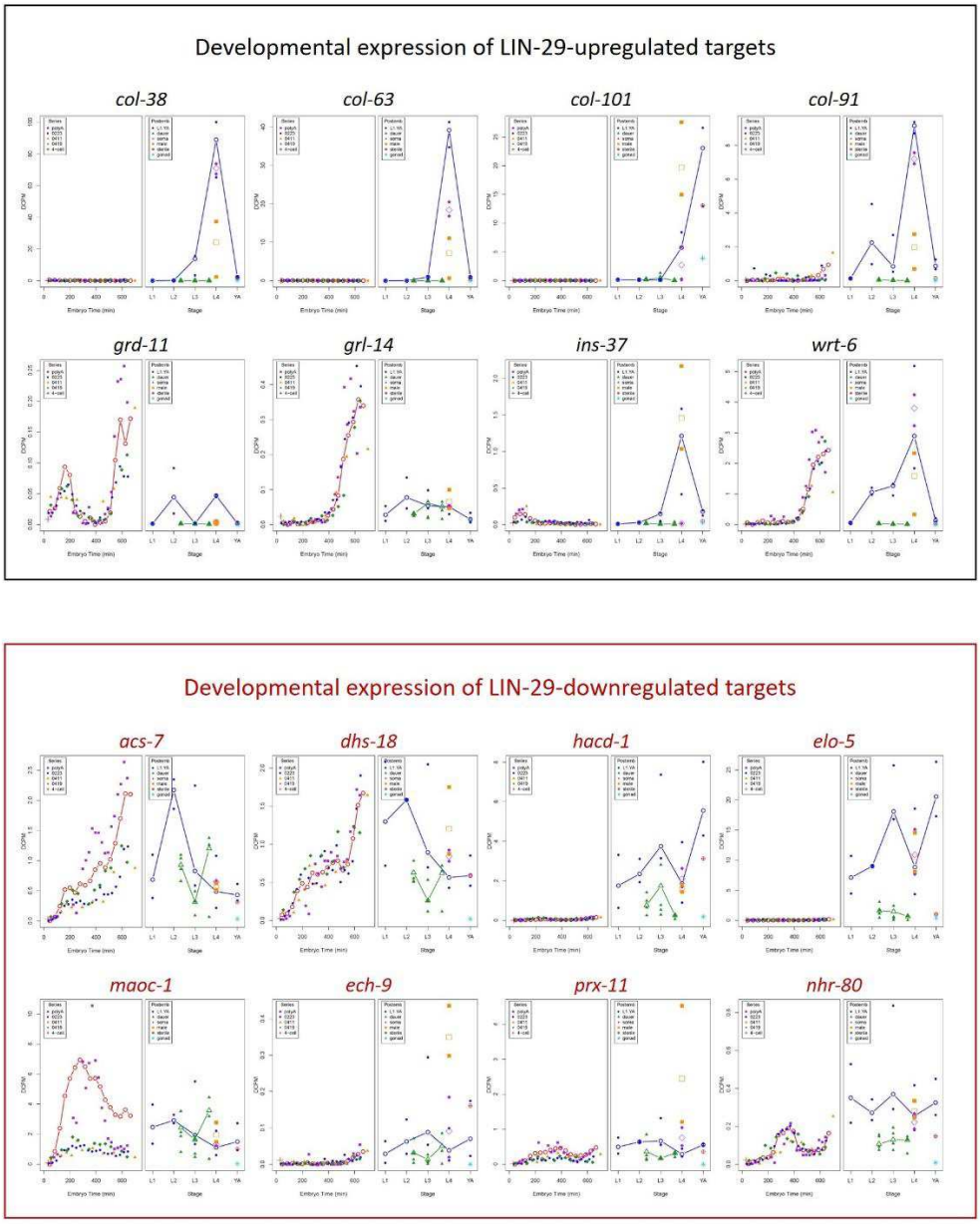


Figure 6. Developmental RNA-Seq data of LIN-29 target genes. Wild type expression data from modENCODE (Gerstein et al., 2010) is shown for eight LIN-29-upregulated (top, black box) and eight LIN-29-downregulated (bottom, red box) targets. For each gene, the left panel shows embryonic expression which is not relevant to this work, while the blue line in the right panel shows postembryonic expression values for each gene in the L1, 2, L3, L4 and young adult stages. Graphical representations were obtained with GExplore 1.4 (Hutter & Suh, 2016).

III.g. Spatial expression patterns of LIN-29-regulated genes include the intestine

We also examined the known expression patterns of our LIN-29-differentially regulated genes in the *C. elegans* database (Supplemental Table; Wormbase; Lee et al., 2018; see Methods). There is published spatial gene expression data for 193 of the 230 upregulated genes; and based on this data, almost 75% of the upregulated genes show expression in at least one tissue known to express LIN-29 protein (Figure 7; Bettinger et al., 1996; Harris & Horvitz, 2011). Consistent with this, when the pattern of spatial expression for the upregulated genes is compared to the genome as a whole by Tissue Enrichment Analysis (TEA; Angeles-Albores et al., 2016), the most overrepresented expression site is the 'epithelial system' with 74/193 of the upregulated genes (P = 1.0E-08; including all of the cuticle *col* genes with known expression) showing expression in this tissue, which is a major site of LIN-29 expression. For the downregulated genes, there is spatial expression information for 316/350; similar to the upregulated genes, 78% of the downregulated genes show expression in at least one tissue known to express LIN-29 (Figure 7). The fact that 75% or more of our up- and down regulated genes express in sites where LIN-29 is present is consistent with these genes being targets of LIN-29 during normal development.

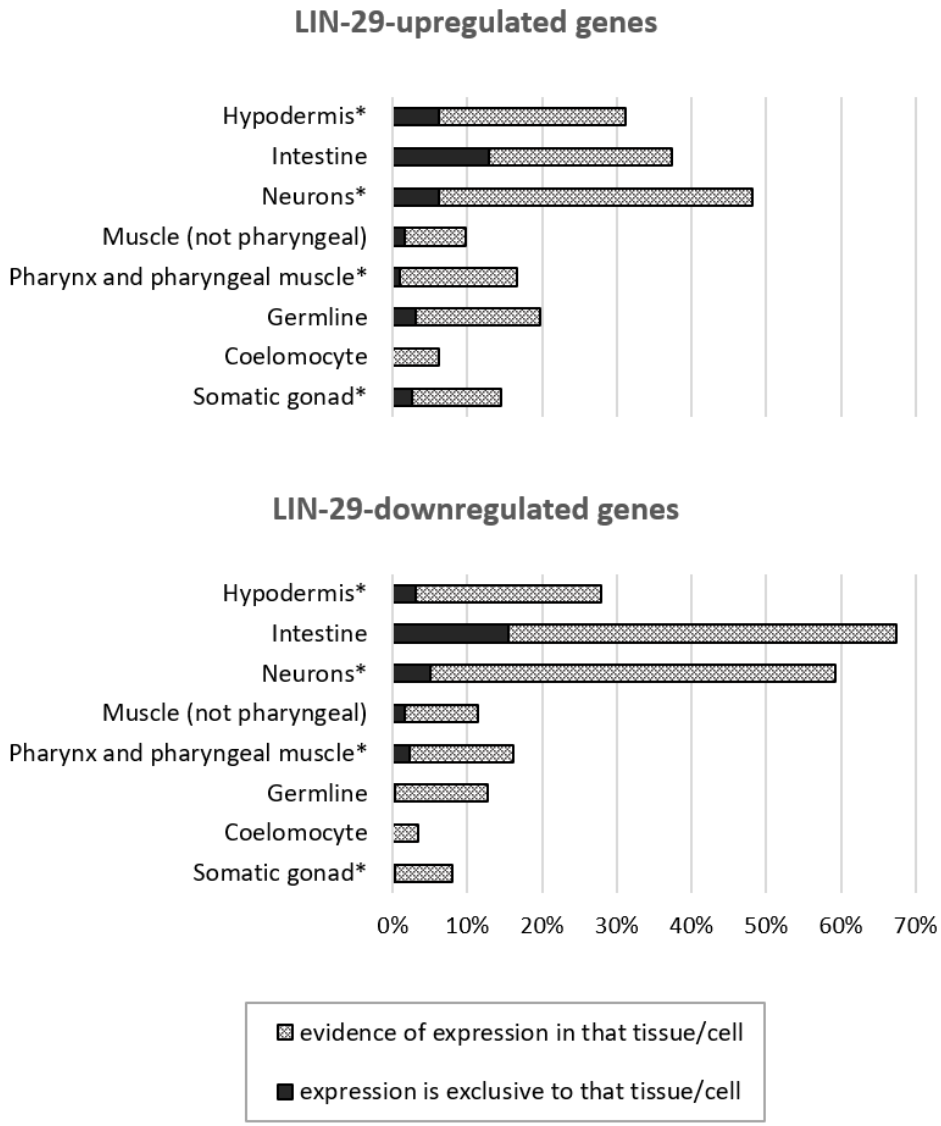


Figure 7. Spatial expression patterns of LIN-29 target genes. Spatial expression data available for 193 of 230 upregulated genes (top) and for 316 of 350 downregulated genes (bottom) was obtained from WormBase (Lee et al., 2018) and plotted as percentages of genes with expression in the indicated tissues. Tissues where LIN-29 is known to be expressed are denoted with an asterisk.

On the other hand, for both the up- and downregulated gene sets, over 20% of the genes show expression in tissues not known to express LIN-29. For both gene sets this site is most often the intestine, although this trend is much more prevalent for the downregulated genes: 213/316 downregulated genes have intestinal expression and 49 of those have only intestinal expression (Figure 7; Supplemental Table). When we performed tissue enrichment analysis for the downregulated gene set, the two overrepresented expression sites with the greatest numbers of genes are 'intestine' (210 genes; TEA P = 1.5E-28) and 'epithelial system' (100 genes; TEA P = 3.20E-06). The same result was obtained when we performed this analysis on the genes that show a significant upregulation in expression in *lin-29(n333)* reduction-of-function mutants (those that are normally repressed by *lin-29* activity; Hunter et al., 2013): 'epithelial system' (144/305 genes) and 'intestine' (130/305 genes). Finally, among the 65 genes involved in lipid metabolism that we identified as possible LIN-29 targets (from the full set of 1,101 significant *hs::lin-29*-responsive genes), 75% of them show expression in the intestine, a major site of metabolic activity in the worm, that is not known to express LIN-29.

One explanation for this result is that expression from the heat shock promoter led to the presence of LIN-29 in the intestine where it is not normally found, which bound to these genes and directly regulated their expression. Alternatively, LIN-29 expression in another tissue could have caused indirect (cell-non-autonomous) regulation of these genes in the intestine. Supporting the latter idea is the fact that at least some of these metabolic genes are almost exclusively expressed in the intestine (i.e. *acs-7*, *dhs-18* and *hacd-1*; Supplemental Table) and are also derepressed by *lin-29* RNAi treatment, a situation in

which no ectopic/intestinal LIN-29 is involved (see above; Figure 4). There is also precedent for the idea of a signal from the hypodermis to the intestine indirectly regulating intestinal gene expression (see below). Together, these observations suggest that in addition to acting cell-autonomously to regulate gene expression in the hypodermis and other tissues in the L4 stage, LIN-29 may also act cell-non-autonomously to regulate expression of genes in the intestine, including many genes involved in lipid metabolism.

III.h. LIN-29-activated signaling molecules, WRT-6 and INS-37, are required for the regulation of LIN-29 intestinal targets

Our results on LIN-29-dependent downregulation of intestinal gene expression suggests a possible role for LIN-29 in coordinating intestinal metabolic activity at the larval to adult transition from other tissues (i.e. hypodermis). Indeed, MacNeil et al. showed that hypodermis-specific transcription factors (e.g. LIN-26) can regulate the expression of reporters for intestinal genes (e.g., *acdh-1,* encoding lipid metabolic enzyme acyl-CoA dehydrogenase) and proposed the existence of a signal that propagates regulatory information from one tissue to another (MacNeil et al., 2015). Furthermore, Dowen et al. hypothesized the activation of an unknown secreted signal by LIN-29 in the hypodermis that mediates LIN-29-dependent regulation of vitellogenin gene expression in the intestine, acting through both DAF-2/insulin receptor signaling and the mTORC2 pathway (Dowen et al., 2016; Murphy et al., 2003).

Interestingly, we identified four genes encoding signaling molecules among our LIN-29 upregulated genes: three encode *C. elegans* proteins related to the Hedgehog family of signaling proteins (*grd-11, grl-14, wrt-6*) and one encodes an insulin-like peptide (*ins-37*). Expression of three of these signal genes goes down in the *lin-29(n333)* data set (Hunter et al., 2013) (*grl-14* (0.57, *P*= 0.065); *wrt-6* (0.17, *P*= 0.046); *ins-37* (0.51, *P*= 0.058). When examining the wild type postembryonic expression data (modENCODE) for these genes, we observed that while *grd-11* and *grl-14* show low expression during larval life, both *wrt-6* and *ins-37* show a marked peak of expression in the L4 stage, when hypodermal LIN-29 protein is active (Figure 6). Moreover, a *wrt-6* reporter is expressed only in the hypodermis and in the socket cells of the amphids (Aspöck, Kagoshima, Niklaus, & Bürglin, 1999).

We individually tested the requirements of *wrt-6* and *ins-37* for both the activation of *vit* genes and the repression of LIN-29-downregulated intestinal metabolic gene targets in the late L4 stage. In the *ins-37* RNAi treated samples, endogenous expression of *vit-1* and *vit-6* showed little or no change, but the expression of *vit-3* was strongly decreased (Figure 8). We also checked the transcript levels of three downregulated, intestinal targets encoding lipid metabolic enzymes (*acs-7*, *dhs-18* and *hacd-1*) in *ins-37*(*RNAi*) animals and found that all three were increased almost 2-fold indicating they are normally repressed in the L4 by *ins-37* activity (Figure 8). In the case of *wrt-6*(*RNAi*) animals, while *acs-7* and *dhs-18* expression showed no obvious change, *hacd-1* transcripts exhibited a marked increase of expression, indicating repression in the L4 stage by *wrt-6* activity. Interestingly, all three *vit* genes showed an almost two-fold reduction in expression in these animals, suggesting a broader WRT-6 requirement for *vit* genes (Figure 8). Together, these results suggest that

both *wrt-6* and *ins-37*, which we identified as upregulated targets of LIN-29, encode signals that play a role in mediating the cell-non-autonomous regulation of at least some of the intestinal targets of LIN-29.

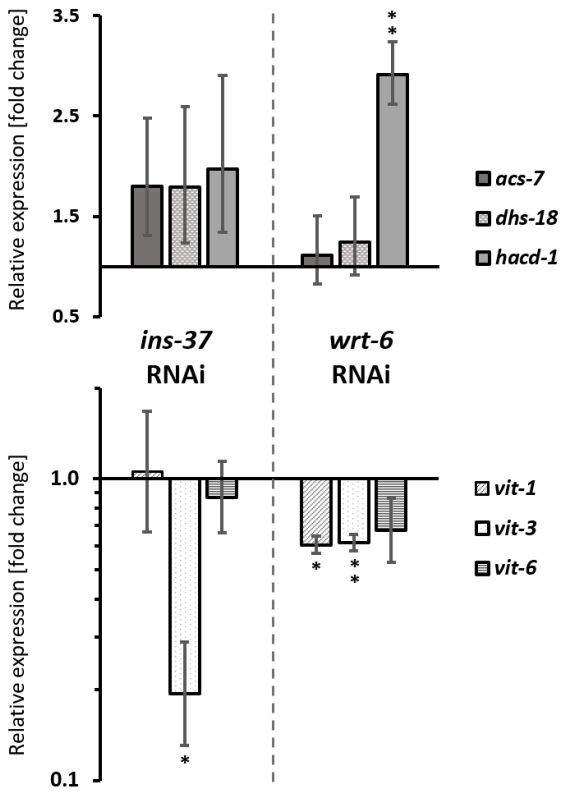


Figure 8. *ins-37* and *wrt-6* are required for proper regulation of gene expression of LIN-29 intestinal target genes in the L4 stage. Endogenous expression of three LIN-29-downregulated metabolic genes (*acs-7*, *dhs-18* and *hacd-1*; top), and three LIN-29-dependent vitellogenin genes (*vit-1*, *vit-3* and *vit-6*; Dowen et al., 2016; bottom) was evaluated by RT-qPCR in late L4-staged *ins-37*(RNAi) or *wrt-6*(RNAi) animals. All six genes express in the intestine and not the hypodermis. Quantifications were relative to expression in animals treated with empty vector RNAi control. Error bars represent standard errors of the mean. $^{**}P<0.001$ and $^{*}P<0.05$ (unpaired *t* test).

III.i. Misexpression of LIN-29 in the adult shortens lifespan

Previous work showed that two heterochronic genes that function early in the larval life, *lin-14* and *lin-4*, can affect lifespan even when their expression is manipulated solely in the adult (Boehm & Slack, 2005). This was a surprise because both *lin-14* and *lin-4* were known as key regulators of the L1 to L2 transition and any effects in the adult were unknown. We know that the expression of *lin-29* in the L4 promotes the normal developmental transition to the adult for a few tissues (i.e. skin and vulva). We wondered whether this LIN-29 'maturing' instruction could have any beneficial or adverse effect on the aging of the animal if overexpressed during adulthood. We tested multiple pulses of LIN-29 in the adult and assessed survival rates in four different conditions: a single heat shock every 24 hrs or every 48 hrs and fed with either dead bacteria or live bacteria (for a total of four assays). Worms subjected to these protocols showed no change in foraging behavior, no altered pharyngeal pumping, no other visible phenotype or sickness, yet all four experiments showed significantly shorter mean lifespan and maximum lifespan (see Figure 9 and Table 7).

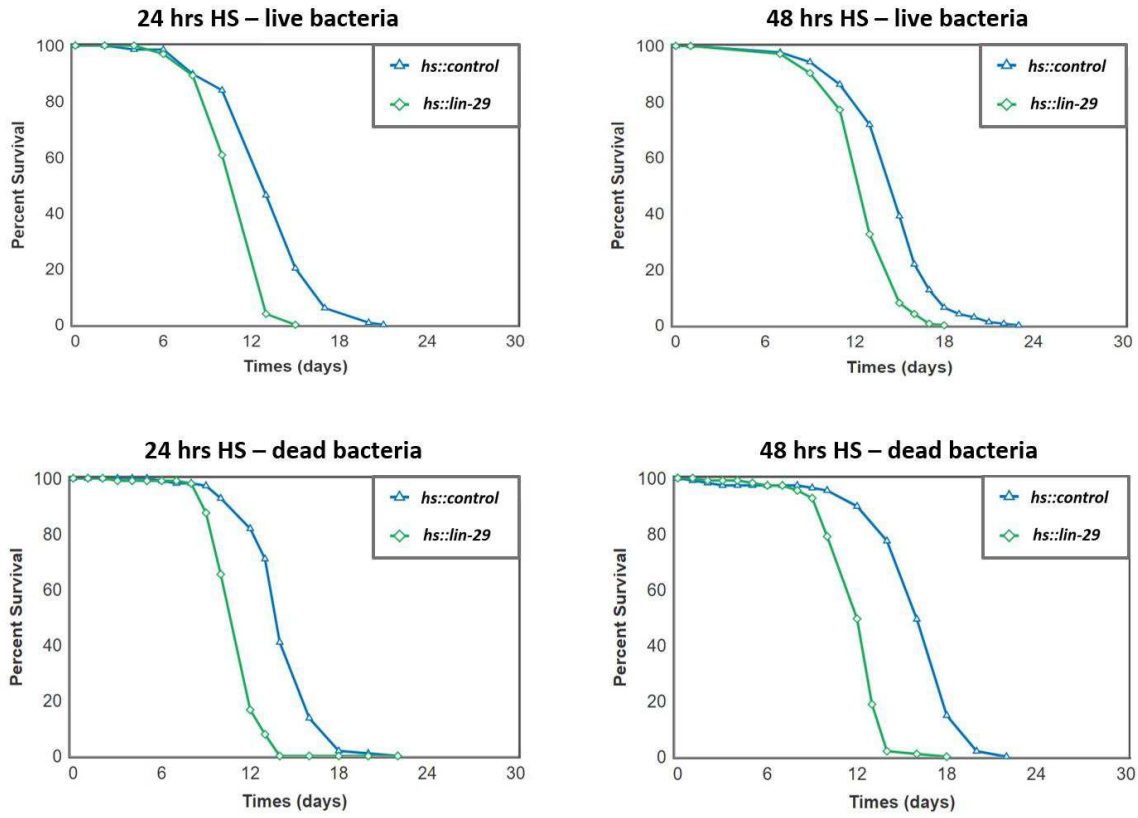


Figure 9. Periodic adult overexpression of LIN-29 shortens lifespan. Synchronized adult animals carrying either *hs::lin-29* or *hs::control*, were periodically exposed to heat shock every 24 or every 48 hours, and fed with either live or dead bacteria. Cohorts were sterilized and followed until the last individual died (see Methods). Lifespan differences were assessed via Kaplan–Meier analysis and *P* values were calculated using log-rank test; in all four conditions both mean and maximum lifespan were significantly shorter (See Table 7). Survival plots were generated with OASIS 2 (Han et al., 2016b).

Table 7. Survival analysis of *hs::lin-29* animals periodically induced in adulthood

Time of heat shock	*E. coli* (diet)	strain	N	Mean lifespan			Abs. max (days)
				days	SE	95% C.I.	
every 24 hrs	live	*hs::control*	140	13.78	0.27	13.26 ~ 14.30	21
		hs::lin-29	140	11.63	0.18	11.27 ~ 11.98	15
every 24 hrs	dead	*hs::control*	114	14.46	0.24	13.99 ~ 14.92	22
		hs::lin-29	109	11.36	0.16	11.05 ~ 11.68	14
every 48 hrs	live	*hs::control*	174	14.89	0.21	14.48 ~ 15.29	23
		hs::lin-29	175	13.07	0.16	12.75 ~ 13.39	18
every 48 hrs	dead	*hs::control*	115	16.36	0.32	15.73 ~ 16.99	22
		hs::lin-29	117	12.11	0.19	11.73 ~ 12.49	18

Survival was analyzed in four independent experiments following the indicated conditions. Mean lifespans were estimated for each group considered separately using the Kaplan-Meier method. SE refers to standard error. Absolute maximum lifespan is the time at which cohorts reached 100% mortality. In all cases $P < 0.0001$ (log-rank test).

While we do not know the cause of this shortened lifespan, our results are consistent with a model in which repeated adult overexpression of *lin-29* over time results in the constant repression of genes encoding metabolic enzymes that are required to keep proper metabolic homeostasis. To corroborate whether LIN-29 is capable of repressing metabolic targets in the adult context, we assessed the expression of three metabolic LIN-29-repressed targets after induction in gravid adults and found that both *dhs-18* and *hacd-1* were downregulated (Figure 10) suggesting that perhaps metabolic functions may indeed be perturbed in these animals.

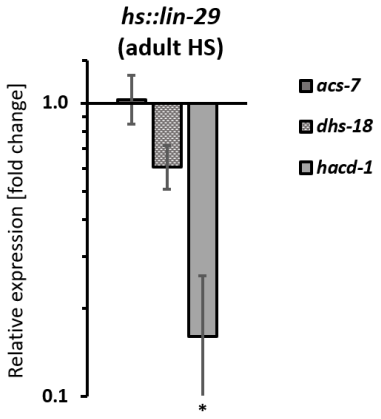


Figure 10. Adult induction of LIN-29 is sufficient to represses expression of *dhs-18* and *hacd-1*. Expression of *acs-7*, *dhs-18* and *hacd-1* in *hs::lin-29* animals after adult heat shock induction was evaluated by RT-qPCR. Quantification was relative to expression in *hs::control* animals. Error bars represent standard errors of the mean. *$P<0.05$ (unpaired *t* test)

III.j. LIN-29-downregulated target proteins and their associations are conserved

Finally, we wondered whether the LIN-29 metabolic targets we identified were conserved and, perhaps part of an evolutionarily-conserved metabolic response. Could LIN-29 be the master switch of a well-established regulatory network which we do not know much about? To learn more about these metabolic gene products and how they relate, we searched for known interactions and associations among several LIN-29 metabolic targets in *C. elegans* as well as among their human homologs, and generated protein-protein association networks with STRING database and search tool (Szklarczyk et al., 2015). Many of these gene products as well as their affiliations appear to be conserved (Figure 11). Given that Egr-1, the proposed mammalian homolog of LIN-29 (see Chapter I section V.f.), plays a role in the regulation of energy expenditure in adipocytes and its expression is highly correlated with dietary-induced obesity and related disorders (Zhang et al., 2013), a role

for LIN-29 in lipid metabolic gene regulation could be conserved in vertebrates as well and should be investigated.

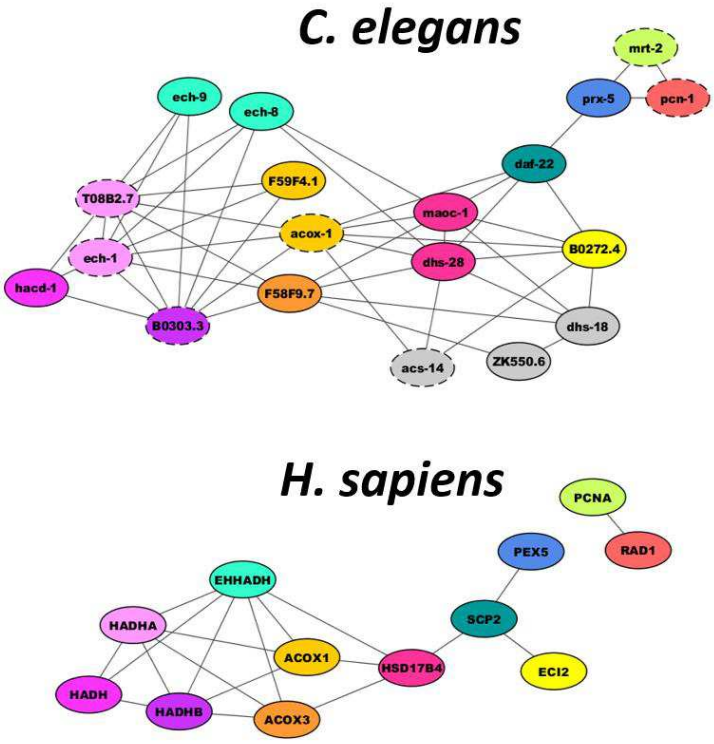


Figure 11. Conserved protein-protein associations among products of fat metabolism genes downregulated by ectopic LIN-29 expression. Homologous gene products are represented by nodes of the same color. Nodes with solid lines are LIN-29 targets. Nodes with dashed lines are not targets yet STRING analysis showed first degree association (see Methods).

IV. Discussion

In this work we used an inducible *lin-29* expression strain to misexpress LIN-29 at an earlier time in development than it is normally expressed, and we performed RNA-Seq analysis with our collaborators to identify differentially expressed genes compared to control animals. We identified 1,101 genes showing significant differential expression upon LIN-29 overexpression, of which 230 were upregulated and 350 were downregulated by 1.7-fold or more. A LIN-29 binding motif was enriched in the promoters of the upregulated genes, and these genes showed spatial and temporal expression patterns consistent with them being targets of LIN-29 in the L4 stage during normal development. The major enriched class of genes upregulated by LIN-29 were those encoding cuticle collagens. While this role for LIN-29 was previously shown (Abete-Luzi & Eisenmann, 2018; Liu et al., 1995; Rougvie & Ambros, 1995), the large number of *col* genes potentially regulated by LIN-29 was not known (we identified 33/187 *col* genes as upregulated by LIN-29). Surprisingly, a major class of overenriched genes among the downregulated targets was those encoding enzymes involved in lipid metabolism, and a second category was genes involved in assembly and function of the peroxisome, where fatty acid beta-oxidation also takes place. Together these results suggest a role for LIN-29 in regulation of lipid metabolism at the larval to adult transition, perhaps via repression of genes involved in fatty acid oxidation. Many of these lipid metabolism genes are expressed in the intestine, which does not normally express *lin-29*, however we showed that expression of three of these genes is increased in the L4 stage upon *lin-29* RNAi. Furthermore, we identified two LIN-29 target genes that encode signaling molecules (*wrt-6* and *ins-37*) that

alter expression of intestinal target genes when their function is reduced. Recently, the Ruvkun lab showed that LIN-29 expressed in the hypodermis acts through insulin and mTORC2 signaling pathways to regulate expression of vitellogenin genes in the intestine (Dowen et al., 2016). Our work suggests that LIN-29 plays an even larger, unanticipated role in lipid metabolic regulation at the L4-to-adult transition than simply regulation of vitellogenin gene expression.

In the following subsections I will discuss different aspects of this work, and in Chapter 4 I will discuss future directions based on the results reported here.

IV.a. Advantages and disadvantages of our gain-of-function misexpression approach

While most previous work characterizing LIN-29 was done using loss-of-function approaches to study *lin-29* function, or by examining the effects of mutations in upstream heterochronic factors that derepress *lin-29* expression, to our knowledge, this work is the first to use an inducible gain-of-function approach to directly express LIN-29 in order to identify target genes of LIN-29 as well as to further investigate its developmental roles. In general, when it comes to investigating transcription factors and their activity, we do not think that either the loss-of-function or the gain-of-function approach is superior, but rather they complement each other. Either gain- or loss-of-function methods will reveal or fail to

reveal biological information depending on the gene interactions and the biological context. For example, target genes that are redundantly regulated by additional factors may not be affected using a loss-of-function approach, yet differences in expression of such target genes might be detected in a gain-of-function approach if expression of the transcription factor of interest is sufficient for regulation. On the other hand, there are cases where a gain-of-function expression approach will not be sufficient to see changes in target gene expression, such as if the transcription factor induction is done in the wrong context (i.e. in the absence of necessary regulatory partners or the required chromatin state).

The activity of a transcription factor can always be analyzed in the *right* context (in our case in the L4 stage) via loss-of-function. Nevertheless, another caveat of the latter approach is that the loss of a transcription factor may have had other developmental consequences during the life of the animal, and some of the observed differences in gene expression at a particular time point in development could be the result of an indirect effect rather than due to direct target gene regulation. In a gain-of-function experiment such as ours, where a wild type genomic background suffers a pulse of transcription factor activity at a discrete time and RNA is collected shortly after, indirect downstream consequences are less likely to have occurred.

The degree of *loss* in a loss-of-function approach adds another variable to consider in this technique: different mutations may result in different final gene products, thus different loss-of-function alleles may affect target genes differently. We believe this could be the

case for the transcriptomic data from microarray analysis of *lin-29*(*n333*) L4-staged animals (Hunter et al., 2013). In this data, cuticle collagen genes such as *col-49*, *col-63*, and *col-138*, and lipid metabolism genes such as *hacd-1*, *acs-7,* and *dhs-18* did not show statistically significant changes in expression ($p < 0.05$; Hunter et al., 2013). Nevertheless, those genes were significantly affected in both *lin-29*(RNAi) (Abete-Luzi & Eisenmann, 2018) and gain-of-function *hs::lin-29* animals (see above section III), as well as in *let-7* loss-of-function animals in which LIN-29 is consequently misexpressed (Hunter et al., 2013). Therefore, we believe that future experiments with *lin-29* should not be done with this 'canonical' allele until the nature of the *n333* mutation is resolved.

A possible caveat that is specific to an approach using a heat shock inducible construct like ours is the systemic expression of a transcription factor and its potential activity in cell types where genes that are not normal targets are now susceptible to its regulation. However in our case, our previous work showed that systemic expression of *lin-29* is not always sufficient to regulate direct target genes ectopically: temporal misexpression of *hs::lin-29* led to temporal misexpression of a *col-38p::yfp* reporter, however it was spatially restricted to the tissue where *col-38* is normally expressed, the hypodermis (Abete-Luzi & Eisenmann, 2018), despite the fact that the heat shock promoter is presumed to be active in all tissues following heat shock. In the same experiment, LIN-29 regulation of *col-38p::yfp* was somewhat temporally restricted as well: while *col-38p::yfp* was expressed after heat shock in the L2/L3 molt and in the adult stage, no activation of the *col-38* reporter was observed when heat shock induction was done in the embryo or in the L1 stage (Abete-Luzi & Eisenmann, 2018), indicating that *hs::lin-29* does not necessarily

regulate target genes only because it was induced, but other context is necessary. Consistent with our results, in the case of the zinc finger transcription factor CHE-1, which is expressed in a small subset of neurons, no ectopic expression of CHE-1 target genes was observed in non-neuronal tissues following induction of *hs::che-1*, because it is prevented by the activity of regulators of histone modification and chromatin organization (Kolundzic et al., 2018; Patel, Tursun, Rahe, & Hobert, 2012; Tursun, Patel, Kratsios, & Hobert, 2011). Together, these examples suggest that what we could initially consider a reason to be cautious when using a heat shock promoter induction, may not be an issue after all.

Broadly, both loss- and gain-of-function approaches have their caveats and their strengths, however each case should be considered differently and the approach to follow should be carefully chosen taking into account the availability of reagents as well as the biological context of the study. In our case, while both approaches tend to complement each other, our gain-of-function protocol using *hs::lin-29* induction seems to be a robust way to study LIN-29 transcription factor activity.

IV.b. Regulation of known targets in *hs::lin-29* animals and identification of *mab-10* as a LIN-29 target

A number of target genes regulated by *lin*-29 have been reported (see Chapter 1 section V) and we were curious to determine which of these known targets were identified by our

gain-of-function approach. Our approach identified several previously known targets responding as expected, specifically L4-specific cuticle collagen genes *col-38 col-49*, *col-63*, *col-138*, *bli-1* and adult-specific *col-19*, all of which showed increased expression by a range of 3- to 48-fold. Also, although below our 1.7-fold change cutoff, known targets *aff-1* and *cki-1* were significantly upregulated by 1.67- and 1.25-fold change, respectively. On the other hand, other known target genes showed no substantial change, such as *cdk-1*, *ceh-2*, *egl-17*, *zmp-1*, *lin-11*, *lag-2*, *blmp-1* and *unc-5*. However, this is not surprising since while these genes are indeed targets of LIN-29, their expression is restricted to one or a few cell types and the RNA used in our approach comes from whole-body samples, making it difficult to detect cell-specific subtle changes with significance. Unexpectedly, we also saw no difference in *nhr-23* expression and an actual increase of 1.45-fold change for *nhr-25* in our *hs::lin-29* animals with respect to control. This result is the opposite from what we would expect given the published repressible role of LIN-29 over these gene(s) (Harris & Horvitz, 2011). In work done by Hunter et al., *lin-29*(*n333*) microarrays also showed no change for *nhr-23*, however *nhr-25* expression was increased (1.90-fold; P= 0.005), which is consistent with LIN-29's repressive effect over this gene (Hunter et al., 2013). One possible explanation is that regulation of *nhr-25* is more intricate than we suspected, and perhaps excess of LIN-29 during earlier stages (i.e. L3 stage) promotes an indirect positive response on the regulation of this gene. Further investigation on the regulation of *nhr-25* would be necessary to address this question.

Based on the decrease in expression of vitellogenin genes *vit-1–vit-6* seen by RT-qPCR analysis of *lin-29(n333)* and *lin-29(RNAi)* animals (Dowen et al., 2016) we expected higher

transcript levels for these genes in the *hs::lin-29* samples. However, we only observed a slight increase in expression for *vit-2* and *vit-6*, which was not statistically significant (Supplemental Table). We can think of two explanations for this result. First, it is possible that LIN-29 is necessary for *vit* gene regulation but not sufficient, perhaps because the proper context for regulation of *vit* genes by LIN-29 (eg., chromatin state or additional regulatory factors) is not available in the L3 stage when we heat-shocked our worms. A second possibility is that *vit* gene expression was repressed by the product of one of our LIN-29 upregulated targets, *mab-3*, which is known to negatively regulate vitellogenin expression in males (Yi & Zarkower, 1999). Interestingly, microarray analysis of gene expression in *lin-29*(*n333*) animals showed slight downregulation of only three *vit* genes (0.50-fold; P= 0.055), *vit-2* (0.66 P=0.07) and *vit-6* (0.72 P=0.01), and only the latter was statistically significant (Hunter et al., 2013). One possibility is that animals used in *lin-29*(*n333*) microarrays (Hunter et al., 2013) were collected in the early L4 stage rather than the late L4 stage when *vit* expression is substantially higher (personal observation, data not shown).

Finally, our approach has identified an important LIN-29 target which was previously overlooked and deserves further comment. *mab-10* encodes a transcription cofactor that physically interacts with LIN-29 and is required for several *lin*-29-dependent processes at the larval to adult transition (Harris & Horvitz, 2011). Previous work indicated that transcription of *mab-10* was not dependent on *lin-29* (Harris & Horvitz, 2011). However, we saw that *mab-10* transcripts levels increase 2.3-fold upon overexpression of *lin-29* (Supplemental Table). Furthermore, *mab-10* levels were reduced in both *lin-29*(*n333*) and

let-7(*n2853*) mutants as well (Hunter et al., 2013). We do not know why this transcriptional regulation of *mab-10* was not observed in the work done by Harris et al., which relied on single molecule FISH analysis of *mab-10* transcripts in *lin-29(n836)* animals, however we believe the regulation of *mab-10* by LIN-29 that we observed is authentic. In fact, this transcriptional regulation may be evolutionarily-conserved, as the LIN-29 orthologs EGR1, EGR2 and EGR3 all regulate expression of the MAB-10 ortholog NAB2 through binding sites in the NAB2 gene promoter (Kumbrink, Gerlinger, & Johnson, 2005; Kumbrink, Kirsch, & Johnson, 2010).

IV.c. LIN-29 effects on vulval and seam cell development

Our early induction experiments showed LIN-29 effects on vulva and seam cell development when overexpressed before the L4 stage. Given LIN-29 known expression pattern and regulatory roles in such cell types, some of the results we obtained were expected. Nevertheless, some of the observed phenotypes were quite intriguing.

LIN-29 normally accumulates in the anchor cell and vulval cells in the L3 stage and is required for proper vulva morphogenesis (Chapter 1 section V). Not surprisingly, its overexpression around the L2 and L3 stages led to vulval abnormalities, especially seen in the adult. In these animals, however, we observed a particularly interesting phenotype which to our knowledge has not been described before: a vulva developmental arrest

around the mid L4 stage that remained in animals even after they reached an adult reproductive stage. Why an excess of LIN-29 in the L3 stage results in the freezing of vulva development a life stage later remains a mystery. One way to explain this is to imagine that perhaps excessive (and prolonged) accumulation of this transcription factor brings in a 'maturing' instruction which prevents further development at a time before vulval eversion takes place. Perhaps further study of worms displaying this phenotype and our LIN-29 target genes may provide clues as to the molecular and cellular nature of this phenotype.

At the L4 to adult transition, the hypodermal seam cells exit the cell cycle and stop dividing, fuse together into one long syncytial seam cell, and secrete the adult cuticle structure called alae (see Chapter 1, section IV). When *lin-29* function is compromised, all three of these processes fail to occur correctly, showing *lin-29* is necessary for all of them (Ambros & Horvitz, 1987). Further, reduction of function of upstream heterochronic pathway genes that repress *lin-29* expression causes these processes to occur precociously (Fielenbach et al., 2007; Lin et al., 2003b; Pasquinelli et al., 2000). We also saw that early expression of LIN-29 alone in the L2 or the L3 can promote precocious seam cell fusion before its normal occurrence in the late L4. Therefore, our work showed that LIN-29 is not only necessary for seam cell fusion but is also sufficient. Surprisingly, LIN-29 alone was not sufficient for the seam cells to form alae precociously. This result suggests that perhaps other factors regulated by the upstream heterochronic pathway members in addition to LIN-29 are also required for alae secretion.

IV.d. LIN-29 and the regulation of body morphology

We found that overexpression of LIN-29 early in development caused body morphology phenotypes in adulthood. Previous work showed that a few *col* genes are required for normal body size, acting either as positive (i.e. *rol-6*, *sqt-1*) or negative (i.e. *lon-3*) regulators of final body size (Nyström et al., 2002; Suzuki, Morris, Han, & Wood, 2002). While *rol-6* was slightly downregulated by LIN-29 in our study, the change was not significant. However, *lon-3* is one of the *bona fide* LIN-29 upregulated genes and its increased levels due to *lin-29* early induction could contribute to the Small phenotype we observed, since overexpression of *lon-3* leads to smaller animals (Nyström et al., 2002).

Several *col* genes have been found to act downstream of DBL-1 (Dpp BMP-like) which is a member of the TGFβ superfamily, and the DBL-1 signaling pathway is well known for its role in the regulation of body size in *C. elegans* (Gumienny & Savage-Dunn, 2013; J. Liang, Yu, Yin, & Savage-Dunn, 2007; Roberts, Gumienny, Gleason, Wang, & Padgett, 2010). Interestingly, other downstream targets of this pathway include metabolic genes, and *hedgehog*-like (*warthog*) and insulin-like signals (J. Liang et al., 2007; Roberts et al., 2010). Specifically, the *warthog* genes *wrt-1* and *wrt-8* (both targets of the DBL-1 pathway) showed Small phenotypes when overexpressed individually (Roberts et al., 2010). Moreover, in separate work, RNAi experiments of other *warthog* genes (including our target genes *wrt-6* and *grd-11*) showed growth defects suggesting again their involvement in body growth regulation (Zugasti, Rajan, & Kuwabara, 2005). These results are consistent with an additional or alternative hypothesis in which perhaps our LIN-29

warthog targets (*wrt-6*, *grd-11*, *grl-14*) and/or *ins-37* contribute to the Small size observed among the *lin-29*-heat-shocked animals. This idea would imply a novel role for LIN-29 as a negative regulator of body size that is consistent with LIN-29 activity at a time when somatic growth rates characteristic of larval life start to decline.

IV.e. LIN-29 regulates intestinal genes via the activation of signaling molecules

We saw a large number of LIN-29-regulated target genes showed normal spatial expression in the intestine, a tissue that does not normally express *lin-29*. Previous work suggested the existence of signals that allow regulation of gene expression in the intestine specifically from the hypodermis (Dowen et al., 2016; MacNeil et al., 2015). In this study, we found two genes encoding extracellular signals that are upregulated by LIN-29 and that participate in the regulation of intestinal gene expression. While intestinal expression changes observed in our *hs::lin-29* animals could have been due to systemic expression from this construct, our results suggest that this may not be the case. First, normally low L4 expression of intestinal metabolic enzymes was increased (derepressed) after *lin-29(RNAi)* treatment, indicating that without any overexpression, LIN-29 activity affects expression of intestinal target genes during normal development. Second, intestinal gene expression was dependent on LIN-29-upregulated signals. The same intestinal genes we tested with *lin-29*(RNAi) showed increased levels after *ins-37*(RNAi) treatment; and in the case of *hacd-1,* expression was derepressed in *wrt-6*(RNAi) as well. Based on these results we believe we are not seeing artifactual expression changes due to use of the heat shock

promoter but are seeing a real cell-non-autonomous effect of LIN-29 on intestinal gene expression.

Previously, expression of intestinal vitellogenin genes was shown to require hypodermally-expressed LIN-29, and an unknown signal was hypothesized to mediate such regulation (Dowen et al., 2016). The identification of *wrt-6* and *ins-37* as targets of LIN-29 prompted us to test if these signaling genes also functioned in vitellogenin gene regulation in the intestine. Strikingly, we found that *wrt-6* is required for proper expression of *vit-1*, *vit-3* and *vit-6* in the L4, and *ins-37* was required for normal L4 expression of *vit-3* as well. Our demonstration of regulation of intestinal gene expression (*vit* genes and lipid metabolic enzyme genes) by WRT-6 and INS-37 makes them the first signaling molecules identified that may mediate heterochronic pathway regulation of intestinal gene expression from the skin during the larva-to-adult switch.

IV.f. LIN-29 and lipid metabolism during the larva-to-adult transition

Although a role for LIN-29 in the cell-non-autonomous regulation of intestinal vitellogenin genes was recently reported (Dowen et al., 2016), a larger role for *lin-29* in the regulation of lipid storage or utilization was not indicated. Given this, one of the most surprising results to us to come from our LIN-29 target gene analysis was the identification of numerous genes involved in lipid metabolism. Genes involved in fatty acid metabolism

and peroxisomal organization were highly enriched among LIN-29 downregulated target genes. In fact, 65 of the LIN-29 differentially-regulated targets were specifically classified as lipid metabolic genes (Y. Zhang et al., 2013) Most of these genes (57/65) were downregulated upon overexpression of *lin-29* (25 of them did so by at least 1.7-fold), and 62% of them show a decrease in transcript levels from the L3 stage to the L4 stage during normal development (modENCODE data; Gerstein et al., 2010). Some of these genes encode enzymes or proteins involved in lipid synthesis (e.g. *fat-5*), storage (e.g. *dgat-2*) and mobilization (e.g. *lips-17*) (Figure 12A and B) but many others (e.g. *hacd-1*, *acs-7*) encode enzymes that participate in beta-oxidation—the process by which fatty acids are broken down in mitochondria and peroxisomes (Figure 12C). Therefore, the analysis of our LIN-29 target search results suggests that the role of LIN-29 in the regulation of intestinal genes, particularly those involved in lipid metabolism, is more substantial than originally thought.

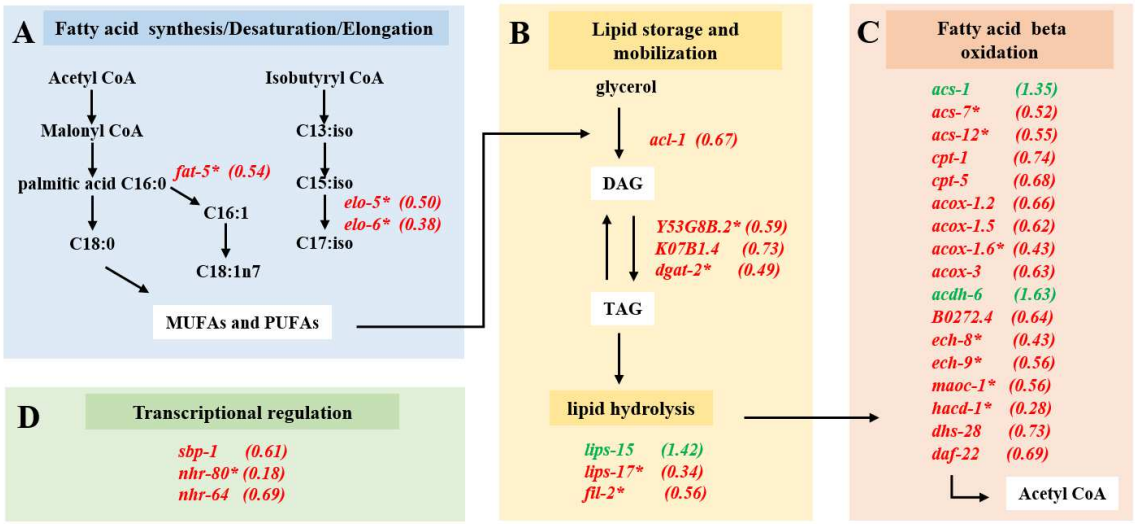


Figure 12. LIN-29-regulated genes involved in lipid metabolism. Genes that were upregulated (green font) or downregulated (red font) upon misexpression of LIN-29 are shown,

with fold change in parenthesis. Genes are grouped into broad categories (boxes A-D) based on their gene product function in lipid metabolism. Note that there are more enzymes involved in these processes, however only genes with a significant change (P<0.05) upon LIN-29 overexpression are shown here. * LIN-29 target genes that were regulated 1.7-fold or more.

In order to determine the consequence of reduction of expression for these genes, we examined published literature and databases for known roles in lipid metabolism or phenotypes associated with that process. A recent RNAi screen identified 78 genes that cause changes in lipid droplet size (and therefore fat storage) when their function was reduced (Zhu et al., 2018). Strikingly, ten of these 78 genes were also regulated by LIN-29 in our analysis (Figure 13). Only one of these genes (*W03D8.8*) was upregulated in *hs::lin-29* animals, the rest were downregulated. Four of these LIN-29-downregulated genes (*acox-1*, *maoc-1*, *dhs-28* and *daf-22*) participate in beta-oxidation whereas five others encode peroxisome assembly factors: three inside the 1.7-down-fold cutoff (*prx-1*, *prx-5* and *prx-11*) and two slightly below (*prx-12* and *prx-13*). Reduction of function by RNAi for all ten genes caused an increase in lipid droplet size, ie., increased fat content (Zhu et al., 2018). Since most of these genes were downregulated upon LIN-29 overexpression, this would suggest a similar increase in lipid droplet size should be taking place upon induction of *lin-29* expression.

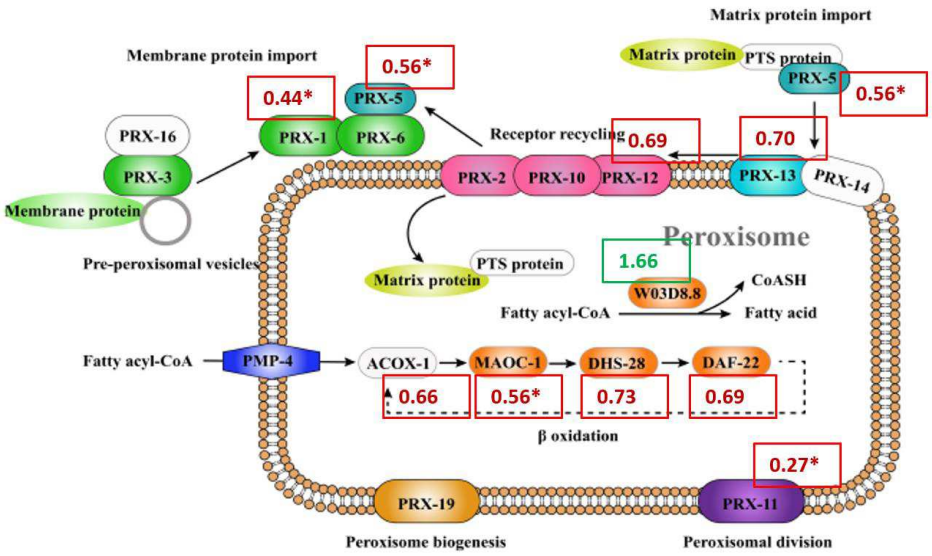


Figure 13. LIN-29-regulated genes involved in peroxisome function. The figure shows those peroxisomal proteins that were identified in a genetic screen for genes that alter lipid droplet size or accumulation when their function is reduced. Those gene products that showed a significant change (P<0.05) in gene expression upon induction of LIN-29 in the L3 stage are shown with a box; the number inside the box indicates the fold change in expression relative to control. Green indicates upregulation; red indicates downregulation. * indicates genes regulation of 1.7X or more. Adapted from Zhu et al. (2018).

In addition to this lipid droplet screen, over the years several laboratories have conducted genome-wide screens to identify genes that alter fat content upon RNAi treatment (e.g. Ashrafi et al., 2003) and this data is available in WormBase (Lee et al., 2018). For eight of the LIN-29 downregulated lipid metabolic genes (*lbp-5, ech-8, ech-9, maoc-1, dhs-16, daf-22, fil-2,* and *acox-1.2*) the RNAi phenotypes include 'fat content increased' (Supplemental Table). However, the opposite phenotype is observed for our downregulated targets *cyp-35A2* and *lagr-1*, where fat content is reduced when these genes are silenced.

Besides metabolic enzymes, among the downregulated targets we identified three genes (*sbp-1*, *nhr-80* and *nhr-64*) encoding transcription factors known to play a role in fat metabolism (Figure 12D). Like its mammalian homolog, SREBP-1c, *C. elegans* SBP-1 is a major regulator of lipid synthesis and storage; it is a transcription factor found in the intestine and it is required for wild type levels of lipogenic enzymes (including *fat-5*, *elo-5* and *elo-6*) so reduction of *sbp-1* function leads to reduced fat content (Kniazeva, Crawford, Seiber, Wang, & Han, 2004; MacNeil et al., 2015; McKay, McKay, Avery, & Graff, 2003; Nomura, Horikawa, Shimamura, Hashimoto, & Sakamoto, 2010). *nhr-80* encodes a nuclear hormone receptor that is another transcriptional regulator of *fat-5;* in addition, NHR-80 physically interacts with NHR-49, another nuclear hormone receptor that regulates lipid metabolic gene expression (Brock et al., 2006; Pathare, Lin, Bornfeldt, Taubert, & Van Gilst, 2012b; Van Gilst et al., 2005). Both *nhr-80* and *fat-5* were found as downregulated targets of LIN-29. One other nuclear hormone receptor gene that was downregulated by LIN-29, *nhr-64*, also has ties to fatty acid metabolism: reduction of *nhr-64* function suppresses the decreased fate storage phenotype of *sbp-1* mutants, and leads to an increase in fat storage (B. Liang, Ferguson, Kadyk, & Watts, 2010).

In summary, we believe our *lin-29* gain-of-function overexpression analysis has identified a previously unappreciated role in the regulation of lipid metabolism for the normal increase in LIN-29 levels that occur at the transition to adulthood. While most of the lipid metabolic gene targets we identified are downregulated by LIN-29, the known lipid storage phenotypes for the genes with available data go in both directions; some genes show an

increase in fat content upon RNAi, while others show a decrease (Supplemental Table). Considering this data alone leaves the role for LIN-29 unclear. However, when examining the general nature of the gene products that are potentially downregulated by LIN-29, and their roles in lipid metabolism (Figures 12 and 13), we suggest that there should be decreased beta-oxidation and peroxisomal function, both of which should result in an increase in fat stores. The idea that LIN-29 promotes an increase in lipid storage is consistent with the observation that *lin-29*(*n333*) mutant animals had slightly decreased fat content as adults when assayed using Oil Red O staining (Dowen et al., 2016; see Chapter 1 section V.e.).

The concept of LIN-29 as a negative regulator of fatty acid catabolic processes at the transition to adulthood is not an unreasonable one. During *C. elegans* larval life, from the hatching of the L1 larva to the transition to adulthood, animals invest the available nutrients in rapid body growth and development. During this relatively short period, the volume of the worm's body undergoes a greater than 30-fold increase; in contrast, once becoming a reproductive adult, the body grows to a much lower extent (~ 6-fold increase) over a longer time (Knight, Patel, Azevedo, & Leroi, 2002). The growth rate of the worm increases during larval life with the fastest growth in the L4 stage (Figure 14; Knight et al., 2002; Uppaluri & Brangwynne, 2015).

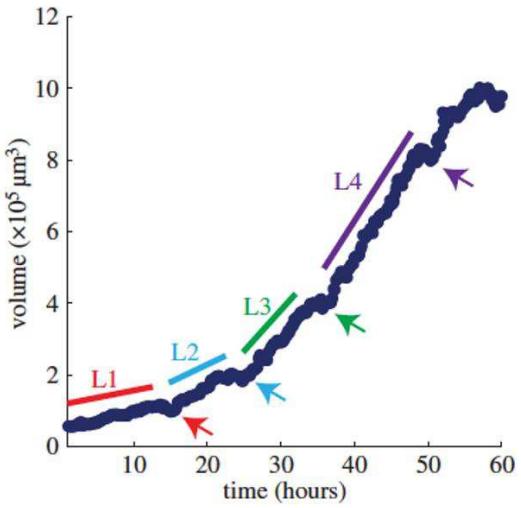


Figure 14. Growth rate during larval life. Shown is the increase in body volume over time. Note that the body grows at the highest rate during the L4 stage. Figure from Uppaluri & Brangwynne (2015)

Consistent with this growth pattern, metabolic rate assays showed an increasing trend in the early larval life reaching the highest values between the L2 and the L4 stages, followed by a steep decline upon becoming an adult (Houthoofd et al., 2002; Figure 15). From that point on, metabolic rates showed a gradual decrease towards aging. After the worm reaches the adult size, the available energy is stored as lipid droplets in the intestine, from where is later mobilized to the germline by vitellogenins (Lemieux & Ashrafi, 2015). This metabolic switch takes place at the time when most of LIN-29 protein is active, and part of LIN-29's 'maturing instruction' could be to contribute to such a change in energy management (including vitellogenesis).

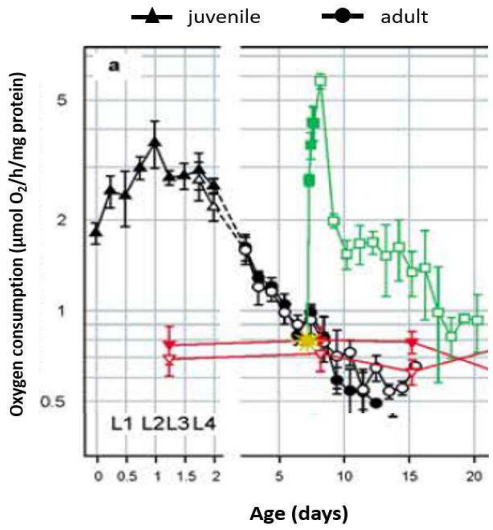


Figure 15. *C. elegans* metabolic expenditure over time. Metabolic rates (black line) were analyzed as oxygen consumption per hour normalized to protein content across larval (L1-L4) stages and during adulthood. Ignore green and red lines. Taken from Houthoofd et al. (2002).

If our hypothesis is correct and LIN-29 is in fact a negative regulator of fatty acid catabolism, perhaps periodic overexpression of this transcription factor throughout the adult life could explain the shorter lifespan we observed without any behavioral change nor visible sickness. Two of three intestinal metabolic genes we tested were indeed downregulated after *lin-29* induction in gravid adults, suggesting that at least some metabolic regulation was still possible past the normal LIN-29-repression at the L4/adult transition. While our work cannot prove that catabolic inhibition is responsible for affecting lifespan, our results are consistent with the idea that these short-lived worms may have suffered perturbance of lipid metabolic homeostasis at least to some degree.

Interestingly, expression of the LIN-29 mammalian homolog Egr-1 in white adipose tissue has been associated with diet-induced obesity, insulin resistance and hyperlipidemia in

both, mice and humans (J. Zhang et al., 2013). Egr-1 null mice, however, were shown to be protected against such pathologies and to have increased metabolic rates and energy expenditure (J. Zhang et al., 2013). Therefore, it is possible that a role for LIN-29 in the inhibition of energy consumption is conserved—at least many of its metabolic targets and their associations (pathways) are (Figure 11). If this is true, we could learn more about vertebrate lipid metabolic regulation from further study of *lin-29* in *C. elegans*.

In conclusion, our results support and expand a previously proposed idea that LIN-29 is responsible for changes in the intestine that are characteristic of the transition into adulthood (Dowen et al., 2016). Specifically, we identified a large set of specific genes functioning in lipid metabolism that may be downregulated by LIN-29 during the transition to adulthood, and we identified two genes encoding signals that may mediate LIN-29's regulation of gene expression in the intestine from the hypodermis. While more studies are required to further elucidate the LIN-29 downstream regulatory network and its implications for the exit from the juvenile stage, this study has set the ground work for the investigation of potentially conserved mechanisms regulating metabolic changes, perhaps in response to development and puberty.

Acknowledgements

We thank A. Golden for plasmids. We thank members of the Eisenmann lab for advice and support. Some strains were provided by the Caenorhabditis Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Funding for this work was provided by the University of Maryland Baltimore County.

References

Abbott, A. L., Alvarez-Saavedra, E., Miska, E. A., Lau, N. C., Bartel, D. P., Horvitz, H. R., & Ambros, V. (2005). The let-7 MicroRNA Family Members mir-48, mir-84, and mir-241 Function Together to Regulate Developmental Timing in Caenorhabditis elegans. *Developmental Cell*, *9*(3), 403–414. https://doi.org/10.1016/j.devcel.2005.07.009

Abete-Luzi, P., & Eisenmann, D. M. (2018, April 23). Regulation of C. elegans L4 cuticle collagen genes by the heterochronic protein LIN-29. *Genesis*. Wiley-Blackwell. https://doi.org/10.1002/dvg.23106

Abrahante, J. E., Daul, A. L., Li, M., Volk, M. L., Tennessen, J. M., Miller, E. A., & Rougvie, A. E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. *Developmental Cell*,

4(5), 625–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12737799

Abrahante, J. E., Miller, E. A., & Rougvie, A. E. (1998). Identification of heterochronic mutants in Caenorhabditis elegans: Temporal misexpression of a collagen :: green fluorescent protein fusion gene. *Genetics*, *149*(3), 1335–1351.

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., & Großhans, H. (2017). LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms. *Molecular Cell*, *65*(3), 476–489.e4. https://doi.org/10.1016/j.molcel.2016.12.010

Altun, Z. F., & Hall, D. H. (2009). Introduction to C. elegans anatomy. In *WormAtlas.*

Ambros, V., & Horvitz, H. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. *Science*, *226*(4673), 409–416. https://doi.org/10.1126/science.6494891

Ambros, V., & Horvitz, H. R. (1987). The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. *Genes & Development*, *1*(4), 398–414. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3678829

Angeles-Albores, D., N. Lee, R. Y., Chan, J., & Sternberg, P. W. (2016). Tissue enrichment analysis for C. elegans genomics. *BMC Bioinformatics*, *17*(1), 366. https://doi.org/10.1186/s12859-016-1229-9

Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J., & Ruvkun, G. (2003). Genome-wide RNAi analysis of Caenorhabditis elegans fat

regulatory genes. *Nature*, *421*(6920), 268–272. https://doi.org/10.1038/nature01279

Aspöck, G., Kagoshima, H., Niklaus, G., & Bürglin, T. R. (1999). Caenorhabditis elegans has scores of hedgehog-related genes: sequence and expression analysis. *Genome Research*, *9*(10), 909–23. https://doi.org/10.1101/GR.9.10.909

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., … Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, *37*(Web Server issue), W202-8. https://doi.org/10.1093/nar/gkp335

Bettinger, J. C., Euling, S., & Rougvie, A. E. (1997). The terminal differentiation factor LIN-29 is required for proper vulvar morphogenesis and egg laying in Caenorhabditis elegans. *Development*, *124*(21), 4333–4342.

Bettinger, J. C., Lee, K., & Rougvie, a E. (1996). Stage-specific accumulation of the terminal differentiation factor LIN-29 during Caenorhabditis elegans development. *Development*, *122*(8), 2517–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8756296

Boehm, M., & Slack, F. J. (2005). A Developmental Timing MicroRNA and Its Target Regulate Life Span in C. elegans. *Science*, *310*(5756), 1954–1957. https://doi.org/10.1126/science.1115596

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, *77*(1), 71–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4366476

Brock, T. J., Browse, J., & Watts, J. L. (2006). Genetic regulation of unsaturated fatty acid composition in C. elegans. *PLoS Genetics*, *2*(7), 0997-1005.

https://doi.org/10.1371/journal.pgen.0020108

Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., Lewis, S., … Web Presence Working Group, the W. P. W. (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics (Oxford, England)*, *25*(2), 288–9. https://doi.org/10.1093/bioinformatics/btn615

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A Sequence Logo Generator. *Genome Research*, *14*(6), 1188–1190. https://doi.org/10.1101/gr.849004

Ding, X. C., & Grosshans, H. (2009). Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. *The EMBO Journal*, *28*(3), 213–22. https://doi.org/10.1038/emboj.2008.275

Dowen, R. H., Breen, P. C., Tullius, T., Conery, A. L., & Ruvkun, G. (2016). A microRNA program in the C. elegans hypodermis couples to intestinal mTORC2/PQM-1 signaling to modulate fat transport. *Genes and Development*, *30*(13), 1515–1528. https://doi.org/10.1101/gad.283895.116

Fielenbach, N., Guardavaccaro, D., Neubert, K., Chan, T., Li, D., Feng, Q., … Antebi, A. (2007). DRE-1: An Evolutionarily Conserved F Box Protein that Regulates C. elegans Developmental Age. *Developmental Cell*, *12*(3), 443–455. https://doi.org/10.1016/J.DEVCEL.2007.01.018

Friedlander-shani, L., & Podbilewicz, B. (2011). Cell Fusion in Health and Disease, *950*, 5–11. https://doi.org/10.1007/978-94-007-0782-5

Frøkjær-Jensen, C. (2015). Transposon-Assisted Genetic Engineering with Mos1-Mediated Single-Copy Insertion (MosSCI) (pp. 49–58). Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-4939-2842-2_5

Frøkjær-Jensen, C., Davis, M. W., Ailion, M., & Jorgensen, E. M. (2012). Improved Mos1-mediated transgenesis in C. elegans. *Nature Methods*, *9*(2), 117–118. https://doi.org/10.1038/nmeth.1865

Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., … Trifonov, T. (2010). Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. *Science*, *330*(6012), 1775 LP-1787. https://doi.org/10.2436/20.7010.01.53

Gorrepati, L., Krause, M. W., Chen, W., Brodigan, T. M., Correa-Mendez, M., & Eisenmann, D. M. (2015). Identification of Wnt Pathway Target Genes Regulating the Division and Differentiation of Larval Seam Cells and Vulval Precursor Cells in *Caenorhabditis elegans*. *G3: Genes|Genomes|Genetics*, *5*(8), 1551–1566. https://doi.org/10.1534/g3.115.017715

Gumienny, T. L., & Savage-Dunn, C. (2013). TGF-β signaling in C. elegans. *WormBook : The Online Review of C. Elegans Biology*, 1–34. https://doi.org/10.1895/wormbook.1.22.2

Gupta, B. P., Hanna-Rose, W., & Sternberg, P. W. (2012). Morphogenesis of the vulva and the vulval-uterine connection*. *WormBook The C. Elegans Research Community*, *11521*(10). https://doi.org/10.1895/wormbook.1.152.1

Han, S. K., Lee, D., Lee, H., Kim, D., Son, H. G., Yang, J.-S., … Kim, S. (2016a).

OASIS 2: online application for survival analysis 2 with features for the analysis of maximal lifespan and healthspan in aging research. *Oncotarget*, *7*(35), 56147–56152. https://doi.org/10.18632/oncotarget.11269

Han, S. K., Lee, D., Lee, H., Kim, D., Son, H. G., Yang, J.-S., … Kim, S. (2016b). OASIS 2: online application for survival analysis 2 with features for the analysis of maximal lifespan and healthspan in aging research. *Oncotarget*, *7*(35), 56147–56152. https://doi.org/10.18632/oncotarget.11269

Harris, D. T., & Horvitz, H. R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. *Development*, *138*(18), 4051–4062. https://doi.org/10.1242/dev.065417

Hong, Y., Roy, R., & Ambros, V. (1998). Developmental regulation of a cyclin-dependent kinase inhibitor controls postembryonic cell cycle progression in Caenorhabditis elegans, *3597*, 3585–3597.

Houthoofd, K., Braeckman, B. P., Lenaerts, I., Brys, K., De Vreese, A., Van Eygen, S., & Vanfleteren, J. R. (2002). Ageing is reversed, and metabolism is reset to young levels in recovering dauer larvae of C. elegans. *Experimental Gerontology*, *37*(8–9), 1015–1021. https://doi.org/10.1016/S0531-5565(02)00063-3

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

Huang, T. F., Cho, C. Y., Cheng, Y. T., Huang, J. W., Wu, Y. Z., Yeh, A. Y. C., … Wu, Y. C. (2014). BLMP-1/Blimp-1 Regulates the Spatiotemporal Cell Migration Pattern

in C. elegans. *PLoS Genetics*, *10*(6). https://doi.org/10.1371/journal.pgen.1004428

Hunter, S. E., Finnegan, E. F., Zisoulis, D. G., Lovci, M. T., Melnik-Martinez, K. V., Yeo, G. W., & Pasquinelli, A. E. (2013). Functional Genomic Analysis of the let-7 Regulatory Network in Caenorhabditis elegans. *PLoS Genetics*, *9*(3). https://doi.org/10.1371/journal.pgen.1003353

Hutter, H., & Suh, J. (2016). GExplore 1.4: An expanded web interface for queries on *Caenorhabditis elegans* protein and gene function. *Worm*, *5*(4), e1234659. https://doi.org/10.1080/21624054.2016.1234659

Inoue, T., Wang, M., Ririe, T. O., Fernandes, J. S., & Sternberg, P. W. (2005). Transcriptional network underlying Caenorhabditis elegans vulval development. *Proc Natl Acad Sci U S A*, *102*(14), 4972–4977. https://doi.org/0408122102 [pii]\r10.1073/pnas.0408122102

Jackson, B. M., Abete-Luzi, P., Krause, M. W., & Eisenmann, D. M. (2014). Use of an Activated Beta-Catenin to Identify Wnt Pathway Target Genes in Caenorhabditis elegans , Including a Subset of Collagen Genes Expressed in Late Larval Development. *G3: Genes|Genomes|Genetics*, *4*(4), 733–747. https://doi.org/10.1534/g3.113.009522

Kamath, R. S., & Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. *Methods (San Diego, Calif.)*, *30*(4), 313–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12828945

Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., & Ahringer, J. (2000). Effectiveness of specific RNA-mediated interference through ingested double-

stranded RNA in Caenorhabditis elegans. *Genome Biology*, *2*(1), research0002.1. https://doi.org/10.1186/gb-2000-2-1-research0002

Kniazeva, M., Crawford, Q. T., Seiber, M., Wang, C.-Y., & Han, M. (2004). Monomethyl Branched-Chain Fatty Acids Play an Essential Role in Caenorhabditis elegans Development. *PLoS Biology*, *2*(9), e257. https://doi.org/10.1371/journal.pbio.0020257

Knight, C. G., Patel, M. N., Azevedo, R. B. R., & Leroi, A. M. (2002). A novel mode of ecdysozoan growth in Caenorhabditis elegans. *Evolution and Development*, *4*(1), 16–27. https://doi.org/10.1046/j.1525-142x.2002.01058.x

Kolundzic, E., Ofenbauer, A., Bulut, S. I., Uyar, B., Baytek, G., Sommermeier, A., … Tursun, B. (2018). FACT Sets a Barrier for Cell Fate Reprogramming in Caenorhabditis elegans and Human Cells. *Developmental Cell*. https://doi.org/10.1016/J.DEVCEL.2018.07.006

Kumbrink, J., Gerlinger, M., & Johnson, J. P. (2005). Egr-1 induces the expression of its corepressor Nab2 by activation of the Nab2 promoter thereby establishing a negative feedback loop. *Journal of Biological Chemistry*, *280*(52), 42785–42793. https://doi.org/10.1074/jbc.M511079200

Kumbrink, J., Kirsch, K. H., & Johnson, J. P. (2010). EGR1, EGR2, and EGR3 activate the expression of their coregulator NAB2 establishing a negative feedback loop in cells of neuroectodermal and epithelial origin. *Journal of Cellular Biochemistry*, *111*(1), 207–217. https://doi.org/10.1002/jcb.22690

Lee, R. Y. N., Howe, K. L., Harris, T. W., Arnaboldi, V., Cain, S., Chan, J., … Sternberg,

P. W. (2018). WormBase 2017: molting into a new stage. *Nucleic Acids Research*, *46*(D1), D869–D874. https://doi.org/10.1093/nar/gkx998

Lemieux, G. A., & Ashrafi, K. (2015). Insights and challenges in using *C. elegans* for investigation of fat metabolism. *Critical Reviews in Biochemistry and Molecular Biology*, *50*(1), 69–84. https://doi.org/10.3109/10409238.2014.959890

Liang, B., Ferguson, K., Kadyk, L., & Watts, J. L. (2010). The Role of Nuclear Receptor NHR-64 in Fat Storage Regulation in Caenorhabditis elegans. *PLoS ONE*, *5*(3), e9869. https://doi.org/10.1371/journal.pone.0009869

Liang, J., Yu, L., Yin, J., & Savage-Dunn, C. (2007). Transcriptional repressor and activator activities of SMA-9 contribute differentially to BMP-related signaling outputs. *Developmental Biology*, *305*(2), 714–725. https://doi.org/10.1016/J.YDBIO.2007.02.038

Lin, S.-Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., … Slack, F. J. (2003a). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Developmental Cell*, *4*(5), 639–50. https://doi.org/10.1016/S1534-5807(03)00124-2

Lin, S.-Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., … Slack, F. J. (2003b). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Developmental Cell*, *4*(5), 639–50. https://doi.org/10.1016/S1534-5807(03)00124-2

Liu, Z., Kirch, S., & Ambros, V. (1995). The Caenorhabditis elegans heterochronic gene pathway controls stage-specific transcription of collagen genes. *Development*

(Cambridge, England), *121*(8), 2471–2478.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

MacNeil, L. T., Pons, C., Arda, H. E., Giese, G. E., Myers, C. L., & Walhout, A. J. M. (2015). Transcription Factor Activity Mapping of a Tissue-Specific in Vivo Gene Regulatory Network. *Cell Systems*, *1*(2), 152–162. https://doi.org/10.1016/j.cels.2015.08.003

Mak, H. Y. (2012). Lipid droplets as fat storage organelles in Caenorhabditis elegans: Thematic Review Series: Lipid Droplet Synthesis and Metabolism: from Yeast to Man. *Journal of Lipid Research*, *53*(1), 28–33. https://doi.org/10.1194/jlr.R021006

McKay, R. M., McKay, J. P., Avery, L., & Graff, J. M. (2003). C elegans: a model for exploring the genetics of fat storage. *Developmental Cell*, *4*(1), 131–42. https://doi.org/10.1016/S1534-5807(02)00411-2

Mello, C., & Fire, A. (1995). DNA transformation. *Methods in Cell Biology*, *48*, 451–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8531738

Moss, E. G. (2007). Heterochronic Genes and the Nature of Developmental Time. *Current Biology*, *17*(11), 425–434. https://doi.org/10.1016/j.cub.2007.03.043

Moss, E. G., & Romer-Seibert, J. (2014). Cell-intrinsic timing in animal development. *Wiley Interdisciplinary Reviews: Developmental Biology*, *3*(5), 365–377. https://doi.org/10.1002/wdev.145

Munoz-Torres, M., & Carbon, S. (2017). Get GO! Retrieving GO Data Using AmiGO, QuickGO, API, Files, and Tools (pp. 149–160). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-3743-1_11

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., … Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, *424*(6946), 277–284. https://doi.org/10.1038/nature01789

Narasimhan, K., Lambert, S. A., Yang, A. W. H., Riddell, J., Mnaimneh, S., Zheng, H., … Hughes, T. R. (2015). Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. *ELife*, *2015*(4), 1–53. https://doi.org/10.7554/eLife.06967

Newman, A. P., Inoue, T., Wang, M., & Sternberg, P. W. (2000). The Caenorhabditis elegans heterochronic gene lin-29 coordinates the vulval-uterine-epidermal connections. *Current Biology*, *10*(23), 1479–1488. https://doi.org/10.1016/S0960-9822(00)00827-7

Nimmo, R. A., & Slack, F. J. (2009). An elegant miRror: microRNAs in stem cells, developmental timing and cancer. *Chromosoma*. https://doi.org/10.1007/s00412-009-0210-z

Nomura, T., Horikawa, M., Shimamura, S., Hashimoto, T., & Sakamoto, K. (2010). Fat accumulation in Caenorhabditis elegans is mediated by SREBP homolog SBP-1. *Genes and Nutrition*, *5*(1), 17–27. https://doi.org/10.1007/s12263-009-0157-y

Nyström, J., Shen, Z.-Z., Aili, M., Flemming, A. J., Leroi, A., & Tuck, S. (2002).

Increased or decreased levels of Caenorhabditis elegans lon-3, a gene encoding a collagen, cause reciprocal changes in body length. *Genetics*, *161*(1), 83–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12019225

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., … Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, *408*, 86. Retrieved from http://dx.doi.org/10.1038/35040556

Patel, T., Tursun, B., Rahe, D. P., & Hobert, O. (2012). Removal of Polycomb Repressive Complex 2 Makes C. elegans Germ Cells Susceptible to Direct Conversion into Specific Somatic Cell Types. *Cell Reports*, *2*(5), 1178–1186. https://doi.org/10.1016/j.celrep.2012.09.020

Pathare, P. P., Lin, A., Bornfeldt, K. E., Taubert, S., & Van Gilst, M. R. (2012a). Coordinate regulation of lipid metabolism by novel nuclear receptor partnerships. *PLoS Genetics*, *8*(4), e1002645. https://doi.org/10.1371/journal.pgen.1002645

Pathare, P. P., Lin, A., Bornfeldt, K. E., Taubert, S., & Van Gilst, M. R. (2012b). Coordinate Regulation of Lipid Metabolism by Novel Nuclear Receptor Partnerships. *PLoS Genetics*, *8*(4), e1002645. https://doi.org/10.1371/journal.pgen.1002645

Rausch, M., Ecsedi, M., Bartake, H., Müllner, A., & Großhans, H. (2015). A genetic interactome of the let-7 microRNA in C. elegans. *Developmental Biology*, *401*(2), 276–286. https://doi.org/10.1016/J.YDBIO.2015.02.013

Ririe, T. O., Fernandes, J. S., & Sternberg, P. W. (2008). The Caenorhabditis elegans

vulva: a post-embryonic gene regulatory network controlling organogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20095–9. https://doi.org/10.1073/pnas.0806377105

Roberts, A. F., Gumienny, T. L., Gleason, R. J., Wang, H., & Padgett, R. W. (2010). Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in Caenorhabditis elegans. *BMC Developmental Biology*, *10*, 61. https://doi.org/10.1186/1471-213X-10-61

Rougvie, A. E., & Ambros, V. (1995). The Heterochronic Gene Lin-29 Encodes a Zinc-Finger Protein that Controls a Terminal Differentiation Event in Caenorhabditis-Elegans. *Development*, *121*(8), 2491–2500. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7671813

Rougvie, A. E., & Moss, E. G. (2013). Developmental Transitions in C. elegans Larval Stages (pp. 153–180). https://doi.org/10.1016/B978-0-12-396968-2.00006-3

Schindler, A. J., Baugh, L. R., & Sherwood, D. R. (2014). Identification of Late Larval Stage Developmental Checkpoints in Caenorhabditis elegans Regulated by Insulin/IGF and Steroid Hormone Signaling Pathways. *PLoS Genetics*, *10*(6), 13–16. https://doi.org/10.1371/journal.pgen.1004426

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R., & Ruvkun, G. (2000). The

lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Molecular Cell*, *5*(4), 659–669. https://doi.org/10.1016/S1097-2765(00)80245-2

Suzuki, Y., Morris, G. A., Han, M., & Wood, W. B. (2002). A cuticle collagen encoded by the lon-3 gene may be a target of TGF-β signaling in determining Caenorhabditis elegans body shape. *Genetics*, *162*(4), 1631–1639. Retrieved from http://www.genetics.org/content/genetics/162/4/1631.full.pdf

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., … von Mering, C. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, *43*(D1), D447–D452. https://doi.org/10.1093/nar/gku1003

Tocchini, C., Keusch, J. J., Miller, S. B., Finger, S., Gut, H., Stadler, M. B., & Ciosk, R. (2014). The TRIM-NHL Protein LIN-41 Controls the Onset of Developmental Plasticity in Caenorhabditis elegans. *PLoS Genetics*, *10*(8), e1004533. https://doi.org/10.1371/journal.pgen.1004533

Tursun, B., Patel, T., Kratsios, P., & Hobert, O. (2011). Direct Conversion of C. elegans Germ Cells into Specific Neuron Types. *Science*, *331*(6015), 304–308. https://doi.org/10.1126/science.1199082

Uppaluri, S., & Brangwynne, C. P. (2015). A size threshold governs Caenorhabditis elegans developmental progression. *Proceedings. Biological Sciences*, *282*(1813), 20151283. https://doi.org/10.1098/rspb.2015.1283

Van Gilst, M. R., Hadjivassiliou, H., Jolly, A., & Yamamoto, K. R. (2005). Nuclear

Hormone Receptor NHR-49 Controls Fat Consumption and Fatty Acid Composition in C. elegans. *PLoS Biology*, *3*(2), e53. https://doi.org/10.1371/journal.pbio.0030053

Wildwater, M., Sander, N., de Vreede, G., & van den Heuvel, S. (2011). Cell shape and Wnt signaling redundantly control the division axis of C. elegans epithelial stem cells. *Development*, *138*(20), 4375–4385. https://doi.org/10.1242/dev.066431

Yi, W., & Zarkower, D. (1999). Similarity of DNA binding and transcriptional regulation by Caenorhabditis elegans MAB-3 and Drosophila melanogaster DSX suggests conservation of sex determining mechanisms. *Development (Cambridge, England)*, *126*(5), 873–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9927589

Zhang, J., Zhang, Y., Sun, T., Guo, F., Huang, S., Chandalia, M., … Fu, M. (2013). Dietary obesity-induced Egr-1 in adipocytes facilitates energy storage via suppression of FOXC2. *Scientific Reports*, *3*, 1–10. https://doi.org/10.1038/srep01476

Zhang, S. O., Box, A. C., Xu, N., Le Men, J., Yu, J., Guo, F., … Mak, H. Y. (2010). Genetic and dietary regulation of lipid droplet expansion in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(10), 4640–5. https://doi.org/10.1073/pnas.0912308107

Zhang, Y., Zou, X., Ding, Y., Wang, H., Wu, X., & Liang, B. (2013). Comparative genomics and functional study of lipid metabolic genes in Caenorhabditis elegans. *BMC Genomics*, *14*(1), 164. https://doi.org/10.1186/1471-2164-14-164

Zhu, X., Liu, Y., Zhang, H., & Liu, P. (2018). Whole-genome RNAi screen identifies methylation-related genes influencing lipid metabolism in Caenorhabditis elegans.

Journal of Genetics and Genomics = Yi Chuan Xue Bao.
https://doi.org/10.1016/j.jgg.2018.03.005

Zugasti, O., Rajan, J., & Kuwabara, P. E. (2005). The function and expansion of the Patched- and Hedgehog-related homologs in C. elegans. *Genome Research*, *15*(10), 1402–10. https://doi.org/10.1101/gr.3935405

Chapter 4: Discussion

In this thesis, I have described my work on *C. elegans* as divided into two projects which will result in two publications. First, the study of temporal regulation of cuticle collagen (*col*) gene expression, specifically a subset of *col* genes that show peak expression in the L4 stage (Chapter 2; Abete-Luzi & Eisenmann, 2018) and second, a search for LIN-29 target genes and their implications for the larva-to-adult transition (Chapter 3; manuscript in preparation). Here, I will discuss my contributions, unanswered questions and future directions for each topic.

I. Investigating temporal control of gene expression in *C. elegans*

I.a. Cuticle collagens as a model to study temporal regulation of gene expression

One of the goals of the Eisenmann Lab is to use *C. elegans* as a model organism to understand the temporal component of gene expression regulation in the context of animal development. Although we know about the general role of the heterochronic pathway in regulating developmental timing in the worm, how this pathway actually regulates gene expression of downstream targets at specific times and places remains mostly unknown. Previous work from this lab identified a large set of genes from the cuticle collagen gene family that show stage-specific expression that might allow us to address this gap in our knowledge. In that work (Jackson, Abete-Luzi, Krause, & Eisenmann, 2014), 116 of 187 *col* genes were categorized based on the life stage in which their transcript levels peak by

comparing developmental RNA-seq data available from modENCODE (Gerstein et al., 2010; Jackson, Abete-Luzi, Krause, & Eisenmann, 2014). Sets of specific *col* genes were identified that show a single peak of expression in the embryo, L2, L3, L4 or adult stage (Chapter 1, Figure 4). This set of stage-specific genes potentially provides a powerful and convenient system for identifying key regulators that control the temporal aspect of gene expression—at least for this gene family in the hypodermis. These genes became the basis for what would become an active goal of the lab: to find stage-specific regulators of *col* gene expression for all stages of the worm.

Broadly, there are at least two complementary ways by which stage-specific subsets of *cols* can be used to identify key temporal regulators: 1) following a bottom-up approach, in which regulatory regions are narrowed down, searched for common sequences among co-expressing *col* promoters (or among other *Caenorhibditis* spp.), tested for functionality (are these regions necessary and/or sufficient for temporally-specific gene expression), and used to identify transcription factors that bind them; and 2) using a top-down approach, where stage-specific *col* gene expression can be assessed for requirements of candidate transcription factors, which can also be tested for binding. With the help of UMBC undergraduate Daniel Stonko, I started by building YFP transcriptional reporter strains for embryonic, L2- and L4-stage *col* genes to verify that predicted temporal expression patterns were accurate—and succeeded. I then initiated this project by using both approaches outlined above to investigate a subset of *col* genes that peak in the L4 stage.

I.b. L4-specific *col* gene regulation and my contributions

In this work, I identified minimal promoter regions for *col-38*, *col-49*, and *col-63* as DNA fragments of less than 300 bp, a length that is surprisingly short yet sufficient to drive expression of YFP reporters specifically in the mid and late L4 stages. I then predicted *cis*-regulatory elements and evaluated their regulatory function *in vivo* via mutagenesis of a *col-38p::yfp* reporter. I used RNAi to study the requirements of four candidate transcription regulators, GATA factor ELT-1 and ELT-3, zinc finger factor LIN-29, and the LIN-29 co-factor MAB-10, and found only LIN-29 to be necessary for the expression of four L4-specific genes (*col-38*, *col-49*, *col-63* and *col-138)*. To further investigate this regulation, I made the first inducible *lin-29* expression construct which, like all other constructs used in this work, I integrated in single copy into the *C. elegans* genome. Using this strain, I found that temporal misexpression of LIN-29 is also sufficient to activate these L4-specific genes at other developmental stages, a fact that reflects the temporal aspect of the control that this transcription factor has over its targets. I also showed that the LIN-29 DNA-binding domain binds the *col-38*, *col-49*, and *col-63* minimal promoters *in vitro* and it does so most likely via more than one site. Moreover, for *col-38*, I showed that the LIN-29 sites necessary for reporter expression *in vivo* are also bound *in vitro*.

All in all, before my research on L4 *col* genes was done, LIN-29 was known to regulate only two cuticle collagens (in addition to non-*col* targets) and its binding sequence was predicted from protein binding arrays (Narasimhan et al., 2015). My results, however, 1) suggest a broader role for LIN-29 in the regulation of L4/adult specific cuticle secretion by

targeting multiple *col* genes; and 2) provide the first identification of specific binding sites for LIN-29 in a known target gene that are bound by LIN-29 *in vitro* and necessary for target gene expression *in vivo*.

I.c. Unanswered questions and future directions

The results obtained in this work indicate that, out of four candidate transcription factors, only LIN-29 is required for L4 *col* expression. This was surprising given that GATA factors ELT-1 and ELT-3 were previously shown to regulate a few *col* genes (Budovskaya et al., 2008; Gilleard, Barry, & Johnstone, 1997; Gilleard & McGhee, 2001; Yin, Madaan, Park, Aftab, & Savage-Dunn, 2015) and that the minimal promoters of the L4 *cols* I tested contain GATA sites. The site requirements (mutagenesis) analysis performed on the *col-38* reporter was somewhat inconclusive in terms of GATA regulation where no effect was seen when losing two of the three predicted GATA sites (G1 and G2), yet a drastic requirement for a single site (G3) was observed. Curiously, in no case was there any form of intermediate expression. As long as G3 remained wild type, no change in expressivity or penetrance was observed for *col-38* reporter, regardless of additional GATA site mutagenesis and/or stringent GATA factor RNAi treatment. It is possible that the G3 site overlaps the binding site of an additional required factor, which would explain the difference between our mutagenesis and RNAi results.

The fact that misexpression of LIN-29 from the heat shock promoter was sufficient to regulate L4 *col* expression in the L2/L3 molt and in the adult but not in the embryo or the L1 stage, also made me consider the possibility of an additional unknown regulator which might be present in the hypodermis during mid to late larval and adult life but be absent at earlier stages. Alternatively, another way to explain this temporal restriction in LIN-29's ability to act, is the *col* gene chromatin context: perhaps it is less permissive for expression in the embryo and the L1 stage. The identification of a potential additional regulator of L4 *col* gene expression could address these questions. Would the unknown factor's binding sequence overlap with GATA site G3? Would it be present in the hypodermis only after the L2 stage? A yeast one-hybrid assay would be one way to identify such a regulator, should there be one. If a new regulator is identified, finding its binding sites in L4-specific *col* genes by doing EMSAs or homologue predictions, and the investigation of its temporal expression could explain my results.

Another major question that remains to be answered is whether LIN-29 regulates these genes directly or indirectly. While my results strongly suggest that it does, an *in vivo* binding experiment should be performed to confirm this. As part of the goals of the modENCODE project, efforts were made to determine genome-wide mappings of transcription factor binding sites at different stages of *C. elegans* via ChIP-chip and ChIP-Seq (Araya et al., 2014; Celniker et al., 2009). To date, there are 242 submissions corresponding to 96 transcription factors; however, none of them is LIN-29, unfortunately. One way to obtain LIN-29 *in vivo* binding data would be to express this protein fused to a tag such as FLAG or GFP, either by adding the tag to the endogenous *lin-29* via genome

editing with CRISPR/Cas9 or by expressing a *lin-29::GFP* transgene, and perform anti-tag ChIP-Seq on L4-stage animals. This method would not only answer the above question, but it would also reveal additional direct targets of LIN-29 at a genomic scale (this topic is discussed further below).

Finally, the big question left to be resolved is what factors temporally regulate the subsets of stage-specific *col* genes expressing at other stages in *C. elegans* life. Following the same approaches used to investigate L4 *col* gene regulation would most certainly allow the identification of key regulators acting at other life stages. However, while the study of temporal regulation of *col* gene expression in all or several life stages was my initial objective, the results obtained in my first attempt (L4 stage *cols*; Chapter 2) motivated me to further explore the L4-to-adult temporal regulation, and in particular, the role of the heterochronic protein LIN-29 in that process.

II. LIN-29 regulation in the larva-to-adult transition

II.a. A search for LIN-29 targets

The zinc finger transcription factor LIN-29 is the most downstream effector of the heterochronic pathway and is known for being required for a subset of developmental events to take place in coordination with the larval-to-adult switch. Some of these events are seam cell fusion and final differentiation, the cessation of the molting cycle and the coordination of the vulval-uterine-seam cell connection (Bettinger, Euling, & Rougvie, 1997; Bettinger, Lee, & Rougvie, 1996; Harris & Horvitz, 2011; Hayes & Ruvkun, 2006; Newman, Inoue, Wang, & Sternberg, 2000; Rougvie & Ambros, 1995). Whereas the role of LIN-29 in these developmental changes has been well studied, only a few target genes have been characterized and linked to such processes. In this work, I wanted to explore LIN-29-mediated regulation at a genomic level by identifying LIN-29 targets, which might also point to potential new roles for this transcription factor.

My advisor and I decided to search for LIN-29 targets using a gain-of-function approach in which overexpression of *lin-29* is followed by the assessment of differential gene expression. This type of approach to search for target genes has been successfully used in the Eisenmann Lab, and results have been published (Gorrepati et al., 2015; Jackson et al., 2014). In my case, a strain carrying inducible *lin-29* was already available from my previous project (Chapter 2; Abete-Luzi & Eisenmann, 2018); and a collaboration with the Krause Lab (NIH/NIDDK) was established to perform RNA-Seq from samples processed

after induction. The outcome of such experiments and the success of the obtained results are discussed in the following section.

II.b. Results and major contributions

To the extent of our knowledge, this work is the first one to use an inducible gain-of-function and RNA-seq approach to identify target genes of LIN-29. In this study, we identified 1,101 genes with significant differential expression, and using an arbitrary cutoff we broke the set down to 230 and 350 genes that were upregulated and downregulated, respectively. Enrichment analysis indicated the presence of 33 cuticle collagens (most of which are specific to the L4/adult stages) among activated targets, suggesting a rather preponderant role for LIN-29 in cuticle secretion at the last molt. This group of genes included all the L4 *cols* found to be regulated by LIN-29 in Chapter 2, a fact that added confidence in our data and together reinforced the idea of LIN-29 as a major regulator of adult cuticle synthesis.

Doing similar analyses, we identified many genes involved in lipid metabolism that were significantly overrepresented among LIN-29 downregulated targets. Most of these metabolic genes, many of which encode enzymes involved in fatty acid catabolism, are normally expressed in the intestine and repressed in the L4 stage. Interestingly, we also found enrichment of peroxisomal genes in the downregulated set. We reasoned that if LIN-

29 repression of these genes compromises peroxisomal function, then this would also negatively affect the beta-oxidation of fatty acids. Given the nature of these repressed targets, it is possible that LIN-29 has a negative effect over fatty acid catabolic processes, which could possibly result in the increase of fat storage.

Regulation of intestinal vitellogenin genes by LIN-29 was shown previously, however, the signal or signals propagating such regulation remained unidentified (Dowen, Breen, Tullius, Conery, & Ruvkun, 2016). Here, we found that two LIN-29-upregulated signaling molecules, the hypodermally-expressed *warthog* protein WRT-6, and the insulin-like peptide INS-37, are required for both activation of expression of intestinal vitellogenins (eg., *vit-3*), and for the repression of some intestinal metabolic enzyme genes, including those involved in fatty acid beta-oxidation (eg., *hacd-1*). I believe that such signals are at least part of the missing link between hypodermal LIN-29 upregulation and intestinal gene response.

For the past two decades, LIN-29 has been mostly known for its role in the temporal regulation of larva-to-adult transition events that are directly related to hypodermal cells, whereas its role in other tissues has been less characterized. Our work on LIN-29 has contributed to a previously unappreciated aspect of this regulator: in addition to the recently shown function for LIN-29 in the control of vitellogenin expression, our results support the idea that LIN-29 expressed in the hypodermis may also play a major role in the intestinal control of lipid metabolism. Perhaps, as a major regulator of the larva-to-adult transition,

LIN-29 also contributes to a metabolic restructuring which, along with molting control, reproductive organogenesis and fertility, may be necessary for the adult life.

II.c. Unanswered questions and future directions

An important gap in our knowledge of LIN-29 targets is knowing whether specific genes are direct or indirect targets of this transcription factor. In the course of this work and also in collaboration with the Krause Lab, we tried to investigate which targets are directly bound by LIN-29 *in vivo* by performing Dam-mediated identification of LIN-29 binding sites across the genome. This technique uses a bacterial DNA adenine methyltransferase (Dam) fused to a eukaryotic transcription factor of interest (e.g. LIN-29). Once expressed in the organism of choice, the fusion protein (if functional) is recruited to the transcription factor's binding site where it methylates GATC in the neighboring genomic region. Genomic DNA from the whole organism is treated with *Dpn*I leaving only fragments flanked by methylated GATC to be amplified and then analyzed by high throughput genomic DNA sequencing (Aughey & Southall, 2016). We generated the necessary strains expressing Dam::LIN-29 and GFP::Dam following the Schuster and Gomez-Saldivar/Askjaer protocols (Askjaer, 2014; Gómez-saldivar, Meister, & Askjaer, 2016) and proceeded with DNA sequencing and LIN-29 binding site mapping. Unfortunately, neither we nor our collaborators were confident in the results obtained, and this data will not be published nor discussed here. At this stage, the best way to answer the question of direct/indirect target genes will be by doing conventional ChIP-Seq analysis of L4-staged

animals expressing a tagged LIN-29 protein as outlined above (subsection I.c.). I would expect to find LIN-29 bound to the promoters of hypodermal genes such as those encoding cuticle collagens (e.g., *col-38, col-49*) and if our hypothesis of a cell-non-autonomous role for LIN-29 is correct, I would expect LIN-29 not to be bound to genes with expression exclusively in the intestine**.** Such results would extend our understanding of LIN-29 function in the larva-to-adult transition by identifying direct targets of LIN-29 activity, as well as validate our hypothesis of an indirect role for LIN-29 in intestinal gene regulation.

Another remaining question is whether the other two LIN-29 regulated signals we identified (*grd-11* and *grl-14*) also mediate intestinal gene regulation by this transcription factor in a similar way that *wrt-6* and *ins-37* seem to do. As was done when investigating *ins-37* and *wrt-6* in Chapter 3, requirements for the other two signals could be tested by performing RNAi treatment on each signal gene, followed by RT-qPCR analysis of intestinal target gene expression in the L4 stage. If either or both signals are playing a role, I predict we would see vitellogenin gene expression decrease and/or lipid metabolic gene expression increase (derepression) in the RNAi-treated animals. In addition, a larger number of LIN-29 intestinal target genes should be interrogated for all four signal requirements than I have currently done. Also, simultaneous deficiency of more than one signal (e.g. double RNAi treatment) could reveal redundant activity if any, in which case I would expect to find their effect on intestinal gene expression to be exacerbated. In any case, finding all of the signals responsible for LIN-29-dependent intestinal regulation is a crucial step in understanding the role of the heterochronic network in the newly-uncovered

hypodermal to intestinal interaction. I expect to carry out these experiments on additional signal genes in the coming weeks before our manuscript is submitted for publication.

Finally, the most interesting question of all is whether LIN-29 is a *bona fide* regulator of lipid metabolism in the worm at the transition to adulthood. Assuming that LIN-29 does affect lipid management, we would like to know if overexpression of LIN-29 can cause physiological changes in fat storage. Known RNAi fat content phenotypes for LIN-29 metabolic targets appear somewhat conflicting: some of them show 'fat content increased' while others show 'fat content reduced'. Given these opposite phenotypes, it is theoretically difficult to predict whether the net effect of early overexpression of LIN-29 would be an increase or a decrease of lipid content, or no change at all. One common method for testing lipid content in worms is using Oil Red O (ORO), a dye which stains total fat and can be digitally quantified (Wählby et al., 2014). Previous ORO staining done on *lin-29*(*lf*) mutant adults showed slightly reduced fat content with respect to wild type (Dowen et al., 2016; Chapter I, section V). Therefore, if this effect is reciprocal, I would expect to find an increased amount of stored lipid in animals that overexpressed *lin-29*. However, if the result is 'no change', no conclusion could be drawn since a major caveat of this technique is that its use is restricted to general quantitative measurements. A more accurate way to assess potential changes in fat metabolism in response to LIN-29 is to do Gas Chromatography-Mass Spectrometry (GC-MS) analysis of lipids. This technique not only would allow the quantification of lipid content, but it would also enable the identification of fatty acid composition thus allowing the detection of qualitative changes as well (Watts & Browse, 2002). For instance, LIN-29 may not affect total lipid content

(storage) but may alter specific fatty acid pathways in such way that the relative amount of a given fatty acid has significantly changed instead. This type of phenotype would not be observed with ORO, yet it would be detected with GC-MS. My advisor and I consider that GC-MS analysis of lipids from LIN-29-induced animals is key to confirming LIN-29's role in lipid metabolism. Therefore, I will attempt to perform such experiments in the following weeks before submission of this work for publication.

II.d. Expanding our knowledge and exploring LIN-29 translational value

One of LIN-29's proposed mammalian homologs, EGR1, has been recently shown to play a role in the negative regulation of energy expenditure in white adipose cells, and its expression has been highly correlated with diet-induced obesity and insulin resistance in both mice and humans (Zhang et al., 2013). Moreover, many of our metabolic targets are conserved among different phyla. If LIN-29 truly triggers a metabolic switch, and some of the interactions of this transcription factor and other genes are conserved (i.e. with co factor NAB/Mab-10), our results could serve as the ground work for a new research direction, in which the *C. elegans* temporal developmental program and metabolic switch could be used as a model to study the genetic interactions underlying changes in metabolic decisions (i.e. expenditure versus storage), perhaps as a consequence of developmental transitions, in other organisms. Last but not least, if LIN-29 regulation indeed shares homology with EGR1, I believe our system (*C. elegans*) could be used to further explore potentially

conserved downstream interactions and learn how these regulate lipid metabolism or how they relate to metabolic diseases such as obesity and diabetes in humans.

References

Abete-Luzi, P., & Eisenmann, D. M. (2018, April 23). Regulation of C. elegans L4 cuticle collagen genes by the heterochronic protein LIN-29. *Genesis*. Wiley-Blackwell. https://doi.org/10.1002/dvg.23106

Araya, C. L., Kawli, T., Kundaje, A., Jiang, L., Wu, B., Vafeados, D., … Snyder, M. (2014). Regulatory analysis of the C. elegans genome with spatiotemporal resolution. *Nature*, *512*(7515), 400–405. https://doi.org/10.1038/nature13497

Askjaer, P. (2014). Modern techniques for the analysis of chromatin and nuclear organization in C. elegans. *WormBook*, 1–35. https://doi.org/10.1895/wormbook.1.169.1

Aughey, G. N., & Southall, T. D. (2016). Dam it's good! DamID profiling of protein-DNA interactions. *Wiley Interdisciplinary Reviews: Developmental Biology*, *5*(1), 25–37. https://doi.org/10.1002/wdev.205

Bettinger, J. C., Euling, S., & Rougvie, A. E. (1997). The terminal differentiation factor LIN-29 is required for proper vulvar morphogenesis and egg laying in

Caenorhabditis elegans. *Development*, *124*(21), 4333–4342.

Bettinger, J. C., Lee, K., & Rougvie, a E. (1996). Stage-specific accumulation of the terminal differentiation factor LIN-29 during Caenorhabditis elegans development. *Development*, *122*(8), 2517–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8756296

Budovskaya, Y. V., Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., & Kim, S. K. (2008). An elt-3/elt-5/elt-6 GATA Transcription Circuit Guides Aging in C. elegans. *Cell*, *134*(2), 291–303. https://doi.org/10.1016/j.cell.2008.05.044

Celniker, S. E., Dillon, L. A. L., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., … Waterston, R. H. (2009, June 18). Unlocking the secrets of the genome. *Nature*. NIH Public Access. https://doi.org/10.1038/459927a

Dowen, R. H., Breen, P. C., Tullius, T., Conery, A. L., & Ruvkun, G. (2016). A microRNA program in the C. elegans hypodermis couples to intestinal mTORC2/PQM-1 signaling to modulate fat transport. *Genes and Development*, *30*(13), 1515–1528. https://doi.org/10.1101/gad.283895.116

Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., … Trifonov, T. (2010). Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. *Science*, *330*(6012), 1775 LP-1787. https://doi.org/10.2436/20.7010.01.53

Gilleard, J. S., Barry, J. D., & Johnstone, I. L. (1997). cis Regulatory Requirements for Hypodermal Cell-Specific Expression of the Caenorhabditis elegans Cuticle Collagen Gene dpy-7, *17*(4), 2301–2311. Retrieved from

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC232079/pdf/172301.pdf

Gilleard, J. S., & McGhee, J. D. (2001). Activation of hypodermal differentiation in the Caenorhabditis elegans embryo by GATA transcription factors ELT-1 and ELT-3. *Molecular and Cellular Biology*, *21*(7), 2533–44. https://doi.org/10.1128/MCB.21.7.2533-2544.2001

Gómez-saldivar, G., Meister, P., & Askjaer, P. (2016). DamID Analysis of Nuclear Organization in Caenorhabditis elegans, *1411*. https://doi.org/10.1007/978-1-4939-3530-7

Gorrepati, L., Krause, M. W., Chen, W., Brodigan, T. M., Correa-Mendez, M., & Eisenmann, D. M. (2015). Identification of Wnt Pathway Target Genes Regulating the Division and Differentiation of Larval Seam Cells and Vulval Precursor Cells in *Caenorhabditis elegans*. *G3: Genes|Genomes|Genetics*, *5*(8), 1551–1566. https://doi.org/10.1534/g3.115.017715

Harris, D. T., & Horvitz, H. R. (2011). MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in C. elegans. *Development*, *138*(18), 4051–4062. https://doi.org/10.1242/dev.065417

Hayes, G. D., & Ruvkun, G. (2006). Misexpression of the Caenorhabditis elegans miRNA let-7 is sufficient to drive developmental programs. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 71, pp. 21–27). https://doi.org/10.1101/sqb.2006.71.018

Jackson, B. M., Abete-Luzi, P., Krause, M. W., & Eisenmann, D. M. (2014). Use of an Activated Beta-Catenin to Identify Wnt Pathway Target Genes in Caenorhabditis

elegans , Including a Subset of Collagen Genes Expressed in Late Larval Development. *G3: Genes|Genomes|Genetics*, *4*(4), 733–747. https://doi.org/10.1534/g3.113.009522

Narasimhan, K., Lambert, S. A., Yang, A. W. H., Riddell, J., Mnaimneh, S., Zheng, H., … Hughes, T. R. (2015). Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. *ELife*, *2015*(4), 1–53. https://doi.org/10.7554/eLife.06967

Newman, A. P., Inoue, T., Wang, M., & Sternberg, P. W. (2000). The Caenorhabditis elegans heterochronic gene lin-29 coordinates the vulval-uterine-epidermal connections. *Current Biology*, *10*(23), 1479–1488. https://doi.org/10.1016/S0960-9822(00)00827-7

Rougvie, A. E., & Ambros, V. (1995). The Heterochronic Gene Lin-29 Encodes a Zinc-Finger Protein that Controls a Terminal Differentiation Event in Caenorhabditis-Elegans. *Development*, *121*(8), 2491–2500. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7671813

Wählby, C., Lee Conery, A., Bray, M.-A., Kamentsky, L., Larkins-Ford, J., Sokolnicki, K. L., … O'Rourke, E. J. (2014). High- and low-throughput scoring of fat mass and body fat distribution in C. elegans. *Methods*, *68*(3), 492–499. https://doi.org/10.1016/j.ymeth.2014.04.017

Watts, J. L., & Browse, J. (2002). Genetic dissection of polyunsaturated fatty acid synthesis in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(9), 5854–5859.

https://doi.org/10.1073/pnas.092064799

Yin, J., Madaan, U., Park, A., Aftab, N., & Savage-Dunn, C. (2015). Multiple cis elements and GATA factors regulate a cuticle collagen gene in Caenorhabditis elegans. *Genesis*, *53*(3–4), 278–284. https://doi.org/10.1002/dvg.22847

Zhang, J., Zhang, Y., Sun, T., Guo, F., Huang, S., Chandalia, M., … Fu, M. (2013). Dietary obesity-induced Egr-1 in adipocytes facilitates energy storage via suppression of FOXC2. *Scientific Reports*, *3*, 1–10. https://doi.org/10.1038/srep01476