Analysis of *lin-42/period* through characterization of a null allele reveals the implications of all three isoforms

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Theresa Leigh Brix Edelman

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Ann E. Rougvie

December 2016

Acknowledgements

I am very grateful to all who have contributed to my success by supporting me through graduate school. I would first like to extend my gratitude to my advisor, Dr. Ann Rougvie. Thank you for sharing with me your expertise and love for science, for providing me with support and patience, and allowing me the freedom to grow. It has been an honor to work for such an exceptional mentor. Thank you for molding me into the scientist, and person, I am today.

Throughout most of my time in the Rougvie Lab, I have had the pleasure of working alongside several remarkable colleagues, particularly Dr. Tamar Resnick, Dr. Kate McCulloch and Karla Opperman. Thank you for providing me with a constant source of feedback and encouragement, and modeling what it is to be a great scientist. You will always be esteemed and treasured friends.

To the UMN *C. elegans* community, thank you for your generosity with reagents, time, wisdom, and friendship. I take with me many wonderful memories of the time we shared around the lab, at weekly Worm Club meetings, and at International *C. elegans* meetings.

To my parents, family and friends, thank you for pouring your love into me, Matt, and the kids along this vibrant and challenging journey. It takes a village to raise a family, and even more so while establishing a career. Thank you for being my village.

To my devoted husband, and treasured children, thank you for your unconditional love and sacrifice. You are and always will be my great joy.

And most importantly, to God, thank you for giving me the strength, perseverance, and wisdom to complete this great achievement.

Dedication

I dedicate this thesis to disabled people around the world, especially people with dwarfism, who are denied the right to an education, or a chance to live a respectable life because of their difference.

Abstract

The timing of post-embryonic development is regulated by the heterochronic gene pathway in *C. elegans*. An important member of this pathway is *lin-42*, the worm homolog of the circadian clock gene *period*. Existing *lin-42* alleles have implicated this gene in developmental timing, molting, and the decision to enter the alternative dauer state. *lin-42* is a complex locus, encoding overlapping and non-overlapping isoforms, and as a result hypomorphic alleles leave at least one isoform intact. A null allele is needed to understand all of the roles of *lin-42* and its individual isoforms. To remedy this problem, and discover the null phenotype, we engineered an allele that deletes the entire *lin-42* protein coding region. *lin-42* null mutants are homozygous viable, but have more severe phenotypes than observed in previously characterized hypomorphic alleles. Additional evidence is also provided for this conclusion by using the null allele as a base for reintroducing different isoforms, showing that each isoform can provide heterochronic and molting pathway activities. *lin-42* functions in the heterochronic pathway to regulate transcription of primary miRNAs, including the *let-7* miRNA family. One member of the *let-7* family is *mir-48*. To better understand how *mir-48* fits in the pathway a *mir-48* gain-of-function suppressor screen was performed. This screen yielded multiple candidates that could function with *mir-48* in the heterochronic gene pathway, as miR-48 target genes, or as regulators of miRNA biogenesis and function. Analysis of these mutants should provide insights into control of developmental time.

Table of Contents

List of Figures

List of Tables

Clarification of contributions

Chapter II

Primers and targeting plasmid used to generate the *lin-42(0)* were made by Angela Barr. Injections and isolation of the *lin-42(0)* alleles were done by Christian Frøkjær-Jensen while in the laboratory of Erik Jorgensen. Analysis of the *lin-42(0)* molting phenotypes were done in collaboration with Katherine McCulloch and Ann Rougvie.

Chapter III

The *mir-48* overexpression screen was performed by Sarah Malmquist, Greg Sindberg, Kristen Reyna, Sarah Smestad. Initial analysis of the suppressors was performed in collaboration with Sarah Malmquist, Greg Sindberg, Kristen Reyna, Sarah Smestad. Analysis of the *lin-66* mutants was done by Sarah Malmquist and Tamar Resnick.

CHAPTER 1: INTRODUCTION

Temporal regulation is important for proper development.

As an embryo develops into a reproductive adult, each morphological change is coordinated by a network of molecular mechanisms. Spatial cues instruct cells about their position and how to behave appropriately, while temporal cues ensure developmental events occur in the correct order and at the precise time. Ordering determines the sequence of events (A → B), whereas interval timing determines the duration between each event (A….B). Temporal regulation is not well understood and is an important area of study.

In humans, one example of a temporally regulated developmental transition is puberty, in which an individual undergoes physiological changes necessary to reach sexual maturity (ROSENFIELD *et al.* 2009). Puberty normally occurs between the ages of 8 and 13 in girls, and 9 and 14 in boys, and is stimulated by the peptide kisspeptin (HARRINGTON and PALMERT 2012; SORENSEN *et al.* 2012; SKORUPSKAITE *et al.* 2014). Kisspeptin activates neurons in the hypothalamus to release a pulse of gonadotropin-releasing hormone (GnRH) (SKORUPSKAITE *et al.* 2014). GnRH stimulates the release of luteinizing hormone (LH) and follicle stimulating hormone (FSH) from the pituitary gland. LH and FSH then activate secretion of estrogen from the ovaries in females, and testosterone from the testes in males, which in turn initiates the appropriate sexual maturation program (SIZONENKO 1978).

Individuals that undergo puberty before the age of 8 in girls, or 9 in boys, are considered to have Central Precocious Puberty (CPP) disorder. Precocious

puberty is associated with an increased risk for a number of diseases later in life, including cardiovascular disease, cancer, and type 2 diabetes (ELKS *et al.* 2013; PRENTICE and VINER 2013; BODICOAT *et al.* 2014). Experiments in primates and genome wide association studies in humans have implicated a number of genes as possible causes of precocious puberty, including KISS1, KISS1R, and LIN-28B; however, mutations in these genes are rarely identified in patients, which suggests that timing of puberty is complex and likely has multiple origins (BRITO *et al.* 2006; FREITAS *et al.* 2007; TOMMISKA *et al.* 2011; SILVEIRA-NETO *et al.* 2012). Thus, additional work is necessary to understand how developmental transitions, such as puberty, are regulated. Since molecular mechanisms in higher eukaryotes are complex, a simple biological model can provide the tools necessary to investigate mechanisms of developmental timing.

***C. elegans* is an exceptional model to study timing.**

C. elegans is an exceptional system to study genetic mechanisms that regulate developmental timing, because it is a powerful genetic model and has a precisely timed developmental program (BRENNER 1974; SULSTON and HORVITZ 1977). The worm has a short life cycle of 2.5 days from hatching to adulthood which makes it feasible to rapidly perform complex genetic experiments. *C. elegans* exist as two sexes, self-fertilizing hermaphrodites (XX) and males (XO), a feature that is useful in genetic experimentation because homozygous hermaphrodites produce genetically identical offspring, and males can be used to move mutations between strains. Furthermore, there are multiple transgenic,

RNAi and gene editing approaches to investigate the function of a particular gene of interest.

The precisely timed development of *C. elegans* makes it an especially useful model to study temporal development. *C. elegans* development goes through embryogenesis and four larval stages (L1-L4) before reaching adulthood. Each stage ends with a larval molt, in which the animal becomes lethargic, produces a new cuticle, and sheds the old one (SINGH AND SULSTON 1978). Molting is a transition in which one developmental program ends and a new one begins. In *C. elegans*, these stages and transitions occur at precise times in development. Furthermore, the cell lineage of the worm is essentially invariant, which makes it possible to track developmental events as specific as a single cell division (SULSTON and HORVITZ 1977). One set of cells that is especially useful to track developmental time are the hypodermal seam cells because they execute stage specific programs that are easy to identify. The hypodermal seam cells are epithelial stem cells, which are located in two lateral rows along either side of the worm and contribute to secreting collagens that form the cuticle (Figure 1A). At the beginning of each larval stage they divide in a stem cell-like division pattern (Figure 1C). Upon each division the anterior daughter exits the seam and fuses with the surrounding syncytial hypodermis while the posterior daughter retains a seam cell fate. Two additional events are easy to visualize and are used to track developmental time. The first is at the beginning of the L2 stage, when a subset of seam cells divide symmetrically, increasing the total number of seam cells on each lateral side from 10 to 16. The second event occurs at the L4-Adult

transition when seam cells terminally differentiate and fuse to form a syncytium (Figure 1C). The syncytium secretes collagens that form an adult specific cuticle containing lateral ridges along the length of the worm, called alae (Figure 1B).

Stage-specific programs, such as the seam cell proliferative division and terminal differentiation, are useful because they follow a temporal pattern, and mutants can be identified in which these temporal patterns are disrupted. In animals where early temporal events are skipped, subsequent events occur too early, and mutants are referred to as "precocious mutants." Alternatively, when earlier events are reiterated, subsequent events occur too late, or not at all, and such mutants are referred to as "retarded mutants" (Figure 1C). Gene mutations resulting in precocious and retarded phenotypes were initially identified and organized into a regulatory pathway, now called the heterochronic gene pathway (AMBROS and HORVITZ 1984; AMBROS and MOSS 1994; SLACK and RUVKUN 1997).

The heterochronic pathway regulates the timing of *C. elegans* development.

The *C. elegans* heterochronic gene pathway comprises a regulatory cascade that ensures developmental programs are executed in the correct order and at the precise time. Each step within the pathway is controlled by key input from a specific gene. The L1-to-L2 transition is regulated by *lin-14 (AMBROS and HORVITZ 1987)*, the L2-to-L3 transition is regulated by *hbl-1 (ABRAHANTE et al. 2003; LIN et al. 2003)*, and the L4-to-Adult transition is regulated by *lin-29* (Figure 2) (ROUGVIE and AMBROS 1995). Both *lin-14* and *hbl-1* promote early larval

stages and must be down-regulated for animals to execute subsequent larval programs (LEE *et al.* 1993; WIGHTMAN *et al.* 1993; ABBOTT *et al.* 2005). However, the L4-to-Adult transition is regulated somewhat differently since *lin-29* promotes adult programs, and must be up-regulated for development to proceed (ROUGVIE and AMBROS 1995). Because *lin-29* is inhibited by *lin-41* and *hbl-1*, downregulation of *lin-41* and *hbl-1* is required to relieve repression on *lin-29* and allow development to progress (Figure 2) (ABRAHANTE *et al.* 2003; LIN *et al.* 2003).

Expression levels of *lin-14*, *hbl-1* and *lin-41* are each temporally downregulated at precise times during development by a miRNA or group of miRNAs (RESNICK *et al.* 2010). Lin-4 and Let-7, the first two miRNAs discovered, were identified in *C. elegan*s because gene mutations resulted in retarded heterochronic phenotypes (Figure 1C) (CHALFIE *et al.* 1981; REINHART *et al.* 2000). *lin-4* mutants reiterate early larval stages, while *let-7* mutants reiterate late larval stages, suggesting they function at different times in development. Further studies demonstrated that *lin-4* is expressed during the L1 stage to down-regulate *lin-14*, while *let-7* expression increases during the L3 stage to downregulate *lin-41* and promote the transition to adulthood (Figure 2) (LEE *et al.* 1993; REINHART *et al.* 2000). Later, a group of miRNAs, miR-48/241/84, were shown to function redundantly during the L2 stage to downregulate *hbl-1* and promote the transition to the L3 stage (ABBOTT *et al.* 2005; LI *et al.* 2005).

Before these findings, RNAs this small had not been recognized as regulators of gene expression. miRNAs are small ~21nt RNAs that bind to the

3’UTR of mRNAs and cause translational inhibition or mRNA degradation. Each miRNA or group of miRNAs is expressed at specific times to temporally regulate their targets. miRNA expression is regulated through a complex and highly regulated process (Figure 3). miRNA biogenesis has recently been reviewed (WINTER *et al.* 2009; LIN and GREGORY 2015). Briefly, miRNAs are transcribed as long primary transcripts (pri-miRNA) are are processed in the nucleus by an enzyme complex called the micro-processor into a 60-70nt stem-loop, or hairpin, precursor miRNA (pre-miRNA) (DENLI *et al.* 2004; GREGORY *et al.* 2004; LEE *et al.* 2004b). This pre-miRNA is then exported into the cytoplasm and then further processed by dicer into a small ~21nt hetero-duplex (PARK *et al.* 2011). The duplex is then loaded onto RISC (RNA-induced silencing complex), and the miRNA complement strand is released from the hetero-duplex through a passive mechanism (KAWAMATA *et al.* 2009). The miRNA-RISC (miRISC) complex then base pairs to the 3’UTR of an mRNA through imperfect pairing and either inhibits translation or targets the mRNA for degradation. Since the interaction is imperfect, each miRNA can bind to multiple target sequences.

miRNAs have since been discovered to be abundant across plant and animal kingdoms and function to regulate almost every biological pathway (FRIEDMAN *et al.* 2009). miRNA levels have been shown to be altered in certain diseases, especially cancer, and have the potential to be used for diagnostics or disease treatments (WANG *et al.* 2009; HATA and KASHIMA 2016; HU *et al.* 2016; KATCHY and WILLIAMS 2016; O'REILLY 2016).

The precise temporal regulation of miRNA expression, processing and function is key to controlling the heterochronic gene pathway, and I was intrigued by how this is achieved. Most of my thesis research centers on the gene *lin-42,* which regulates the expression of miRNA primary transcripts, including *let-7* and *mir-48*. Another aspect of my research specifically investigated the function of *mir-48*. Interestingly, *let-7* and *mir-48* are members of the same miRNA family, but *mir-48* functions to promote the L2-to-L3 transition, while *let-7* functions to regulate the larval-to-adult transition. Since *lin-42* is expressed during each larval stage, it has the opportunity to regulate both miRNAs.

miR-48, miR-241, and miR-84 regulate mid larval programs.

Let-7 and miR-48 are members of a family of miRNAs that share homology in the seed region, which includes nucleotides 2-8 of the miRNA that enables miRNAs to recognize and bind to their targets (LAU *et al.* 2001; LIM *et al.* 2003). miRNAs with a common seed region often have overlapping targets and function redundantly, and thus have been grouped into miRNA families. Three members of the *let-7* family, *mir-48*, *mir-241*, and *mir-84*, collectively promote the *C. elegans* L2-to-L3 transition (ABBOTT *et al.* 2005). Disruption of *mir-48*, *mir-241*, or *mir-84* causes little or no phenotype, whereas disruption of all three miRNAs has a strong heterochronic defect that results in reiteration of the L2 stage pattern of seam cell division. When the L2 stage is reiterated, animals have an increased number of seam cells and a delay in alae formation (Figure

1C). *mir-48/84/241* function at the L2 stage to down-regulate their targets including *hbl-1*, *daf-12*, and possibly *lin-28* (Figure 2).

hbl-1 is the homolog of the *Drosophila hunchback* transcription factor (FAY *et al.* 1999). *C. elegans hbl-1* is essential for embryonic development, but also regulates larval development (ABRAHANTE *et al.* 2003; LIN *et al.* 2003). *hbl-1* promotes L2 programs and inhibits L3 programs; therefore down regulation of *hbl-1* is required for development to proceed to the L3 stage. *hbl-1* is expressed in the hypodermis and seam during the L1 and L2 stages but is no longer detectable by the L3 stage. Post-embryonic RNAi depletion of *hbl-1* causes L2 seam cell programs to be skipped, resulting in a decreased number of seam cells and precocious alae (ABRAHANTE *et al.* 2003; LIN *et al.* 2003). Consistent with *hbl-1* as a target, disruption of *mir-48/241/84* causes expression of *hbl-1* to persist, and thus a reiteration of the L2 stage fates in the hypodermis (ABBOTT *et al.* 2005).

lin-28 functions in opposition to *mir-48/241/84* at the L2 stage to promote L2 programs (MOSS *et al.* 1997; VADLA *et al.* 2012). *lin-28* mutants skip the L2 stage, as demonstrated by a decreased number of seam cells and the presence of precocious alae. *lin-28* promotes L2 programs by supporting *hbl-1* expression (VADLA *et al.* 2012), whereas *mir-48/241/84* inhibit *hbl-1* expression. *lin-28* is controlled by *lin-66*, which is a novel cytoplasmic protein that negatively regulates *lin-28* post-transcriptionally through its 3'UTR (MORITA and HAN 2006) (Figure 2). It is unclear whether *lin-28* is also regulated by *mir-48/84/241* through binding to its 3'UTR. Disruption of a predicted *let-7* family binding site results in

misregulation of a *lin-28::LacZ* reporter (MORITA and HAN 2006). However, when a rescuing *lin-28::gfp* reporter was expressed in a *mir-48/84/241* mutant, its expression was not affected (ABBOTT *et al.* 2005). Further evidence is needed to determine whether the *let-7* family regulates *lin-28*. Since *lin-28* promotes *hbl-1* expression, repression of *lin-28* by the *let-7* family could help reduce *hbl-1* expression and promote developmental progression (Figure 2).

let-7 functions in late larval stages to promote developmental progression.

let-7 functions later in development to regulate the final developmental transition, a point in which the stem-cell-like seam cells terminally differentiate (REINHART *et al.* 2000; VADLA *et al.* 2012). *let-7* mutants execute early programs normally, but reiterate a later larval stage causing alae to form too late, resulting in a retarded phenotype (REINHART *et al.* 2000; VADLA *et al.* 2012).

let-7 miRNA expression quickly increases during the L3 stage (REINHART *et al.* 2000; VADLA *et al.* 2012; MCCULLOCH and ROUGVIE 2014) and functions to target *lin-41 (SLACK et al. 2000; VELLA et al. 2004)*. *lin-41* expression levels begin to decrease during the L3 stage in the epidermis and are significantly reduced by the L4 stage (SLACK *et al.* 2000). *let-7* mutants have a retarded defect, executing a 5th larval stage before transitioning to adulthood, and conversely, *lin-41* mutants have a precocious defect in which seam cells transition to adulthood one stage too early (REINHART *et al.* 2000; SLACK *et al.* 2000). Thus, for wild type development, Let-7 levels increase and downregulate *lin-41*, allowing the transition to adulthood. The regulatory relationship between *let-7* and *lin-41,* is a

well-studied model of miRNA gene regulation. It is an especially fascinating relationship because it is conserved between relatives as distant as nematodes and humans (PASQUINELLI *et al.* 2000; LAGOS-QUINTANA *et al.* 2003).

The conserved target of *let-7*, *lin-41*, encodes an RNA-binding protein that belongs to the TRIM-NHL protein family, which contains a RING finger, B-box, coiled-coil domain, as well as six NHL repeats (NHL = NCL-1, HT2A2, and LIN-41) (SLACK *et al.* 2000). It has many conserved functions in both worms and vertebrates including roles in miRNA processing (RYBAK *et al.* 2009), translational repression, and cell proliferation (CHANG *et al.* 2012; CHEN *et al.* 2013; LOEDIGE *et al.* 2013). In the heterochronic pathway, *lin-41* functions, perhaps indirectly, to downregulate *lin-29*, the most down-stream member of the heterochronic pathway (SLACK *et al.* 2000). LIN-29 is a zinc-finger transcription factor that regulates the final seam cell differentiation step (ROUGVIE and AMBROS 1995). This larval-to-adult transition requires a reduction in LIN-41 levels, which relieves repression on LIN-29 and in turn promotes adult programs.

Tight regulation *of let-7* ensures that *lin-41* repression does not occur until late larval stages. One of the most well understood regulators of *let-7* is *lin-28*, which was first discovered as a heterochronic gene that promotes L2 programs (AMBROS and HORVITZ 1984). It was later discovered as a regulator of Let-7 processing in humans (HEO *et al.* 2008; NEWMAN *et al.* 2008; RYBAK *et al.* 2008; VISWANATHAN *et al.* 2008), which was surprising since *let-7* was believed to function much later in development than *lin-28*. Subsequently, *lin-28* was confirmed to negatively regulate Let-7 miRNA processing in *C. elegans*

(LEHRBACH *et al.* 2009). In both nematode and vertebrate systems, LIN-28 directly binds to the let-7 precursor and recruits a poly(U) polymerase which promotes degradation of the let-7 pre-miRNA (HEO *et al.* 2008; LEHRBACH *et al.* 2009). Further work in *C. elegans* has shown that *lin-28* functions in a two-step mechanism to regulate development. The first step is through a *let-7* independent mechanism in which *lin-28* promotes *hbl-1* expression. The second step is by regulating *let-7* processing to prevent late developmental transitions from occurring too early (VADLA *et al.* 2012). Growing evidence suggests that a two-step role for *lin-28* is conserved in vertebrates (TSIALIKAS and ROMER-SEIBERT 2015).

lin-28, let-7, and *lin-41* are all conserved in vertebrates and play important roles to regulate stem cell renewal and differentiation. *lin-28* and *lin-41* promote self-renewal in stem and progenitor cells, while *let-7* promotes embryonic stem cell differentiation (FAUNES and LARRAIN 2016). In vertebrates, *let-7* is considered a tumor suppressor or differentiation factor, since mutations in *let-7* lead to tumor formation and over-expression of *let-7* causes premature cell cycle exit, similar to what is seen in seam cells.

lin-42 is a homolog of _period_ and a member of the heterochronic gene pathway.

lin-42 genetically interacts with both *let-7* and *mir-48*, and has been shown to regulate expression of their primary transcripts (Figure 2) (MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014). *lin-42(lf)*

mutants have a precocious heterochronic defect in which seam cell terminal differentiation and alae formation occurs one stage too early at the L3m (Figure 1C) (JEON *et al.* 1999; TENNESSEN *et al.* 2006). *lin-42* is particularly interesting because it has a dynamic expression pattern and complex genomic structure, as described in more detail below (Figure 4A&B). *lin-42* is a homolog of *period*, a member of the circadian clock in flies and mammals (JEON *et al.* 1999). *period* plays a role in regulating the circadian 24-hour light/dark cycle, which was clearly demonstrated by the first *period* mutants found in *Drosophila* that either shorten (per^s), lengthen (per^l), or abolish (per^0) circadian rhythms (KONOPKA and BENZER 1971; HARDIN 2005; YU and HARDIN 2006).

In flies, the circadian clock is regulated by a negative feedback loop in which *clock* and *cycle* inhibit *period* and *timeless* (Figure 5). During the day, CLOCK (CLK) and CYCLE (CYC) heterodimerize to form a transcription factor that activates expression of *period* (*per*) and *timeless* (*tim*) (HAO *et al.* 1997; DARLINGTON *et al.* 1998; MCDONALD *et al.* 2001; WANG *et al.* 2001). Initially, PER and TIM levels remain low because PER is phosphorylated and marked for degradation by CASEIN KINASE 2 (CK2) and DOUBLETIME (DBT) (KLOSS *et al.* 1998; PRICE *et al.* 1998; AKTEN *et al.* 2003; NAWATHEAN and ROSBASH 2004). TIM is destabilized by CRYPTOCHROME (CRY) during the day, however at night TIM begins to stabilize because CRY is only active during the day (ASHMORE and SEHGAL 2003). As TIM accumulates, PER and TIM dimerize and form a complex with DBT (KLOSS *et al.* 2001; SAEZ *et al.* 2007). Next, TIM is phosphorylated by SHAGGY (SGG) and allows the complex to enter the nucleus (MARTINEK *et al.*

2001). TIM-PER then bind with CLK-CYC, which halts transcription of *per*, *tim* and other targets *(*LEE *et al. 1999)*. The following day, light activates CRY which destabilizes TIM, and relieves the repression on CLK-CYC (EMERY *et al.* 1998; LIN *et al.* 2001; BUSZA *et al.* 2004; DISSEL *et al.* 2004). CLK-CYC can then begin the cycle again by re-activate *per* and *tim* transcription.

There are both similarities and differences between clock genes in flies and worms. *lin-42* has recently been shown to negatively regulate transcription, similar to PERIOD (MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014). In *C. elegans*, *lin-42* genetically interacts with kin-20/doubletime and tim-1/timeless because the *lin-42(lf)* precocious defect was enhanced when mutants were fed *kin-20* or *tim-1* RNAi (BANERJEE *et al.* 2005). Additionally, wild type animals fed *kin-20* and *tim-1* RNAi had a mild precocious phenotype, however *lin-42* did not interact with *kin-20* or *tim-1* in a yeast-two hybrid assay (Gardner 2005). Surprisingly, in *C. elegans tim-1* expression does not oscillate like *timeless* does in flies, rather its expression increases as development progresses and plays a role in regulating chromosome cohesion (SEHGAL *et al.* 1995; JEON *et al.* 1999; CHAN *et al.* 2003). Further investigation is needed to better understand the role *kin-20* and *tim-1* play in developmental timing. Mutations in the *C. elegans* clock/cycle homolog, *aha-1*, do not result in a heterochronic defect, nor has AHA-1 been shown to interact with other clock related genes (BANERJEE *et al.* 2005). The reason for this can possibly be explained by the difference between a cyclical diurnal clock and a linear developmental clock. A cyclical diurnal clock regulates the expression of the

same set of genes during each time interval, whereas development requires expression of a unique set of genes at each developmental stage (BANERJEE *et al.* 2005). Since *lin-42*, *kin-20*, *tim-1* regulate the linear developmental clock in *C. elegans*, a negative feedback loop may not be necessary.

LIN-42 and PERIOD have several shared homology domains, including the N-terminal PAS domains and the C-terminal SYQ and LT domains (JEON *et al.* 1999; TENNESSEN *et al.* 2006). PAS domains (Per-Arnt-Sim) are protein interaction domains and required for PERIOD dimerization and nuclear entry in both flies and mammals (YAGITA *et al.* 2000; HENNIG *et al.* 2009). SYQ and LT are smaller domains that reside in a region that interacts with circadian clock proteins, including CLOCK and Casein Kinase Iε, in flies and mammals (CHANG and REPPERT 2003; LEE *et al.* 2004a; SUN *et al.* 2010). *period* in flies and mammals is expressed as one transcriptional unit that encodes all of the conserved homology domains; however, *lin-42* is expressed as three transcriptional units and two of the *lin-42* transcripts, *lin-42a* and *lin-42c,* are non-overlapping (Figure 4A) (TENNESSEN *et al.* 2006). *lin-42a* encodes a protein containing the SYQ and LT domains, while *lin-42c* encodes a protein containing the PAS domains (Figure 4A). In contrast, *lin-42b* encodes all conserved domain structures, and therefore most closely resembles *period* (TENNESSEN et al. 2006).

Similar to PERIOD, LIN-42 is expressed cyclically, however its expression cycles with each developmental stage rather than the light-dark cycle. RT-PCR experiments that detected both *lin-42b* and *lin-42c* transcripts demonstrate that they accumulate during the intermolt and then are degraded so that they are no

longer detectable by the molt (JEON *et al.* 1999). Since *lin-42a* is expressed from a different promoter than *lin-42b/c* there is a possibility that their expression patterns differ. Transcriptional reporters were used to compare *lin-42a* and *lin-42b* expression and the results suggest that their temporal expression profiles differ, but there is disagreement about the order in which they are expressed (MONSALVE *et al.* 2011; PERALES *et al.* 2014). A more stringent analysis is needed to look directly at the expression profiles of these transcripts to determine whether they indeed differ.

The three transcriptional units are a unique feature of *lin-42* and provide the opportunity to study their functions separately. However, genetic analysis of *lin-42* is complex because extant hypomorphic alleles leave one isoform intact. Upstream mutations affect *lin-42b* and *lin-42c,* while downstream mutations affect *lin-42a* and *lin-42b* (Figure 4A). Mutations that affect either LIN-42A or LIN-42C both result in a precocious heterochronic defect, even though they are completely different proteins (TENNESSEN *et al.* 2006). Interestingly, LIN-42A can rescue upstream and downstream mutations when over-expressed, whereas LIN-42C largely can only rescue upstream mutations when over-expressed, demonstrating that LIN-42A plays a particularly important role in the function of LIN-42 (TENNESSEN *et al.* 2006). A *lin-42* null allele is a long awaited tool required to fully investigate how each isoform functions on its own.

Although it is clear that *lin-42* plays an important role in *C. elegans* development, the argument remains as to whether this nematode has circadian rhythms, and whether *lin-42* plays a role in regulating those behaviors. Several

studies have looked at behavioral oscillations using video-tracking (KIPPERT *et al.* 2002), eye inspection recordings (SAIGUSA *et al.* 2002), and an automated tracking system (SIMONETTA and GOLOMBEK 2007; SIMONETTA *et al.* 2009). These studies argue that *C. elegans* has behavioral oscillations in response to light or temperature. Other studies looked at gene expression oscillations during adulthood and provided evidence for entrained gene expression cycling (VAN DER LINDEN *et al.* 2010; OLMEDO *et al.* 2012) during adulthood. One particular study compared behavioral oscillations between wild type and *lin-42* mutants beyond the L4 stage and argued that the circadian period was lengthened in *lin-42* mutants (SIMONETTA *et al.* 2009); however the data from this study were not robust and results varied between animals/samples. A role for LIN-42 in regulating a circadian clock in adulthood is not likely, since *lin-42* mRNA levels are low and have not been observed to oscillate past larval development (JEON *et al.* 1999; TENNESSEN *et al.* 2006; VAN DER LINDEN *et al.* 2010; TEMMERMAN *et al.* 2011; OLMEDO *et al.* 2012).

Although it is unlikely that *lin-42* regulates circadian behaviors, evidence continues to emerge for circadian rhythms in *C. elegans*. Recently, one study demonstrated that *pdf-1* (pigment dispersing factor) is necessary to maintain circadian rhythmic movement behaviors in *C. elegans,* with data that is more robust and consistent than previous circadian behavioral studies (HERRERO *et al.* 2015). *pdf-1* plays an important role in maintaining circadian locomotor activity in *D. melanogaster (RENN et al. 1999)*, and in *C. elegans pdf-1* has been implicated in regulating locomotion and egg-laying, behaviors that are more advantageous

in the daytime when temperatures are elevated (BARRIOS *et al.* 2012; MEELKOP *et al.* 2012). Future studies are needed to better understand how circadian rhythms are regulated in *C. elegans*; however it is clear that they are regulated by a more primitive clock then higher metazoans such as *Drosophila* and mice. Further studies using *C. elegans* could provide insight into the convergence of developmental and circadian clocks.

***lin-42* regulates miRNA expression.**

lin-42 was believed to be the only heterochronic gene that had a cyclical expression pattern, until it was discovered that primary transcripts of *lin-4, mir-48/mir-241/mir-84*, and *let-7* also cycle (VAN WYNSBERGHE *et al.* 2011; MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014). These cyclical expression patterns were unexpected since the mature form of each miRNA is expressed for one block of time. Since primary transcripts are present at times when the mature miRNA does not accumulate, this illustrates the sophisticated regulatory mechanisms required to ensure each miRNA is processed and stabilized at the appropriate time. *lin-42* is a great candidate for a regulator of miRNA expression because its expression pattern cycles with these pri-miRNAs and it genetically interacts with *lin-4* and *let-7* (ABRAHANTE et al. 1998; REINHART et al. 2000; TENNESSEN et al. 2006).

Recently, three studies demonstrated that *lin-42* functions as a negative regulator of miRNA expression, but is not required for the cycling of primary miRNA transcripts. qPCR, northern blot analysis, and transcriptional reporters

demonstrated that levels of both primary and mature forms of *let-7* and *mir-48* increase in *lin-42* mutants (McCulloch and Rougvie 2014; Perales *et al.* 2014; Van Wynsberghe *et al.* 2014). Misexpression of *let-7* and *mir-48* suggests that *lin-42* functions upstream in the heterochronic pathway.

Genetic epistasis analysis was performed between *lin-42* and the *let-7* family miRNAs to investigate their relationships within the heterochronic gene pathway. *lin-42* null mutants (*lin-42(0))* have a strong precocious seam cell defect, and *let-7(0)* and *mir-48(0)* are the only *let-7* family genes to suppress the *lin-42(0)* alae defect (McCulloch and Rougvie 2014). Additionally, *lin-42(0)* suppresses the *let-7(0)* retarded defect and that suppression is dependent on the increase in miR-48, since deletion of *mir-48* abrogates the suppression (McCulloch and Rougvie 2014). These genetic results support a model in which *lin-42* functions upstream of *let-7* and *mir-48* to regulate developmental timing. Furthermore, *Plet-7::gfp* was shown to be misregulated in a *lin-42(lf)* mutant, demonstrating that *lin-42* regulates *let-7* transcription.

Genome wide studies have also linked *lin-42* to the regulation of other miRNAs (Perales *et al.* 2014; Van Wynsberghe *et al.* 2014). These results suggest that *lin-42* has a global role in gene regulation outside the heterochronic pathway. ChIP analysis indicates that LIN-42 associates with chromatin near putative transcription start sites of both miRNAs and mRNAs (Perales *et al.* 2014). Since an extensive number of genes oscillate with the molting cycle (Hendriks *et al.* 2014), it is possible that *lin-42* could contribute to that regulation. One hypothesis is that *lin-42* controls oscillating expression patterns, since

period maintains cyclical expression of circadian genes in other systems (HARDIN *et al.* 1990; HARDIN *et al.* 1992). Surprisingly, heterochronic pri-miRNA transcript levels oscillate in *lin-42* mutants (MCCULLOCH and ROUGVIE 2014). Future genetic and molecular studies will be important to determine the degree to which *lin-42* regulates gene expression.

lin-42 promotes development through the molting pathway.

Developmental timing mechanisms regulate the sequence, spacing, and rhythmicity of biological events. *lin-42* was first identified because the order of development was altered in *lin-42(lf)* mutants (ABRAHANTE *et al.* 1998; JEON *et al.* 1999), but more recently, *lin-42* has also been shown to be important to regulate the rhythmic behavior of molting (MONSALVE *et al.* 2011).

Molting is a characteristic of ecdysozoan development in which animals synthesize a new exoskeleton and shed the old one (AGUINALDO *et al.* 1997). Molting occurs in a series of steps in which the worm decreases activity and stops feeding (lethargus), the old exoskeleton is released from the animal (apolysis) and the worm emerges (ecdysis) to complete the process (Singh and Sulston 1978).

Wild-type animals molt at precise times during development and complete each molt in 1-2 hours. However, *lin-42(ok2385)* mutants have delays entering and exiting each molt (MONSALVE *et al.* 2011). For instance, wild-type animals take an average of 6.6 hrs to complete the L2 stage and 1.7 hrs to complete the L2 molt, whereas *lin-42(ok2385)* mutants take ~12.6 hrs to complete the L2 stage

and ~5.1 hrs to complete the molt. Additionally, more than half of the larvae arrested development and were often found to be trapped in one or more cuticles, supporting the hypothesis that *lin-42* is required to complete the molting process. The *lin-42(ok2385)* allele deletes *lin-42a,* which includes the SYQ and LT domains, and leaves only *lin-42c* intact. Overexpression of *lin-42a* from a multi-copy array is sufficient to suppress the molting defect (MONSALVE *et al.* 2011). This molting phenotype has only been observed when *lin-42a* is deleted and not when *lin-42c* is disrupted, suggesting that *lin-42a* plays a role in molting. A null allele that deletes all *lin-42* transcripts would be helpful to determine if *lin-42a* contributes to the role *lin-42* plays in molting. Additionally, isoform specific rescue experiments of a null allele would also be useful to delineate isoform specific contributions to the role *lin-42* plays to regulate molting.

In arthropods, molts are triggered by a pulse of ecdysone which is converted to the 20-hydroxyecdysone steroid hormone (THUMMEL 1996; GILBERT *et al.* 2002). Neither ecdysone nor its receptors have been detected in nematodes (CHITWOOD 1999), however two nuclear hormone receptors, NHR-23 and NHR-25, which are homologous to ecdysone-responsive gene products DHR3 and Ftz-F1, are found in *C. elegans* and have been shown to be critical molting regulators (KOSTROUCHOVA *et al.* 1998; ASAHINA *et al.* 2000; GISSENDANNER and SLUDER 2000; KOSTROUCHOVA *et al.* 2001). *nhr-23(RNAi)* larvae have been reported as being small, having difficulties shedding old cuticles after molts, and developmentally arresting as larvae (KOSTROUCHOVA *et al.* 1998). *nhr-25(RNAi)* larvae have arrested growth phenotypes and difficulties

shedding cuticles (GISSENDANNER and SLUDER 2000). *nhr-23(RNAi)* and *nhr-25(RNAi)* phenotypes resemble molting defects observed in *lin-42(ok2385)* mutants (MONSALVE *et al.* 2011). Additionally, mutations in *nhr-25* enhance both the heterochronic and molting phenotypes found in *lin-42(lf)* mutants, suggesting it may act in parallel to or downstream of *lin-42* to coordinate larval molts (HADA *et al.* 2010; MONSALVE *et al.* 2011). Further analysis is needed to understand the nature of their relationships.

lin-42 functions with _daf-12_ to regulate dauer entry.

lin-42 plays multiple roles during development to ensure animals progress to adulthood rapidly through its role in regulating molting and miRNA expression. *lin-42* also controls the decision to enter a stress-resistant alternative larval stage, called dauer, and pause development if environmental conditions are not favorable for reproduction (TENNESSEN *et al.* 2010). The nuclear hormone receptor *daf-12* modulates the decision to continue or pause development by either promoting or inhibiting *mir-48/mir-241/mir-84* expression (BETHKE *et al.* 2009). In favorable environmental conditions the DAF-12 ligand, dafachronic acid (DA), is abundant and binds DAF-12. Ligand bound DAF-12 promotes *mir-48/mir-241/mir-84* expression, which in turn promotes developmental progression. In unfavorable environmental conditions DA concentrations are low and DAF-12 binds to its co-repressor DIN-1 (LUDEWIG *et al.* 2004). DIN-1 bound DAF-12 negatively regulates *let-7* family miRNAs (BETHKE *et al.* 2009; HAMMELL *et al.* 2009a) and promotes dauer formation (ANTEBI *et al.* 1998; ANTEBI *et al.*

2000). To ensure development continues during favorable conditions, miR-48/241/84 feeds back to inhibit *daf-12* and prevent a developmental pause (HAMMELL *et al.* 2009a).

lin-42 has been shown to antagonize *daf-12* during both continuous and discontinuous development (TENNESSEN *et al.* 2006; TENNESSEN *et al.* 2010). During continuous development, *daf-12(lf)* mutants have a retarded seam cell defect while *lin-42(RNAi)* animals have a precious alae defect, and both defects are suppressed in *daf-12(lf)*; *lin-42(RNAi)* animals *(TENNESSEN et al. 2006)*. These results suggest *daf-12* and *lin-42* function antagonistically in parallel, since there is mutual suppression. Since ligand bound DAF-12 promotes *mir-48/241/84* expression, while *lin-42* inhibits *mir-48* expression, their function likely converges through the *let-7* family miRNAs (BETHKE *et al.* 2009; HAMMELL *et al.* 2009a; MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014).

The antagonistic relationship between *lin-42* and *daf-12* is maintained when environmental conditions are unfavorable. During mildly stressful conditions, *lin-42(lf)* mutants have a high incident dauer phenotype, whereas *daf-12(lf)* mutants have a dauer-defective phenotype (TENNESSEN *et al.* 2010). These results demonstrate that *lin-42* promotes continuous development during mildly stressful conditions while DIN-1 bound DAF-12 promotes discontinuous development. DAF-12 functions as a sensor of environmental conditions, while LIN-42 functions to inhibit dauer entry. Together LIN-42 and DAF-12 fine-tune

the decision between continuous and discontinuous development and ensure animals develop and reproduce during favorable or mildly stressful conditions.

Impact of studying *C. elegans* heterochronic genes *lin-42* and *mir-48*

lin-42 is an intriguing heterochronic gene that plays multiple roles throughout *C. elegans* development as a regulator of the heterochronic, dauer and molting pathways (JEON *et al.* 1999; TENNESSEN *et al.* 2006; TENNESSEN *et al.* 2010; MONSALVE *et al.* 2011; MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014). Because *lin-42* was found to regulate the expression of heterochronic pri-miRNAs, and possibly miRNA and mRNA transcripts more globally, further investigations into the function of *lin-42* are vital to determine how *lin-42* regulates its miRNA targets and whether *lin-42* has a global role in gene regulation (MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014). Up until now, genetic analysis of *lin-42* has used hypomorphic alleles that all leave one isoform intact, making phenotypic and genetic analysis more difficult to interpret. A null *lin-42* allele is a key reagent that is missing to fully explore the role *lin-42* plays to regulate *C. elegans* development.

In chapter 2, I describe the work done to generate a *lin-42* null allele and analyze the null phenotype. Since extant *lin-42* alleles leave at least one isoform intact, I wanted to determine if null mutants have a more severe phenotype or any novel phenotypes. The *lin-42(0)* has precocious heterochronic defects, as well as molting defects, that occur at a higher penetrance then either upstream or

downstream hypomorphic alleles. This demonstrates that LIN-42A and LIN-42C isoforms, which are expressed in *lin-24(ok2385)* and *lin-42(n1089)* respectively, have some level of function. Rescue analysis of the *lin-42(0)* supports these results, since over-expression of *lin-42a*, *lin-42b*, or *lin-42c* rescue the null allele to some degree. However, because *lin-42a* and *lin-42b* strongly rescue the null, whereas lin-42c only mildly rescues the null, these results demonstrate that SYQ/LT domain containing isoforms are key to LIN-42 function.

Another important question I set out to investigate is the temporal expression profile of each transcript. *lin-42a* is expressed from a different promoter than *lin-42b* and *lin-42c*, and transcriptional reporters suggest that *lin-42a* is expressed at a different time than the other transcripts, however there is disagreement as to the order in which they are expressed. To more fully examine the expression patterns I performed stringent qPCR analysis and demonstrated that *lin42a* and *lin-42b* are expressed at the same time during each larval stage, despite being expressed from two different promoters.

lin-42 regulates the expression of the *let-7* family miRNA transcripts, including *mir-48*. *mir-48* is another interesting member of the heterochronic pathway that functions redundantly to promote the L2-to-L3 transition. To identify genes that interact genetically with *mir-48*, an over-expression suppressor screen was performed. In chapter 3, my goal was to identify suppressors of *mir-48* over-expression. This approach is expected to identify heterochronic genes, miR-48 target genes, and miRNA biogenesis and function genes. Five alleles of the heterochronic gene *lin-66* were identified, which validates the approach. Four

suppressors were linked to chromosomes and sequenced, however suppressors were yet to be identified. Information gathered from analyzing the *mir-48* suppressors will be useful for future investigations.

Figure 1: Heterochronic genes regulate seam cell development. A. The hypodermal seam cells are located on the lateral midline on either side of the worm. **B.** The adult cuticle contains ridges, called alae, along the length of the animal on each side. Terminal differentiation of hypodermal seam cells, executed at the L4-molt in wild type, is required for alae synthesis. **C.** Seam cell lineages of wild type and heterochronic mutants. Precocious mutants: *lin-42(lf)* and *hbl-1(lf)* result in precocious alae formation at the L3-molt (ABRAHANTE *et al.* 1998; ABRAHANTE *et al.* 2003; LIN *et al.* 2003); *lin-28(0)* mutants skip the L2 proliferative division and form alae at the L3-molt (AMBROS and HORVITZ 1984). Retarded mutants: *lin-4(0)* mutants reiterate the L1 stage division pattern throughout larval development and undergo an extra molt(s) (AMBROS 1989); *let-7(0)* mutants reiterate the L3 stage and form alae after a 5th molt (REINHART *et al.* 2000; VADLA *et al.* 2012); *mir-48(0) mir-241(0); mir-84(0)* mutants reiterate the L2 proliferative division (ABBOTT *et al.* 2005).

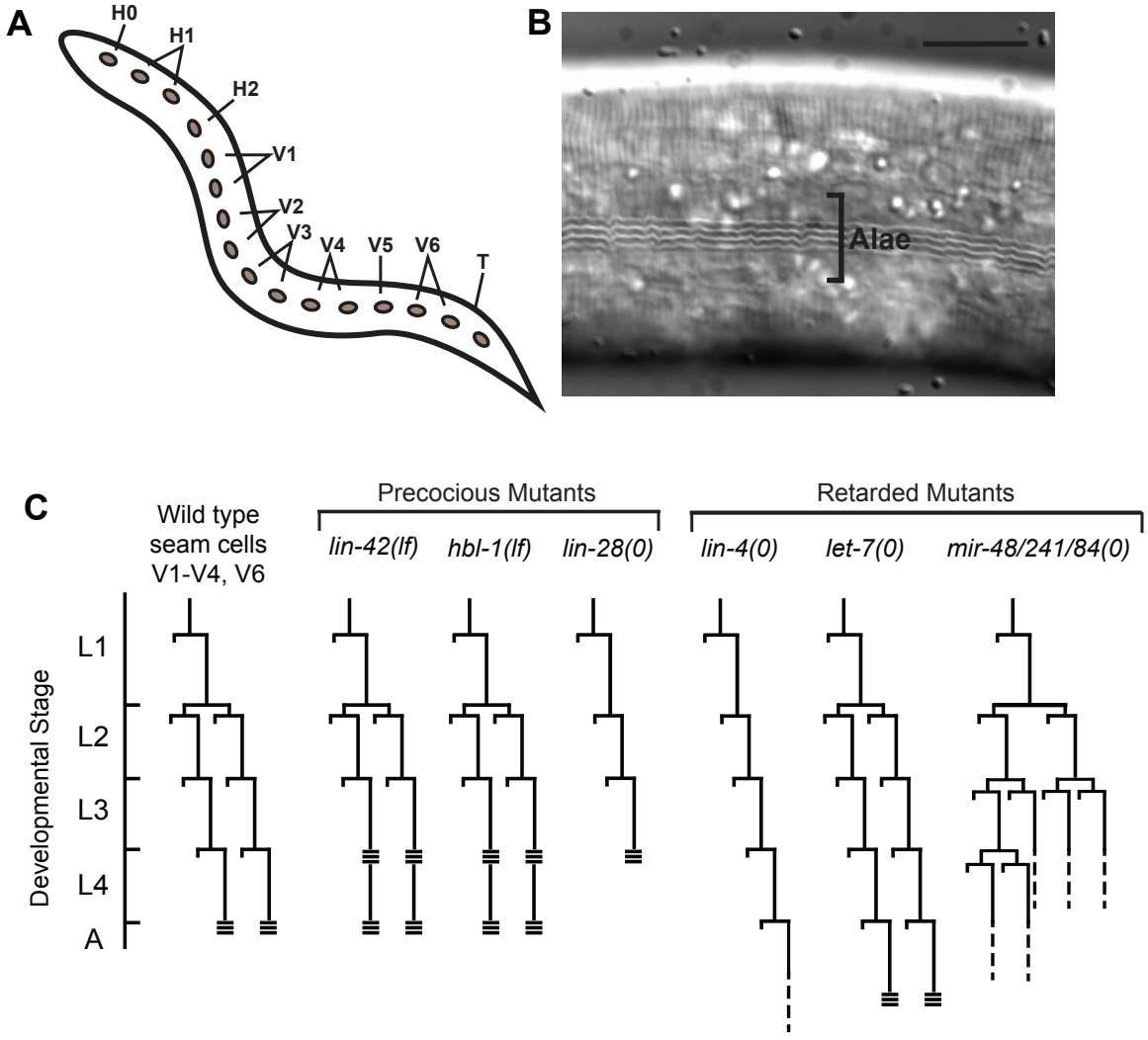

A
H0
H1
H2
V1
V2
V3
V4
V5
V6
T
B
Alae
C
Wild type
seam cells
V1-V4, V6
Precocious Mutants
lin-42(lf)
hbl-1(lf)
lin-28(0)
Retarded Mutants
lin-4(0)
let-7(0)
mir-48/241/84(0)
Developmental Stage
L1
L2
L3
L4
A

Figure 1

Figure 2: The *C. elegans* heterochronic pathway. Arrows indicate activation and blunt ends indicate suppression. Interactions are supported by genetic and biochemical data, except when indicated by a question mark, and are not necessarily direct. Note, *daf-12* interactions included in this pathway are with the ligand bound *daf-12* which promotes continuous development.

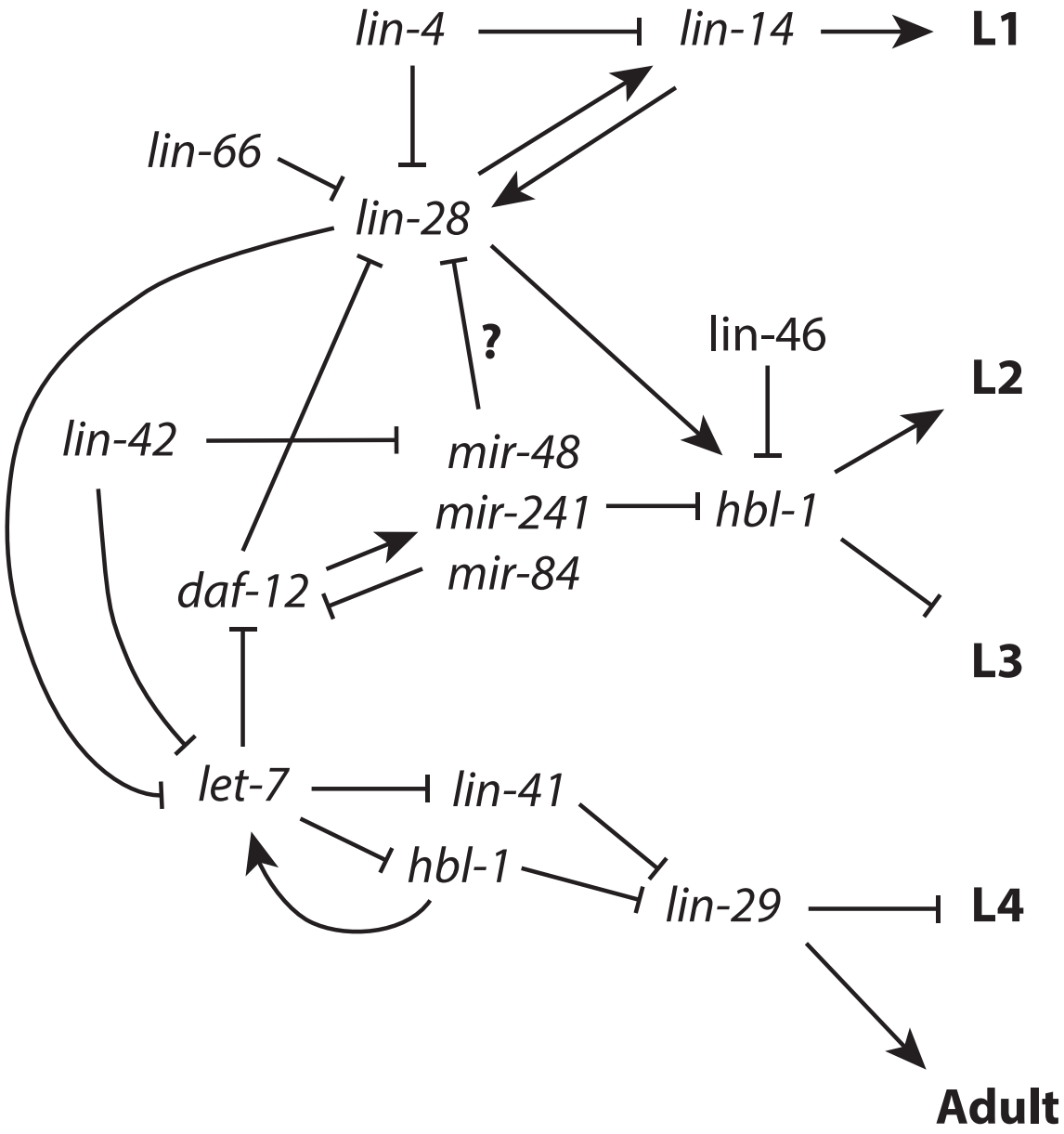

lin-4
lin-14
L1
lin-66
lin-28
?
lin-46
L2
lin-42
mir-48
mir-241
hbl-1
mir-84
daf-12
L3
let-7
lin-41
hbl-1
lin-29
L4
Adult

Figure 2

Figure 3: Highly processed miRNAs regulate post-transcriptional gene expression. A diagram of miRNA processing and function. miRNAs are transcribed by Pol II as long primary transcripts that are capped and poly-adenylated. These pri-miRNAs are then cleaved in the nucleus by the microprocessor to produce a pre-miRNA. The pre-miRNA is exported into the cytoplasm where it is further processed into a short ~21nt miRNA duplex. The duplex is incorporated into the RISC and the miRNA complement is release. The miRISC binds to the 3'UTR of an mRNA target through imperfect paring between the miRNA and mRNA and inhibits translation or targets mRNAs for degradation.

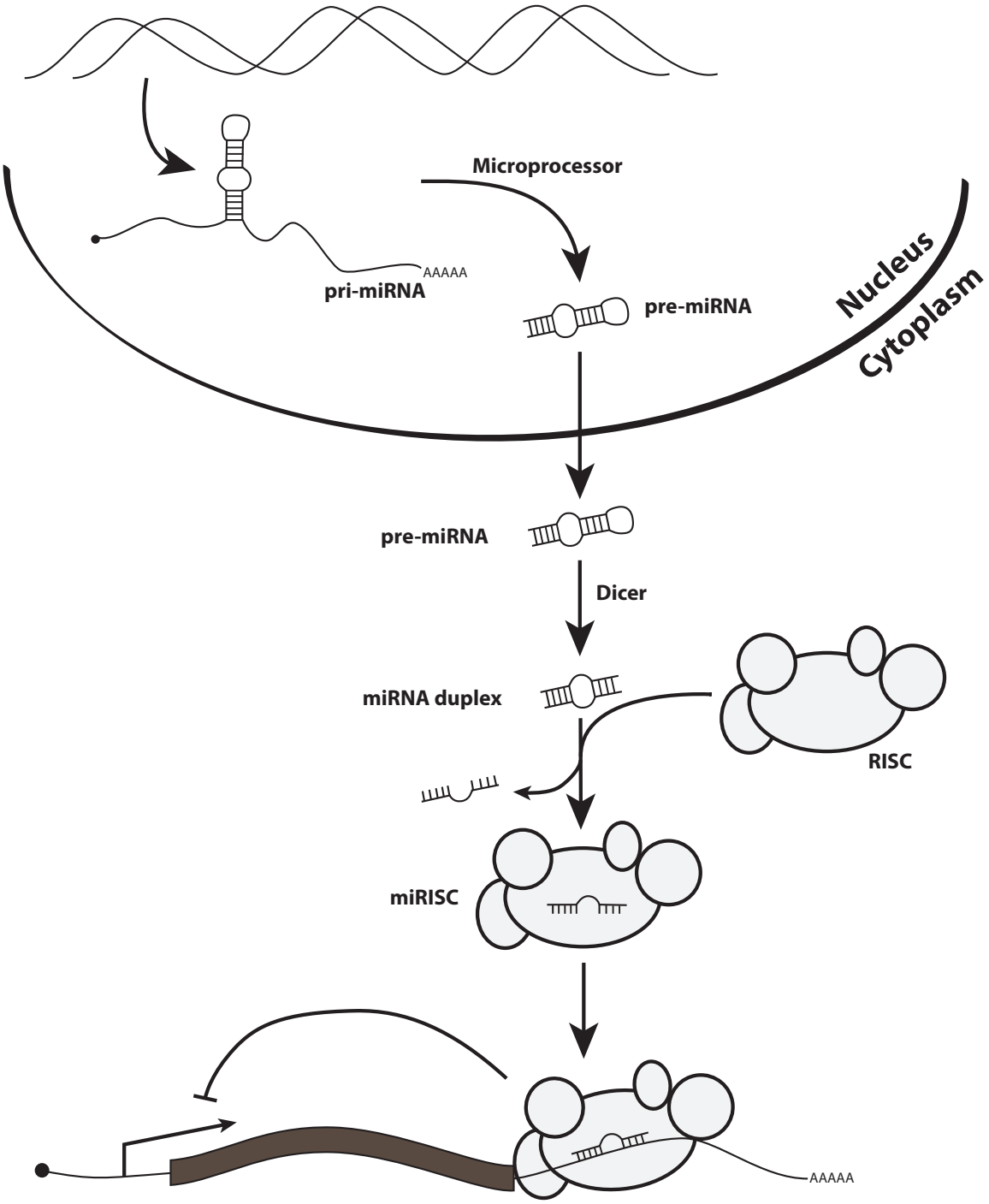

Microprocessor
AAAAA
pri-miRNA
pre-miRNA
Nucleus
Cytoplasm
pre-miRNA
Dicer
miRNA duplex
RISC
miRISC
AAAAA

Figure 3

Figure 4. The *lin-42* locus is complex and produces three transcripts. A. The *lin-42* locus produces three transcripts that encode three LIN-42 isoforms. Two transcripts, *lin-42a* and *lin-42c*, do not overlap. *lin-42a* encodes a protein containing the SYQ and LT domains, while *lin-42c* encodes a protein containing the PAS domains. *lin-42b* is a full length transcript that encodes SYQ, LT, and PAS domains. *n1089* deletes the upstream region and leaves *lin-42a* intact, while *ok2385* deletes the downstream region and leaves *lin-42c* intact. *ox461,* analyzed in this thesis, deletes the entire coding region. **B.** A representation of *lin-42* mRNA expression as reported in Jeon *et al.* Primers used detected *lin-42b* and *lin-42c* (JEON *et al.* 1999).

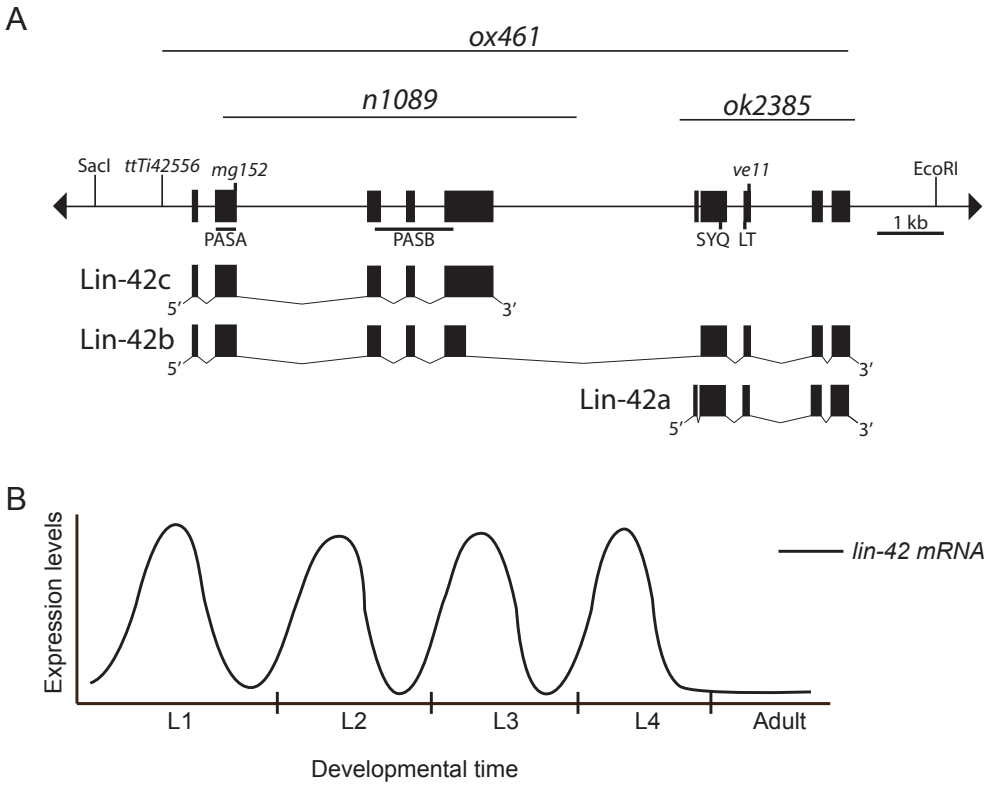

A
ox461
n1089
ok2385
SacI
ttTi42556
mg152
ve11
EcoRI
PASA
PASB
SYQ
LT
1 kb
Lin-42c
5′
3′
Lin-42b
5′
3′
Lin-42a
5′
3′
B
Expression levels
lin-42 mRNA
L1
L2
L3
L4
Adult
Developmental time

Figure 4

Figure 5. The *Drosophila melanogaster* circadian clock is regulated by a negative feedback loop. Shown is a simplified version of the *period/timeless* feedback loop that regulates the Drosophila circadian clock. See text for details. PERIOD (PER), TIMELESS (TIM), casein kinase 2 (CK2), DOUBLETIME (DBT), CRYPTOCHROME (CRY), CLOCK (CLK), CYCLE (CYC).

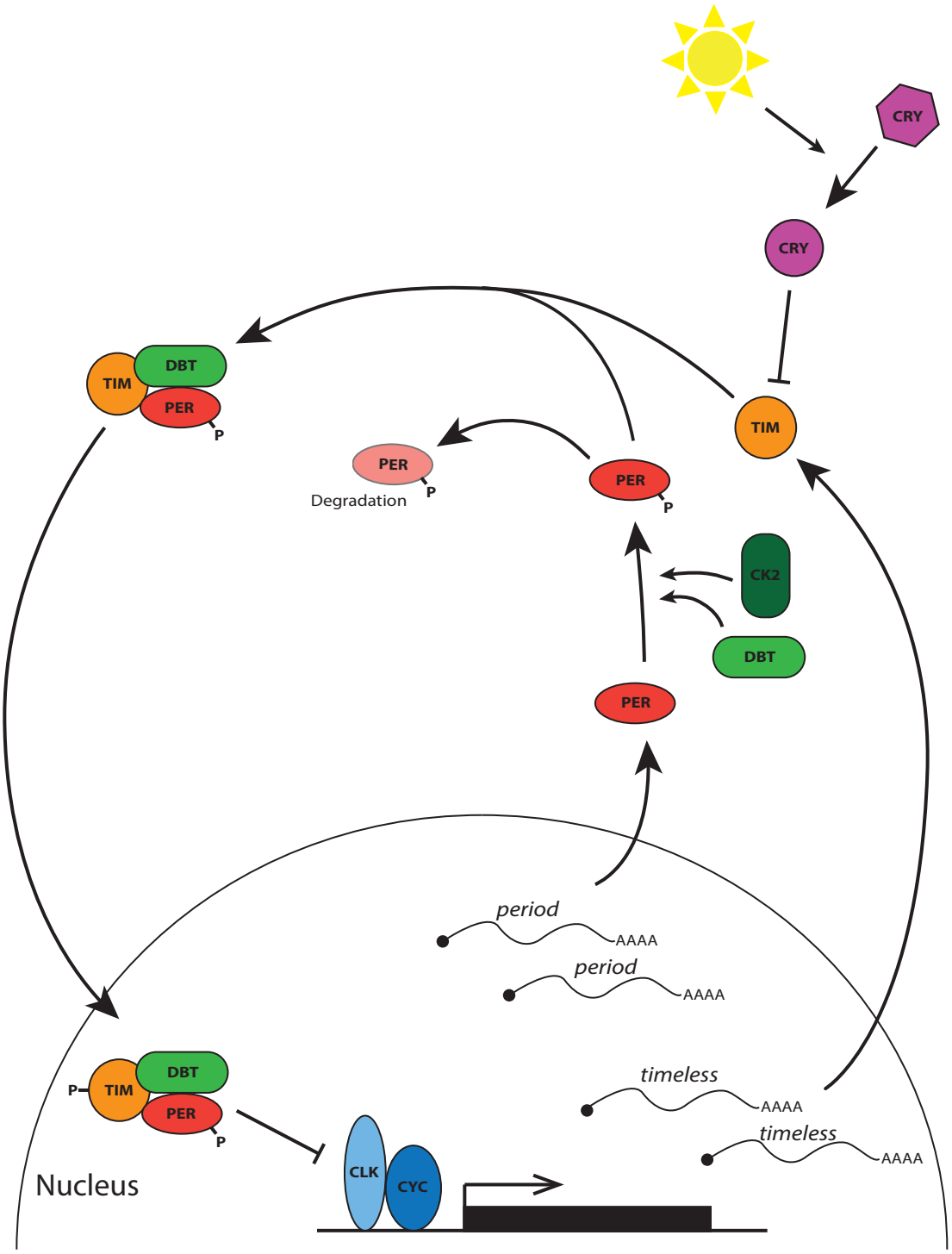

CRY
CRY
TIM
DBT
PER
P
PER
Degradation
P
PER
P
TIM
CK2
DBT
PER
period
AAAA
period
AAAA
timeless
AAAA
timeless
AAAA
P
TIM
DBT
PER
P
CLK
CYC
Nucleus

Figure 5

Analysis of a *lin-42*/*period* null allele implicates all three isoforms in regulation of *C. elegans* molting and developmental timing

Theresa L.B. Edelman[*], Katherine A. McCulloch[*,1], Angela Barr[*], Christian Frøkjær-Jensen[§], Erik M. Jorgensen[§] and Ann E. Rougvie[*]

[*]Department of Genetics, Cell Biology, and Development
University of Minnesota
Minneapolis, MN 55454

[1]Present address: Division of Biological Sciences, Section of Neurobiology
University of California, San Diego
La Jolla, CA 92093

[§]Department of Biology, Howard Hughes Medical Institute, University of Utah
Salt Lake City, Utah 84112

Corresponding Author:
Ann Rougvie
University of Minnesota
6-160 Jackson Hall
321 Church St. SE
Minneapolis, MN 55108
(612) 624-4708
rougvie@umn.edu

Keywords: *C. elegans*, *lin-42*, heterochrony

Running title: Analysis of a *lin-42* null allele

SUMMARY

The *C. elegans* heterochronic gene pathway regulates the relative timing of events during post-embryonic development. *lin-42*, the worm homolog of the circadian clock gene *period*, is a critical element of this pathway. *lin-42* function has been defined by a set of hypomorphic alleles that cause precocious phenotypes, in which later developmental events, such as the terminal differentiation of hypodermal cells, occur too early. A subset of alleles also reveals a significant role for *lin-42* in molting; larval stages are lengthened and ecdysis often fails in these mutant animals. *lin-42* is a complex locus, encoding overlapping and non-overlapping isoforms. Although existing alleles that affect subsets of isoforms have illuminated important and distinct roles for this gene in developmental timing, molting, and the decision to enter the alternative dauer state, it is essential to have a null allele to understand all of the roles of *lin-42* and its individual isoforms. To remedy this problem, and discover the null phenotype, we engineered an allele that deletes the entire *lin-42* protein coding region. *lin-42* null mutants are homozygous viable, but have more severe phenotypes than observed in previously characterized hypomorphic alleles. We also provide additional evidence for this conclusion by using the null allele as a base for reintroducing different isoforms, showing that each isoform can provide heterochronic and molting pathway activities. Transcript levels of the non-overlapping isoforms appear to be under coordinate temporal regulation, despite being driven by independent promoters. The *lin-42* null allele will continue to be

an important tool for dissecting the functions of *lin-42* in molting and developmental timing.

INTRODUCTION

C. elegans is a powerful system for studies of developmental time control because of the invariance and precise temporal orchestration with which its cell division patterns are programmed as development proceeds (SULSTON and HORVITZ 1977; KIMBLE and HIRSH 1979; SULSTON 1983). *C. elegans* develops through embryogenesis and four larval stages, each with a characteristic set of cell divisions and morphogenetic events, prior to becoming reproductively competent adults. Genes that provide temporal cues necessary to specify the appropriate sequence and timing of these postembryonic cell divisions have been identified and termed heterochronic genes (ROUGVIE and MOSS 2013). When heterochronic genes are mutated, specific larval programs are skipped or reiterated causing subsequent events to occur too early or to be delayed, respectively. For example, mutations in the heterochronic gene *lin-42* cause a precocious phenotype, as demonstrated by terminal differentiation of hypodermal cells occurring one stage too early, during the L3, rather than L4, stage (ABRAHANTE *et al.* 1998; JEON *et al.* 1999; TENNESSEN *et al.* 2006). In wild-type animals, *lin-42* temporally restricts this differentiation event, at least in part, by acting as a negative transcriptional regulator of certain miRNA genes, including *let-7* family miRNAs that have prominent roles in the heterochronic gene pathway

(REINHART *et al.* 2000; ABBOTT *et al.* 2005; LI *et al.* 2005; MCCULLOCH and ROUGVIE 2014; PERALES *et al.* 2014; VAN WYNSBERGHE *et al.* 2014).

lin-42 functions more broadly than control of temporal cell fate in the hypodermis. One additional role is in mediating responses to environmental cues. *lin-42* mutants are hypersensitive to entry into the dauer larva stage, a reversible diapause state that affords survival when growth conditions deteriorate, demonstrating that a wild-type function of *lin-42* is to inhibit dauer formation (TENNESSEN *et al.* 2010). Another *lin-42* function is in the molting pathway: certain *lin-42* alleles cause variable delays in the molting cycle and a failure of ecdysis, leading to the proposal that *lin-42* may coordinate these activities with hypodermal development (MONSALVE *et al.* 2011).

lin-42 is a particularly intriguing member of the heterochronic gene pathway because it encodes the worm homolog of PERIOD (JEON *et al.* 1999; TENNESSEN *et al.* 2006), a core component of the circadian clock in flies and vertebrates, thereby providing a link to another biological timing mechanism. LIN-42 and PERIOD share several regions of homology, including the hallmark PAS protein interaction domains, and smaller SYQ and LT domains that have been shown to interact with circadian clock proteins, including CLOCK and Casein Kinase Iε, in flies and mammals (CHANG and REPPERT 2003; LEE *et al.* 2004a; SUN *et al.* 2010). In flies, the circadian clock is regulated by a transcriptional negative feedback loop between PERIOD/TIMELESS and CLOCK/CYCLE. Activity of a PERIOD/TIMELESS complex mediates repression of CLOCK/CYCLE transcriptional activity to drive cyclical gene expression.

Interestingly, while PERIOD contains the PAS and SYQ/LT domains in a single protein, the *lin-42* locus encodes multiple protein isoforms, including non-overlapping proteins that separate these domains. LIN-42B is the longest protein, containing maximal homology to PERIOD proteins as it contains all of the conserved domains (Figure 1A). In contrast, LIN-42A and LIN-42C are non-overlapping and expressed from distinct promoters (TENNESSEN *et al.* 2010). LIN-42A contains the SYQ and LT domains, while LIN-42C contains the PAS domains. These two non-overlapping isoforms provide the opportunity to investigate the function of the PAS and SYQ/LT domains separately. RNA-seq data compiled on WormBase provides strong support for expression of *lin-42*a and *lin-42*b, while support for *lin-42*c comes from 3'RACE (Rapid amplification of cDNA ends) sequence tags (WormBase release WS252). Moreover, a *lin-42* allele that deletes the PAS domain can be rescued by a genomic fragment encompassing the *lin-42c* transcription unit, indicating that *lin-42c* expression can provide function, but this fragment fails to efficiently rescue a premature stop in the LT domain (TENNESSEN *et al.* 2006). In contrast, *lin-42a* expression can rescue mutations that disrupt either *lin-42a* or *lin-42c*, such as *n1089*, *ve11* and *ok2385* (TENNESSEN *et al.* 2006; MONSALVE *et al.* 2011) (Figure 1A). These results suggest that the SYQ and LT domain containing isoform is key to LIN-42 function, whereas the PAS-containing isoforms may play a more regulatory role. Supporting this argument is the observation that *n1089*, the PAS domain deletion, and *mg152*, a premature stop codon predicted to eliminate expression of most of the PAS domain, leave Lin-42A intact yet cause a heterochronic

phenotype; Lin-42A only rescues these alleles when presumably overexpressed from a multicopy array.

A tool long missing from the *lin-42* arsenal is a null allele. A deletion allele that eliminates all three isoforms is needed to reveal the *lin-42* null phenotype and allow functional dissection of individual isoform contributions. To remedy this problem and enable further probing of the *lin-42* mechanisms of action, we generated and characterized a *lin-42* null allele using MosDel technology (FRØKJÆR-JENSEN *et al.* 2010) to delete the entire *lin-42* coding region. *lin-42(null)* mutants are viable, but have more severe molting and developmental timing defects than do previously characterized mutants in which one isoform is left intact.

MATERIALS & METHODS

Strains and Nematode Maintenance

C. elegans were grown and maintained at 20˚ on NGM plates seeded with *E. coli* OP50 as previously described (BRENNER 1974). Full genotypes of strains used in this work are listed in Table 1.

Generation of a *lin-42* null allele

Mos1-mediated deletion of the *lin-42* locus was performed essentially as described (FRØKJÆR-JENSEN *et al.* 2010) using a *Mos1* insertion residing 440 bp upstream of the first *lin-42* exon that was obtained from the nemaGENETAG consortium (ttTi42556; (VALLIN *et al.* 2012) (Figure 1A). The deletion template contained an ~1.5kb homology fragment 5' to the *Mos1* insertion site, and an

~1.4 kb 3' homology fragment from the *lin-42a/b* 3'UTR (Figure 1B). The fragments were amplified from wild-type genomic DNA using the following primers that also contain attB sites for Gateway Cloning: 5' homology fragment, AB25 5'-ggggacagcagctttcttgtacaaagtggaacctaaaactcctcggt-3'/AB26 5'-ggggacaactttgtataataaagttgacgaatcatgttccctgt-3'; 3' homology fragment, AB27 5'-ggggacaactttgtatagaaaagttggactgaaaattggtgtatgaaca-3'/AB28 5'-ggggactgctttttgtacaaacttgccgtcttcccgaaaactt-3'. The *lin-42* locus flanking fragments were assembled into pCFJ66, flanking *Punc-122::gfp* and *C. briggsae unc-119*, to yield pAB8.

To generate transgenic animals, a mixture of pALB8 (50 ng/µl), pJL43.1 (P*glh-2::transposase* 50ng/µl), pGH8 (P*rab-3::mCherry* 10 ng/µl), P*myo-2::tdTomato* (2.5 ng/µl), and pCFJ104 (P*myo-3::mCherry* 5 ng/µl) was injected into RG1514 *ttTi42556; unc-119(ed3)*. Broods of animals containing extrachromosomal arrays of these plasmids were screened for *unc-119*(+) animals that had lost the extra-chromosomal array and thus lacked red fluorescence. These candidate deletion animals were allowed to reproduce and then were genotyped by PCR. DNA sequencing confirmed the appropriate junction fragments using the following primer sets: 5' junction: AB45 5'-gtaccctcaagggtcctcct-3'/AB46 5'-cccagactttgcatcgaaat-3'. 3' junction: AB35 5'-cgaaaatttcaaaaagctcgt-3'/AB37 5'-caattcatcccggtttctgt-3'. In each set, one primer lies within the P*unc-122::gfp*/*C. briggsae unc-119(+)* insertion, while the other primer resides in the genome outside of the homology arms. Two independent *lin-42(0)* strains were identified, EG15911 *lin-42(ox461); unc-119(ed3)* and

EG15910 *lin-42(ox460); unc-119(ed3).* Each strain was outcrossed three times to generate RG1590 *lin-42(ox461)* and RG1591 *lin-42(ox460)*, respectively.

Generation of a *lin-42b* minigene and transgenic animals

pCP2 is a *lin-42*b minigene made by subcloning a 5.5 kb Lin-42b/c promoter fragment from pHG82 (TENNESSEN *et al.* 2010) onto a *lin-42*b cDNA with *gfp* coding sequence added just prior to the stop codon and containing the *unc-54* 3'UTR. pCP2 [P*lin-42b/c::lin-42b::gfp::unc-54*] (5 ng/µl) and *str-1::gfp* (100 ng/µl) were injected into N2 hermaphrodites to generate transgenic lines bearing extrachromosomal arrays which were crossed into *lin-42(ox461)* for analysis. The Lin-42c expressing arrays were made similarly by injecting genomic clone pMJ13 (JEON *et al.* 1999) at 5 ng/µl together with *str-1::gfp* (100 ng/µl) or *sur-5::gfp* (75 ng/µl) as a transformation marker.

Phenotypic analysis

lin-42(ox461) animals are egg-laying defective, causing adult hermaphrodites to die following the internal hatching of eggs (the so-called "bag-of-worms" phenotype). Because *lin-42* mutant animals are sensitive to growth conditions and their heterochronic phenotypes can be suppressed by starvation (ABRAHANTE *et al.* 1998), eggs were isolated from *lin-42(ox461)* animals by hypochlorite treatment, washed in M9 buffer and plated at low density onto seeded plates so that they would hatch into optimal growth conditions. To monitor postembryonic development starting at the L1 molt, newly hatched larvae were monitored using a Kramer FBS10 microscope until they stopped pumping and entered the first molt, at which time they were singly-picked to seeded 30

mm plates. *lin-42(0)* animals were monitored every 2 hrs for pumping and ecdysis for 5 days, whereas wild-type control animals were checked hourly until they reached adulthood. Similarly, to quantify the proportion of larval arrest in *lin-42* mutant populations, L1 larvae were singly picked to seeded plates and monitored daily for growth. Animals that failed to reach adulthood by eight days were classified as arrested. Heterochronic phenotypes were scored at the appropriate developmental stage in animals hatched at low density on seeded plates.

qPCR analysis

Synchronized populations of worms were generated by hypochlorite treatment of gravid adults to isolate eggs, which were then hatched overnight in M9 buffer. Starved L1s were then plated at a density of 5,000 to 10,000 animals per 10 cm plate. Animals were plated twice, 12 hours apart, to stagger collection times. At each timepoint, animals were washed off plates with M9, pelleted and flash frozen in liquid nitrogen. Biological replicates were derived from independent starting populations and performed on different days.

RNA was extracted using Trizol (Life Technologies) and 425-600 µm glass beads (Sigma) to aid in disrupting the cuticle. Total RNA (5 µg) was treated with DNAse I (Ambion Turbo DNAse kit) to remove any genomic DNA, and then 1 µg was reverse transcribed with 10 U Roche Transcriptor using random primers (0.5 µg; Promega Corp.). qPCR reactions with 12.5 ng of reverse transcribed sample were run in triplicate as directed (Roche) on an Eppendorf Realplex Thermocycler using 96-well plates. The following Roche Universal Probe Library

(UPL) probe and primers were used for each assay: *lin-42*a, probe #10/TE55 5'-gtacgatcttgcagagccagt-3'/TE56 5'-gaggcttgagtgatggtggt-3'; *lin-42*b, probe #10/TE65 5'-ctttcgaggatgagctgagaa-3'/TE68 5'-ctgatccttgaggcttgagtg; *lin-42*c, probe #146/TE59 5'-aattagacggcgcgagagt-3'/TE60 5'-gccagcatgtgtacttttgc-3'; *mlt-10*, probe #115/TE61 5'-ggcgttgaagaagttcaagag-3'/TE62 5'-cggaacttttcggcttcag-3'; *ama-1*, probe #165/TE72 5'-ggatggaatgtgggttgaga-3'/TE73 5'-gttgtcggtgaggtccattc. Data was collected and analyzed using Realplex 2.0 software. Each time-point was normalized to an *ama-1* internal control, and each plate was normalized to the 0 hr or 6 hr sample within each time-course as indicated. Data was analyzed using the ΔΔCt method in Microsoft Excel (LIVAK and SCHMITTGEN 2001). Reactions performed on samples where the reverse transcription step was omitted failed to result in detectable signals (data not shown).

Introduction of LIN-42(W258E) mutation using CRISPR-Cas9

The W258E mutation was introduced using the CRISPR-Cas9 gene editing approach (DICKINSON *et al.* 2013; PAIX *et al.* 2014). The pTBE21 sgRNA was engineered as described (DICKINSON *et al.* 2013). Briefly, the targeting sequence was introduced into pDD162 using the Q5 Site-Directed Mutagenesis Kit (NEB), forward primer TE85 5'-gagcgcctacgtgaatccaGTTTTAGAGCTAGAAATAGCAAGT-3' (targeting sequence is in lowercase), and reverse primer TE86 5'-CAAGACATCTCGCAATAGG-3'. Wild-type animals were injected with sgRNA pTBE21 (50 ng/µl), a repair oligo containing the mutated codon (which also

disrupted the PAM site) TE89 [5'-acggatcgatcctccgatgccagaccgaatggagcgcctacgtgaatccaGAGacccgaaaaatggagctagtcgtcgccagacaccgtatctgctccctt-3'] (50 ng/µl), and pCFJ104 [Pmyo-2::mcherry] (4 ng/µl). Transgenic F1 animals and their siblings were screened by amplifying the region that contained the mutation using TE90 5'-aggcttcagcgtaggtttca-3' and TE91 5'-aggcgcagagataaaagtgc-3', and digesting the product with Hpy188III (NEB). Mutants were sequenced to confirm they contained the correct sequence mutation.

Reagent availability

Strains and plasmids used in this work are available upon request.

RESULTS & DISCUSSION

Generation of a *lin-42* null allele

Genetic analysis of *lin-42* has been complicated by the presence of non-overlapping isoforms and the lack of a null allele (TENNESSEN *et al.* 2006). To remedy this situation, the Mos1-mediated transposon system (MosDel; (FRØKJÆR-JENSEN *et al.* 2010)) was used to delete the *lin-42* coding region. The *Mos1* insertion ttTi42556 (VALLIN *et al.* 2012), 440 bp upstream from the ATG start site (Figure 1A), was targeted by the Mos1 transposase to generate a double-strand break that was then repaired from a template plasmid containing homology arms outside the *lin-42* coding region (Figure 1B). Two *lin-42(null)* alleles were isolated and confirmed by PCR and sequencing: *lin-42(ox460)* and *lin-42(ox461).* The two deletions are identical, removing 10,226 bp of genomic

DNA spanning the *lin-42* coding region and replacing it with *C. briggsae unc-119(+)* and a *Punc-122::GFP* transgene (Figure 1B). Following three outcrosses to wild-type, the strains appeared indistinguishable. *lin-42(ox461)* was chosen for detailed analysis and is hereafter referred to as *lin-42(0)* for simplicity.

lin-42(0) is homozygous viable but causes highly penetrant molting defects

lin-42 is expressed in late stage embryos (JEON *et al.* 1999) raising the possibility that a null allele could cause embryonic lethality. However, *lin-42(0)* animals can be maintained in a homozygous state, and the vast majority of embryos hatch (98%; n=52), indicating that *lin-42* is not essential during embryogenesis.

To further characterize the *lin-42(0)* phenotype, we compared it to the phenotypes of two representative hypomorphic alleles, *lin-42(n1089)* and *lin-42(ok2385),* which each delete one of the two non-overlapping transcription units. *lin-42(n1089)* deletes the majority of *lin-42c* including the PAS domain (and the corresponding portion of *lin-42b*), but leaves *lin-42a* intact (TENNESSEN *et al.* 2006). Conversely, *lin-42(ok2385)* removes *lin-42a* and the SYQ and LT domains, while leaving *lin-42c* intact and truncating *lin-42b* (MONSALVE *et al.* 2011). Of these two alleles, *lin-42(n1089)* causes the mildest phenotype, largely restricted to heterochronic defects in the hypodermis, while *lin-42(ok2385)* is notably more severe, with animals also exhibiting larval arrest and molting defects(MONSALVE *et al.* 2011).

lin-42(0) mutants exhibit larval arrest, developmental delays, and molting defects. When the development of individual animals was tracked, 75% of *lin-42(0)* animals failed to reach adulthood; they arrested at an early larval stage (Figure 2A). This phenotype is similar to that of *lin-42(ok2385)* animals, but has a statistically higher penetrance, indicating the upstream region contributes some function (P=0.003, Fisher's Exact Test). Strikingly, the larval arrest phenotype is observed in *lin-42(0)* and *lin-42(ok2385)*, but not *lin-42(n1089)* mutants (Figure 2A) as previously noted (MONSALVE *et al.* 2011), indicating that the PAS domain is dispensable for larval stage progression when *lin-42*a is intact. Arrested animals were often seen trapped in the previous stage cuticle indicating an inability to ecdyse (Figure 2B), suggesting that molting errors are a causal factor in the larval arrest phenotype.

The *lin-42(0)* molting defect appears to result in a growth arrest, perhaps as a consequence of reduced nutrition, rather than a complete developmental arrest in some tissues. Arrested animals exhibited developmental progression based on expression of the motor neuron marker *del-1::gfp* (WINNIER *et al.* 1999). In wild-type animals, *del-1::gfp* expression comes on in the VB motor neurons in the early L2, but is repressed in the VA motor neurons, which are establishing synaptic inputs during the L2 stage. From the late L2 to the adult, *del-1::gfp* expression is progressively established in the VAs, appearing in an anterior-to-posterior wave. In arrested *lin-42(0)* animals, *del-1::gfp* expression was initially observed in the 11VB motor neurons, and was progressively activated in the VAs such that by day five post-hatching, 16/20 arrested animals exhibited expression

in all VA and VB motor neurons, with the remaining four lacking expression in one or two posterior VAs. In addition, vulval and gonadal development was observed in a small number of arrested animals (Figure 2B, right inset), indicating that these tissues could also developmentally progress. However, the rarity of these animals has so far precluded a thorough analysis of this phenotype.

To understand better the temporal dynamics of postembryonic development in *lin-42(0)* mutants, we tracked individual animals beginning at the L1 molt (see Materials and Methods). Animals were hatched in the presence of food and synchronized by monitoring entry into L1 lethargus (Figure 2C), a characteristic behavior lasting approximately 2 hrs at the end of each larval stage, in which the worms cease movement and pharyngeal pumping as a new cuticle is synthesized prior to ecdysis. When compared to wild type, L1 lethargus of *lin-42(0)* mutants was lengthened by about 1 hr, but more than 90% of animals examined completed ecdysis as evidenced by shed cuticles (Figure 2C and data not shown, n=62). The developmental delays and molting defects became more severe as animals progressed through later larval stages; the population became increasingly more asynchronous and a large percentage of animals failed to shed cuticles. Most *lin-42(0)* animals that bypassed early larval arrest executed the third larval molt about 2.5 days after hatching. The fourth larval stage exhibited the longest delay, with animals requiring at least 24 hrs after the L3 molt before entering the final period of lethargus (data not shown). These phenotypes mimic those previously described for *lin-42(ok2385)* (MONSALVE *et al.* 2011).

To compare the relative developmental delay caused by the different *lin-42* alleles, the time required to reach adulthood was assessed for animals that escaped larval arrest. *lin-42(0)* mutants had the most severe delay with a majority of animals requiring five days to mature (Figure 2D) while most *lin-42(ok2385)* mutants became adults on day four. A few *lin-42(0)* and *lin-42(ok2385)* animals required 6 to 8 days to reach adulthood; these animals were observed to temporarily arrest in an early larval stage before eventually escaping from their cuticle and continuing through development. In contrast, *lin-42(n1089)* animals, which have a deletion of the 5' end of the locus that encodes the PAS domain, exhibit only a mild developmental delay, with ~70% of animals maturing on day 3, similar to wild-type animals. This analysis reinforces the idea that the SYQ/LT-containing isoforms play a major role in control of molting; however, the enhanced severity of *lin-42(0)* relative to the hypomorphic alleles indicates a contribution from the conserved PAS domains encoded by *lin-42*b/c transcripts.

The three *lin-42* alleles examined also form an allelic series with respect to an egg-laying defective (Egl) phenotype: *ox461* > *ok2385* > *n1089* > wild-type (Table 2). All *lin-42(0)* animals that became fertile adults were egg-laying defective (Egl), compared to 78% of *lin-42(ok2385)* and 18% of *lin-42(n1089)* mutants. The proportion of egg-laying competent animals presumably mirrors the relative ability of animals to shed the L4 stage cuticle, thereby allowing eggs to exit through the vulva. A corresponding decrease in average brood size was observed in each strain, likely due to internal hatching of embryos, which

eventually leads to death of the hermaphrodite parent, as well as a minor uncharacterized reduction in fertility (Table 2).

lin-42(0) heterochronic defects are more severe than those caused by hypomorphic alleles

lin-42 was defined by mutations with a precocious heterochronic phenotype in which hypodermal seam cells terminally differentiate and produce a characteristic adult cuticle with alae one stage too early at the L3 molt (L3m) (ABRAHANTE *et al.* 1998; JEON *et al.* 1999) (Figure 3A,B). In order to score this phenotype in a population of *lin-42(ox461)* animals, which develop asynchronously through the larval stages, it was necessary to reliably identify L3 stage animals. Tracking of individual animals revealed that in *lin-42(0)* animals that did not arrest, vulval divisions occurred during the L3 stage as in the wild type, and vulval morphogenesis had proceeded to the invagination stage by the L3m, appearing at least superficially wild-type (Figure 3C). Vulval morphology was thus used to identify L3m *lin-42(0)* animals from roughly staged populations of animals.

Partial or complete precocious alae was observed in all *lin-42(0)* animals at the L3m. This defect is significantly stronger than that of either *lin-42(n1089)* or *lin-42(ok2385)* mutants ($P<0.0001$, Fisher's Exact Test), with a larger proportion of animals forming complete alae at the L3m (Figure 3B,C). The observation that *lin-42(0)* is more severe than either *lin-42(n1089)* or *lin-42(ok2385)* suggests the

isoform present in each hypomorphic allele contributes some level of function, even though they are distinct proteins.

In addition to defects in the timing of seam cell terminal differentiation, several heterochronic mutants have altered seam cell number due to the skipping or reiteration of the proliferative division normally observed at the beginning of the second larval stage in wild-type animals (Figure 3A), when seam cell number is increased from 10 to 16 on each lateral side (SULSTON and HORVITZ 1977). Prior results that link *lin-42* activity to control of the proliferative division prompted assessment of seam cell number in *lin-42(0)* animals. Notably, *lin-42* negatively regulates expression of *mir-48* (MCCULLOCH and ROUGVIE 2014; PERALES et al. 2014; VAN WYNSBERGHE et al. 2014), a miRNA that when over expressed results in reduced seam cell number (LI *et al.* 2005). In addition, when grown at the permissive temperature of 15˚, *lin-14(n179ts); lin-42(n1089)* animals skip the L2 stage proliferative division, a phenotype not observed in either single mutant (LIU 1990). We used the P*elt-5::his-24::mCherry* (LIU *et al.* 2009) hypodermal marker as an aid in scoring seam cell number. *lin-42(0)*; P*elt-5::his-24::mCherry* animals hatch with a wild-type number of seam cells on each lateral side (9.95±0.22) that proliferate appropriately, resulting in adults with the full complement of 16±0.56 (Table 2). Thus, *lin-42* activity is not essential for seam cell specification or proliferation.

P*elt-5::his-24::mCherry* was used to monitor seam cell number rather than *scm::gfp*, a standard tool used in the assessment of seam cell number and morphology in heterochronic mutants (ANTEBI *et al.* 2000; LEHRBACH *et al.* 2009;

HADA *et al.* 2010; CHU *et al.* 2014), because the latter reporter strongly suppressed the *lin-42(0)* precocious alae phenotype (Table 2). The observed suppression was not specific to the integrated *scm::gfp* transgene used (*wIs51* [*scm::gfp* + *unc-119*(+)]); similar suppression was also observed in animals with an integrated *mlt-10::gfp* reporter (*mgIs49 [Pmlt-10::gfp-pest + Pttx-3::gfp]*) *(HAYES et al. 2006)*. One commonality between these transgene reporters is that *Pmlt-10::gfp* and *scm::gfp* are both expressed in the hypodermis, although primarily in distinct cells. *Pmlt-10::gfp* has a cyclical expression pattern in hyp7, the main body hypodermal syncytium, peaking ~4 hours prior to ecdysis (MELI *et al.* 2010), whereas *scm::gfp* is predominantly expressed in the hypodermal seam cells throughout the larval stages (TERNS *et al.* 1997; MOHLER *et al.* 1998; KOPPEN *et al.* 2001). One possibility is that given the multicopy nature of the transgene arrays, the *scm::gfp* and *Pmlt-10::gfp* promoters may titrate out a factor whose precocious expression is required for alae synthesis at the L3 molt. Although the mechanism of suppression is not clear, caution should be used when scoring heterochronic phenotypes in mutants containing these reporters. Even though *Pelt-5::his-24::mCherry* is expressed in seam and hyp7 cells during larval stages and maintained in adult seam cells, suppression was not observed in animals with this reporter (Table 2).

Each isoform can provide some *lin-42* activity

Previous studies showed that expression of the SYQ/LT domain encoding isoform Lin-42a from multicopy arrays can rescue both *lin-42(ve11)*, a premature

stop in *lin-42a* (Figure 1A), as well as the *lin-42*c deletion *lin-42(n1089),* whereas expression of *lin-42c* only rescues *lin-42(n1089)* efficiently (TENNESSEN et al. 2006). The availability of a null allele afforded us the opportunity to more fully test the ability of each isoform to rescue the heterochronic, larval arrest, and egg-laying phenotypes (JEON *et al.* 1999; TENNESSEN *et al.* 2006). Transgenic expression of *lin-42a* strongly rescued both the heterochronic and larval arrest phenotypes of *lin-42(0)* animals (Table 2), indicating that this isoform is sufficient to control these processes. However, there is likely to be a contribution from overexpression at play, because the *n1089* and *mg152* alleles result in heterochronic defects, demonstrating that endogenous *lin-42a*, although sufficient for allowing larval progression, is insufficient to regulate the timing of hypodermal fates.

To test the rescuing ability of *lin-42b*, it was first necessary to construct a *lin-42b* minigene by fusing the *lin-42b/c* promoter to a *lin-42b* cDNA with *gfp* inserted prior to the stop codon. Use of a *lin-42b* cDNA omits the 3.6 kb intron between exons 5 and 6 which harbors the *lin-42a* promoter and thus prevents independent expression of *lin-42a* (Figure 1A). P*lin-42b/c::lin-42b(cDNA)::gfp::unc-54 3'UTR* expression rescued both the heterochronic and molting defects of *lin-42(0)* animals at least as well as did *lin-42a* expression (Table 2). Moreover, this result indicates that the essential spatial and temporal enhancers that drive *lin-42b/c* expression are located upstream of the *lin-42* coding region and/or reside in exons.

The ability of *lin-42*c expression to rescue the null was less robust than was observed for either *lin-42*b or *lin-42*a (Table 2). *lin-42*c only weakly rescued the precocious alae phenotype, and four of five lines tested showed little, if any, rescue of the larval arrest phenotype. As in non-transgenic null mutants, the arrested animals were often observed to be trapped in a previous stage cuticle, indicating failure to rescue molting defects. The effects on egg-laying ability, fertility, and brood size were also greatly dampened relative to the rescuing abilities of *lin-42*b or *lin-42*a. We have not ruled out the possibility that the *lin-42b/*c promoter itself contributes to the minimal rescue observed with *lin-42*c. We note that the extent of rescue among *lin-42*c lines appeared somewhat more variable than for *lin-42*a or *lin-42*b rescue, and the variability does not correlate with the co-injection marker used (*str-1::gfp* vs *sur-5::gfp*; see Tables 1, 2). Another possibility is that the variable rescue could reflect a requirement for an optimal expression range that may vary for the different phenotypes scored. For example, GFP is readily detected in *ox461; veEx658* animals, which show rescue of the alae phenotype but not larval arrest, and conversely, GFP is not detected in *ox461; veEx321* animals which show significant rescue of the larval arrest phenotype, but less extensive rescue of precocious alae. Regardless of the reason underlying the *lin-42*c results, the rescue experiments clearly demonstrate that the *lin-42*c-expressing construct is dramatically less efficient than either *lin-42*a or *lin-42*b at providing function.

lin-42 transcripts have similar accumulation patterns

lin-42a is expressed from a different promoter than are *lin-42b* and *lin-42c* (TENNESSEN *et al.* 2006), raising the question as to whether their temporal expression patterns differ. *lin-42* expression levels were reported to be oscillatory using a PCR approach that would detect both *lin-42b* and *lin-42c* (JEON *et al.* 1999), with levels peaking in the intermolts and becoming undetectable during the molts, and subsequent reports confirmed a pulsatile expression pattern for *lin-42* using a variety of methods (GISSENDANNER *et al.* 2004; HENDRIKS *et al.* 2014). However, assays of promoter fusions to destabilized fluorescent reporter proteins have led to conflicting reports about the relative order of *lin-42a* versus *lin-42b* expression (MONSALVE *et al.* 2011; PERALES *et al.* 2014). As an alternative method to address this question, we measured endogenous levels of each *lin-42* transcript, using the Roche Universal Probe Library (UPL) RT-qPCR assay, and compared their temporal expression profiles throughout development. The primer sets were specific, as they failed to detect a product when the corresponding deletion allele was assayed (Figure S1). Primers for *lin-42a* and *lin-42b* spanned introns and were confirmed to detect mature transcripts; however, an intron-spanning assay specific to *lin-42c* cannot be designed, and the assay used could also detect an unspliced *lin-42b* transcript (Figure 4A). Therefore, the most informative comparison is between *lin-42a* and *lin-42b* message levels. RNA was prepared from synchronized populations of wild-type animals collected at two-hour intervals from 6-36 hrs and transcript levels were assayed, using endogenous *mlt-10* expression as an internal control to mark developmental time and monitor synchrony. *mlt-10* message levels also oscillate, peaking ~4 hrs

before each lethargus (Figure 4) (FRAND *et al.* 2005; MCCULLOCH and ROUGVIE 2014). The expression patterns of the *lin-42* transcripts were found to be highly similar to each other (Figure 4), peaking at approximately the same time and returning to a low basal level before each molt (~4 hours after the *mlt-10* peak), despite being derived from two different promoters.

Disruption of a conserved Trp residue, important for mPER dimerization, does not appear to affect LIN-42 function

Dimerization of PER is important for its stabilization and translocation to the nucleus, where it functions to negatively regulate its targets, CLK and CYC (LEE *et al.* 1999; YAGITA *et al.* 2000). PER proteins exist as homo- and heterodimers *in vivo* in mammals, and the PAS domain has been shown in flies to function as a dimerization domain (HUANG *et al.* 1993; YAGITA *et al.* 2000). Crystal structure studies revealed that *Drosophila* PER dimers are stabilized by interaction of a highly conserved tryptophan residue (Trp482) of one molecule, and a hydrophobic cleft of the second molecule (YILDIZ *et al.* 2005). Mutation of the conserved tryptophan residue in mice, W419E, disrupts homodimerization of the PAS domain *in vivo*, since introduction of a negatively charged glutamate residue disrupts interaction with the hydrophobic cleft (HENNIG *et al.* 2009; KUCERA *et al.* 2012). The tryptophan residue is conserved in LIN-42 (W258), and the W258E mutation has been shown to abrogate dimerization of LIN-42 *in vitro* (Carrie Partch, UC Santa Cruz, personal communication). We introduced the W258E mutation into the endogenous *lin-42* locus using homology directed

repair CRISPR-Cas9 technologies. Homozygous animals were scored for molting and heterochronic defects, including larval arrest (n>100), egg-laying (n>100), and L3m precocious alae (n>20) defects. Surprisingly, no defects were observed. Dimerization hasn't been examined *in vivo*, so it is possible that LIN-42 dimerization is stabilized through other residues within the PAS domain, or through non-PAS interactions. Alternatively, dimerization may not be relevant to LIN-42 function *in vivo*.

Further investigation is necessary to determine if the LIN-42 PAS domain functions as a dimerization domain *in vivo*, because the PAS domain is evolutionarily conserved and likely serves an important function. Analysis of the W258E in a more sensitized genetic background, for example in trans to the *lin-42(0)* deletion, would be useful in exploring whether W258E confers some degree of loss of function. Furthermore, targeted mutagenesis of other highly conserved residues within the PAS domain, or deletion of the entire PAS domain, will be useful to examine PAS function. It has been shown that dimerization of mammalian PERIOD, is necessary for the protein to translocate to the nucleus, where it functions (YAGITA *et al.* 2000). Like PEROID, LIN-42 localizes to the nucleus (TENNESSEN *et al.* 2006) and may also require dimerization for nuclear entry. Our lab recently introduced a fluorescent tag into the endogenous *lin-42* locus using CRISPR, and this allele will be useful in future mutagenesis studies to monitor whether the mutant protein is destabilized, and analyze whether mutations might affect protein localization.

CONCLUSION

A longstanding question in the *C. elegans* developmental timing field has been whether complete lack of *lin-42/period* activity would confer novel heterochronic phenotypes. This question has been unanswered due to lack of alleles that eliminated all isoforms. We remedied this problem by generating a null allele that deletes the entire *lin-42* coding region. Phenotypic analysis has so far revealed no new heterochronic defects and that the *lin-42(0)* mutant has similar, but more penetrant, heterochronic and molting defects than the hypomorphic alleles *lin-42(n1089)* and *lin-42(ok2385)*. These results demonstrate that the two non-overlapping transcription units, *lin-42a* and *lin-42c*, present in *lin-42(n1089)* and *lin-42(ok2385)* mutants respectively, each confer some level of function.

The three *lin-42* transcription units appear highly similar in their temporal accumulation patterns, peaking at a similar time during each intermolt and becoming undetectable during each molt. *lin-42a* and *lin-42b/c* maintain cyclical expression patterns in *lin-42(n1089)* and *lin-42(ok2385)* mutants respectively, indicating their dynamic expression patterns do not require the missing isoform(s).

Transgene expression experiments reveal that the most important isoforms are *lin-42*a *and lin-42*b*. lin-42*a lacks PERIOD's hallmark PAS domain, but contains the smaller conserved SYQ/LT domains, and *lin-42b* contains all three*. lin-42a* or *lin-42b* expression strongly rescues *lin-42(0)* mutants for all parameters assessed, whereas PAS domain encoding *lin-42c* only mildly

rescues *lin-42(0)* heterochronic defects. The strong rescuing ability of *lin-42*b indicates that independent expression of the short isoforms is not absolutely required for function.

Aside from the fact that SYQ/LT domain containing isoforms are key, the functions of the conserved motifs in LIN-42 are unclear. Indeed, to date no missense alleles of *lin-42* have been reported that cause phenotypes; all alleles are deletions or result in premature stop codons. The availability of a null allele will now facilitate structure function analysis, and in conjunction with Crispr-Cas9 genome engineering approaches, should allow correlation of functions with isoforms, and more precisely, to the domains they harbor and the amino acids that comprise them.

lin-42 has a well defined role as a negative transcriptional regulator of miRNA genes, including some that function in the heterochronic gene pathway (McCulloch and Rougvie 2014; Perales *et al.* 2014; Van Wynsberghe *et al.* 2014), potentially allowing *lin-42* to indirectly coordinate translation of many downstream messenger RNAs. Initial chromatin immunoprecipitation coupled to sequencing (ChIP-seq) experiments have indeed found LIN-42 associated with chromatin at miRNA promoters, but in addition, LIN-42 is found near the transcription starts of protein coding genes (Perales *et al.* 2014). Key goals for the future will be to validate these targets and begin to partition them among *lin-42*'s various roles, which include the molting, timing, and dauer formation pathways, as well as to determine whether the *lin-42* isoforms are differentially recruited to promoters. Interestingly, levels of LIN-42C present in *lin-42(n1089)*

mutants are sufficient to carry out the molting function, but the not heterochronic function, perhaps reflecting a division of labor among isoforms. One possible model is that the region spanning the SYQ and LT domains is important for LIN-42 association with most targets and similar to PERIOD, acts to block the action of transcriptional activators. The PAS domain may provide a more regulatory function, enhancing the activity of the SYQ/LT region, or allowing it access a wider array of target genes.

Acknowledgements

We thank Dr. Iva Greenwald for thoughtful comments on the manuscript and Dr. Alison Frand for sharing ARF224. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). This work was supported by the National Institutes of Health [grants R01GM50227 (A.E.R), R01GM095817 (E.M.J), and 5T32HD007480 (T.B.E)], the Minnesota Medical Foundation (A.E.R.), the Carlsberg Foundation (C.F.J.), and the Direktør Ib Henriksens Foundation (C.F.J.).

Figure 1: *lin-42(ox461)* deletes the *lin-42* coding region

(A) *lin-42* genomic locus and transcription units. The line with terminal arrowheads represents genomic DNA of the *lin-42* locus drawn with 5' to the left, which is inverted from WormBase. *lin-42* alleles and the *Mos*1 insertion (*ttTi42556*) are indicated above the line. *Sac*I and *Eco*RI sites are included for reference with (B) and (C), but not all recognition sites for these enzymes are shown. The *lin-42* locus produces three transcription units diagrammed below with filled boxes representing exons: *lin-42a*, *lin-42b*, and *lin-42c*. *lin-42a* and *lin-42c* are non-overlapping and expressed from distinct promoters (TENNESSEN *et al.* 2006). Note that the *lin-42* nomenclature used here conforms to that adopted by WormBase and differs from that of pre-2014 publications from the Rougvie lab (e.g. TENNESSEN *et al.* 2006; 2010). (B) *lin-42* deletion alleles. The extent of each deletion is noted in brackets. The *lin-42(n1089)* PASA domain is in brackets since a majority of the domain is deleted. (C) The *lin-42(ox461)* allele deletes the *lin-42* coding region and replaces it with P*unc-122:gfp* and *C. briggsae unc-119*(+). The fragments used as repair templates in creation of the deletion allele are indicated. See Materials and Methods for details.

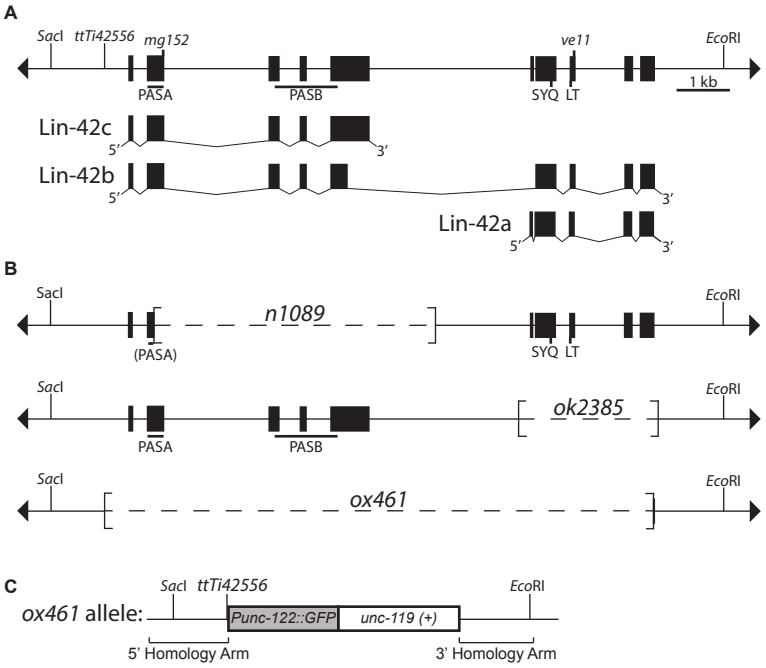

A
SacI
ttTi42556
mg152
ve11
EcoRI
PASA
PASB
SYQ
LT
1 kb
Lin-42c
5′
3′
Lin-42b
5′
3′
Lin-42a
5′
3′
B
SacI
n1089
EcoRI
(PASA)
SYQ
LT
SacI
ok2385
EcoRI
PASA
PASB
SacI
ox461
EcoRI
C
SacI
ttTi42556
EcoRI
ox461 allele:
Punc-122::GFP
unc-119 (+)
5' Homology Arm
3' Homology Arm

Figure 1

Figure 2: *lin-42(0)* animals exhibit highly penetrant larval arrest and developmental delay phenotypes

(A) *lin-42(0)* mutants have a severe larval arrest phenotype. Individually plated wild-type, *lin-42(n1089), lin-42(ok2385),* and *lin-42(ox461)* mutants were monitored for developmental progression. The percent of animals that arrested as larvae and failed to attain adulthood by 8 days post-hatching is shown. n ≥ 100 for each genotype. (B) Micrographs of wild-type and *lin-42(ox461)* animals 72 hrs post-hatching. At this time-point, all wild-type animals had reached adulthood and were laying eggs, whereas <2% of *lin-42(0)* mutants were adults, and they did not yet contain fertilized eggs, only oocytes. Most *lin-42(0)* animals appeared unable to complete the second larval molt. The left *lin-42(0)* panel contains two animals that are the same chronological age, with an inset showing the smaller animal is trapped in an unshed cuticle (white arrow). The right panel shows a 96 hr animal that appears arrested by size, but has nevertheless begun vulval morphogenesis, marked by the white v in inset. (C) Timing of molts in seven representative wild-type animals (top) and *lin-42(0*) mutants (bottom). Each horizontal row represents an individual animal that was monitored for pharyngeal pumping, lethargus, and ecdysis. Gray shading denotes animals in lethargus and vertical red lines indicate that a shed cuticle was observed on the plate. A total of 68 animals were followed for *lin-42(ox461)* and 10 for wild-type. (D) The time of adult onset is delayed in *lin-42* mutants. Animals from (A) that bypassed larval arrest were scored for the day post hatching that they reached

adulthood, and graphed as percent of animals that escaped arrest. n = 20, 125, 60, 23 for wild-type, *n1089*, *ok2385* and *ox461*, respectively.

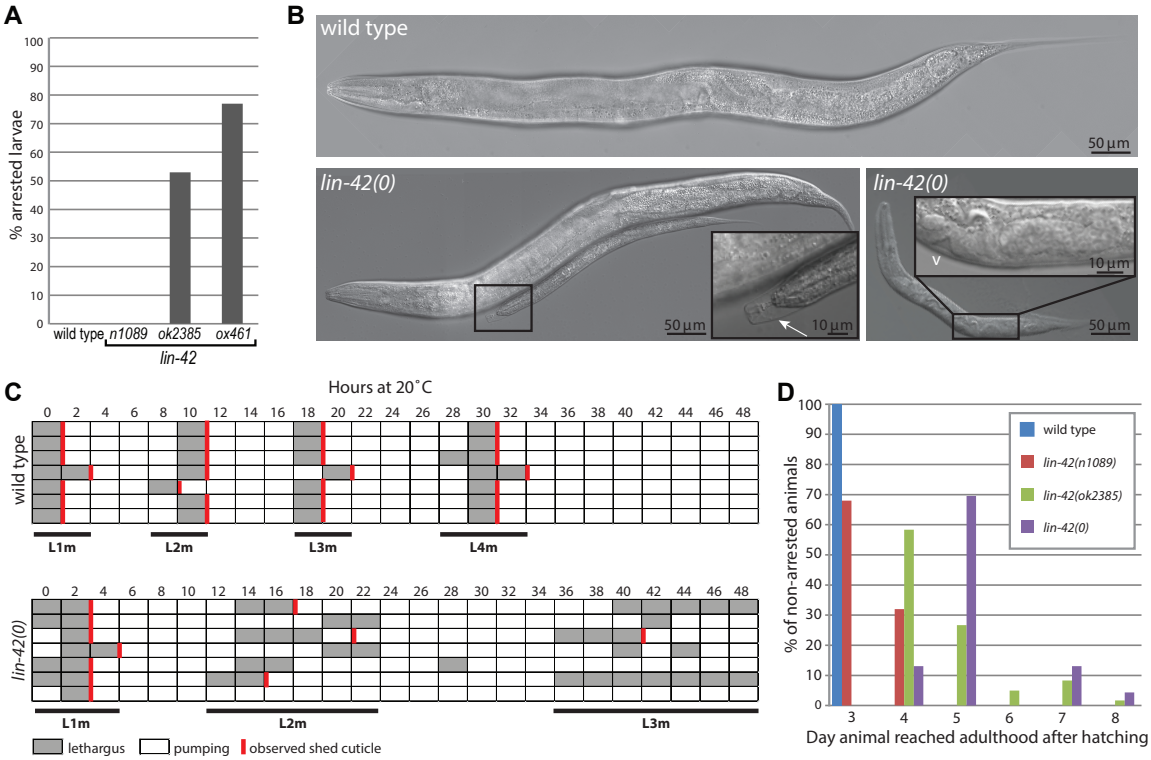

A
% arrested larvae
wild type
n1089
ok2385
ox461
lin-42
B
wild type
50 µm
lin-42(0)
50 µm
10 µm
lin-42(0)
v
10 µm
50 µm
C
Hours at 20˚C
wild type
lin-42(0)
L1m
L2m
L3m
L4m
lethargus
pumping
observed shed cuticle
D
% of non-arrested animals
wild type
lin-42(n1089)
lin-42(ok2385)
lin-42(0)
Day animal reached adulthood after hatching

Figure 2

Figure 3: *lin-42(0)* mutants have a precocious heterochronic phenotype

(A) *lin-42* mutations cause precocious seam cell development. Shown is a representative seam cell lineage (V1-V4, V6) diagram in wild-type and *lin-42(0)* mutants. A horizontal line indicates a cell division, and triple horizontal bars indicate alae formation. Developmental stages are relative to the molts, and the actual intermolt periods of wild-type and *lin-42(0)* animals are different (see Figure 2C). (B) *lin-42(0)* mutants have a more severe heterochronic defect than *lin-42(lf)* animals**.** Wild-type, *lin-42(n1089), lin-42(ok2385),* and *lin-42(ox461)* animals were analyzed at the L3-molt for alae formation. Animals were scored for either complete, partial (alae with gaps), or no alae. n≥20 for each strain. **P<0.0001, Fisher's exact test. (C) *lin-42(0)* animals form precocious alae at the L3-molt stage. Micrographs of representative wild-type and *lin-42(0)* animals at the third molt are shown, with an image of cuticle in the top row and vulval morphogenesis (v) of the animal below to denote staging. Bar, 10 µm, applies to all panels.

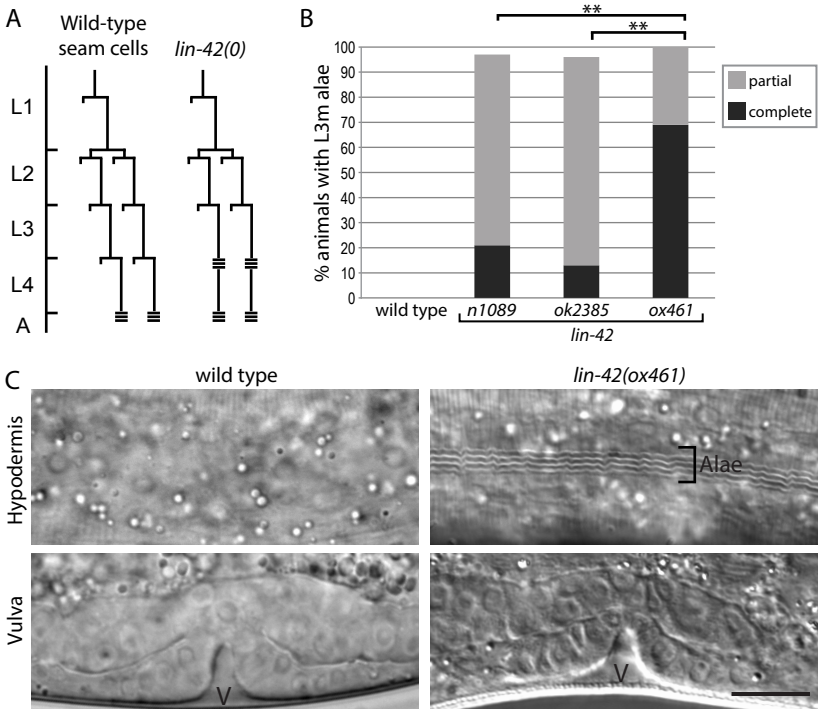

A
Wild-type
seam cells
lin-42(0)
L1
L2
L3
L4
A
B
**
**
% animals with L3m alae
100
90
80
70
60
50
40
30
20
10
0
partial
complete
wild type
n1089
ok2385
ox461
lin-42
C
wild type
lin-42(ox461)
Hypodermis
Alae
Vulva
V
V

Figure 3

Figure 4: Levels of three *lin-42* transcripts cycle in unison.

(A) lin-42 transcript diagrams with small boxes above indicating the locations of primer sets used in qPCR assays: *lin-42a* (red square), *lin-42b* (green triangle), and *lin-42c* (purple X). *lin-42a* and *lin-42b* assays are intron spanning and recognize a single transcript, whereas *lin-42c* does not contain a unique intron; its primer could also amplify the *lin-42b* primary transcript. (B-D) Representative time courses of *lin-42a*, *lin-42b*, and *lin-42c* accumulation, relative to the time of *mlt-10* expression, in wild-type and *lin-42* mutants. Two independent biological replicas for each genotype are shown in Figure S1, along with primer controls. (B) Wild-type with time points normalized to 0 hr. (C-D) *lin-42a* and *lin-42c* levels cycle in *lin-42* mutant backgrounds. Time points within each assay are normalized to 6 hr. (C) *lin-42a* message levels in *lin-42(n1089)* mutants. (D) *lin-42c* message levels in *lin-42(ok2385)* mutants from 6-24 hrs.

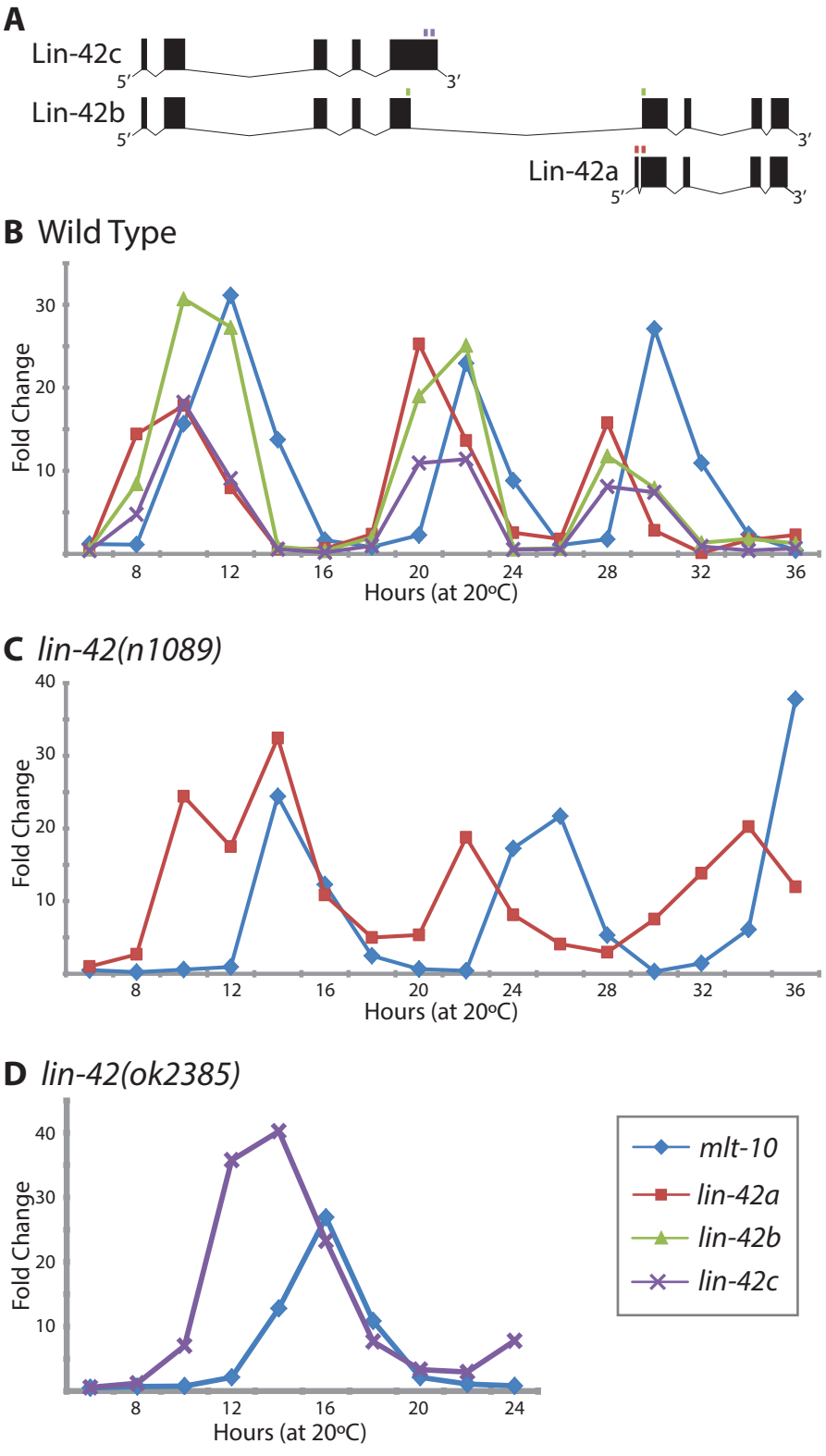

A
Lin-42c
5′
3′
Lin-42b
5′
3′
Lin-42a
5′
3′
B Wild Type
Fold Change
30
20
10
8
12
16
20
24
28
32
36
Hours (at 20ºC)
C lin-42(n1089)
Fold Change
40
30
20
10
8
12
16
20
24
28
32
36
Hours (at 20ºC)
D lin-42(ok2385)
Fold Change
40
30
20
10
8
12
16
20
24
Hours (at 20ºC)
mlt-10
lin-42a
lin-42b
lin-42c

Figure 4

Figure S1: Biological replicates of *lin-42* transcript accumulation patterns. (A-C) Expression of *lin-42a*, *lin-42b*, and *lin-42c* was measured relative to *mlt-10* in a total of three biological replicates (one replicate is displayed in Figure 4) for (A) wild type, (B) *lin-42(n1089)*, and (C) *lin-42(ok2385)*. Time points within each assay are normalized to 0 or 6 hr. (B) The decrease in *lin-42a* expression at 12hrs in each *lin-42(n1089)* time course is an artifact of staggering populations, since the decrease is not observed when animals were collected continuously from 6-16 hrs (third panel). (D) *lin-42* gene expression assays are specific to the intended target. Transcript levels were measured in mixed-stage populations of wild type, *lin-42(n1089)*, *lin-42(ok2385)*, and *lin-42(0)* to demonstrate the specificity of each assay. ΔCt values ([target Ct] – [*ama-1* Ct]) are normalized to wild type.

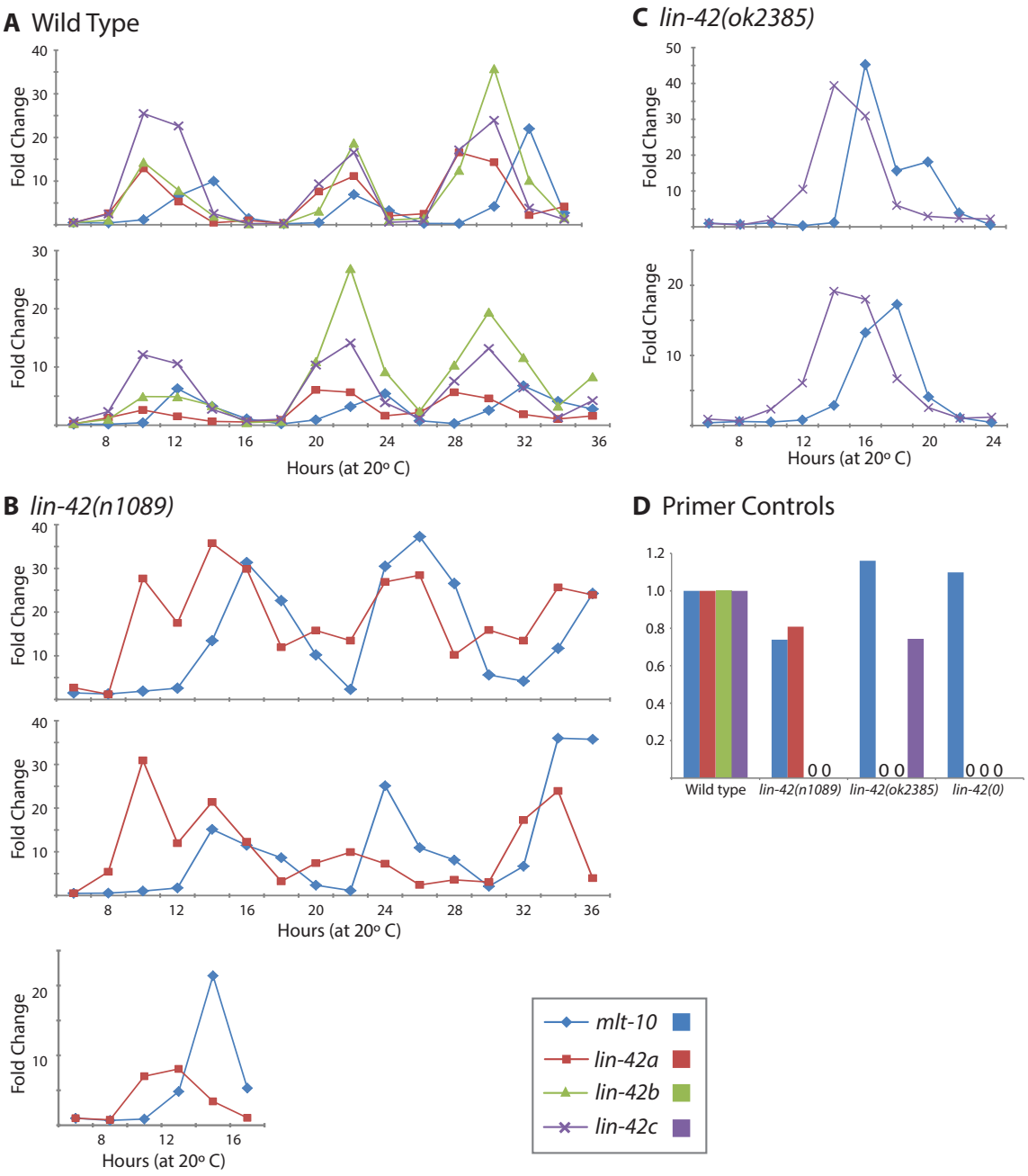

A Wild Type
Fold Change
Hours (at 20° C)
C lin-42(ok2385)
Fold Change
Hours (at 20° C)
B lin-42(n1089)
Fold Change
Hours (at 20° C)
D Primer Controls
Wild type
lin-42(n1089)
lin-42(ok2385)
lin-42(0)
mlt-10
lin-42a
lin-42b
lin-42c

Figure S1

Table 1 Strains used in this study

Strain	Genotype[a]	Reference
ARF224	*lin-42(ok2385)*	Monsalve et al. 2011
EG15910	*lin-42(ox460 [Punc-122::gfp + Cbr-unc-119(+)]); unc-119(ed3)*	This work
EG15911	*lin-42(ox461 [Punc-122::gfp + Cbr-unc-119(+)]); unc-119(ed3)*	This work
GR1395	*mgIs49* [P*mlt-10::gfp-pest* + P*ttx-3::gfp*]	Hayes et al. 2006
IE42556	*ttTi42556*	Vallin et al. 2012
JR667	*unc-119(e2498::Tc1)*; *wIs51* [P*scm::gfp* + *unc-119(+)*]	Antebi et al. 2000
MT2257	*lin-42(n1089)*	Abrahante et al. 1998
N2	Wild type var. Bristol	Brenner 1974
RG1514	*ttTi42556; unc-119(ed3)*	This work
RG1580	*lin-42(ox461); veEx323* [pHG83 *(lin-42a(gDNA)) + str-1::gfp*]	This work; Tennessen et al. 2006
RG1590	*lin-42(ox461)*	This work
RG1650	*lin-42(ox461); wIs51*	This work
RG1665	*lin-42(ox461); veEx593* [pHG83 + *sur-5::gfp*]	This work
RG1739	*lin-42(ox461); veEx651* [pCP2 (P*lin-42b/c::lin-42b(cDNA)::gfp::unc-54 3'UTR*) + *str-1::gfp*]	This work
RG1757	*lin-42(ox461); veEx652* [pCP2 + *str-1::gfp*]	This work
RG1758	*lin-42(ox461); veEx655* [pCP2 + *str-1::gfp*]	This work
RG1786	*lin-42(ox461); veEx594* [pHG83 + *sur-5::gfp*]	This work
RG1791	*lin-42(ox461); gaIs233*	This work
RG1792	*lin-42(ox461); mgIs49*	This work
RG1816	*lin-42(ox461); veEx321* [pMJ13 (P*lin-42b/c::lin-42c(gDNA)::gfp::lin-42 3'UTR*) + *str-1::gfp*]	This work; Tennessen et al. 2006
RG1822	*lin-42(ox461); veEx657* [pMJ13 + *str-1::gfp*]	This work
RG1823	*lin-42(ox461); veEx658* [pMJ13 + *str-1::gfp*]	This work
RG1824	*lin-42(ox461); veEx317* [pMJ13 + *sur-5::gfp*]	This work
RG1825	*lin-42(ox461); veEx319* [pMJ13 + *sur-5::gfp*]	This work; Tennessen et al. 2006
SD1434	*unc-119(ed3); gaIs233* [P*elt-5::HIS-24::mCherry* + *unc-119(+)*]	Liu et al. 2009

[a]Full genotypes are given as appropriate at the first appearance of an allele or transgene array. Key components of plasmids are similarly detailed at the first appearance of the plasmid.

Table 2 *lin-42* heterochronic and larval arrest phenotypes

	Animals with L3m alae (%)					Animals that reached adulthood		
Genotype	Complete	Partial	None	% Arrest[a]	No. Seam cell[b]	Egl (%)	Fertile (%)	Brood Size[c]
wild-type	0	0	100	0	nd	0	100	336.7 ± 37.7
lin-42(n1089)	21	76	3	0	nd	18	100	69.1 ± 67.5
lin-42(ok2385)	13	83	4	53	nd	78	85	36.5 ± 46.1
lin-42(ox461)	69	31	0	77	nd	100	72	7.6 ± 5.2
P*elt-5::mCherry*	0	0	100	nd	15.7 ± 0.75	nd	nd	nd
ox461; P*elt-5::mCherry*	59	41	0	nd	16.0 ± 0.56	nd	nd	nd
P*scm::gfp*	0	0	100	nd	16.0 ± 0.00	nd	nd	nd
ox461; P*scm::gfp*	0	5	95	nd	16.0 ± 0.32	nd	nd	nd
P*mlt-10::gfp*	0	0	100	nd	nd	nd	nd	nd
ox461; P*mlt-10::gfp*	9	14	77	nd	nd	nd	nd	nd
lin-42a rescue								
ox461; veEx323	0	14	86	0	nd	3	100	147.0 ± 25.3
ox461; veEx593	0	28	72	2	nd	1	99	189.3 ± 70.8
ox461; veEx594[d]	0	0	100	4	nd	10	100	106.0 ± 57.9
*lin-42*b rescue								
ox461; veEx651	0	4	96	2	nd	4	100	193.3 ± 71.1
ox461; veEx652	0	4	96	2	nd	6	100	208.8 ± 84.2
ox461; veEx655	0	6	94	14	nd	16	100	186.5 ± 66.1
lin-42c rescue								
ox461; veEx317	17	48	35	59	nd	85	85	12.3 ± 11.7
ox461; veEx319	0	67	33	61	nd	76	86	19.5 ± 30.4
ox461; veEx321	22	74	4	37	nd	97	81	13.1 ± 25.9
ox461; veEx657	0	19	81	83	nd	75	89	20.8 ± 32.1
ox461; veEx658	4	68	28	74	nd	46	97	45.3 ± 49.5

All animals assayed arose from eggs hatched onto seeded plates at low density. n≥20 for all analyses except brood size of *ox461* and *ox461; veEx657* animals where n=16.

[a]Percentage of animals that arrested as young larvae.

[b]P*elt-5::mCherry* strains were scored as young adults, and P*scm::gfp* strains were scored during the L3 molt.

[c]Average number of progeny from fertile adults; sterile animals were not included in the calculation.

[d]This strain exhibited some embryonic lethality; non-Egl adults laid an average of 33 ± 23 dead eggs.

CHAPTER 3: MIR-48 OVER-EXPRESSION SUPPRESSOR SCREEN

SUMMARY

Heterochronic genes regulate the timing of developmental events in *C. elegans*. Animals with mutations in these genes either skip or reiterate certain events, and are classified as either precocious or retarded mutants, respectively. Identification of the relevant mutations of retarded heterochronic mutants led to the discovery of the first two miRNAs, *lin-4* and *let-7*. *let-7* belongs to a miRNA gene family, including *mir-48*, *mir-241*, and *mir-84*, members of which share homology at the 5'end of the mature miRNA product. Disruption of *mir-48*, *mir-241*, or *mir-84* individually results in little or no heterochronic phenotype; however, disruption of all three miRNAs results in a strong retarded defect, indicating that these three miRNAs function redundantly. Over-expression of *mir-48* from a multicopy array causes a gain-of-function precocious phenotype and also results in a strong egg-laying defect. To identify genes that interact with *mir-48*, we performed a suppressor screen and identified suppressors that restored egg-laying. This screen is powerful because several classes of suppressors are expected to be recovered, including new heterochronic mutations, miR-48 target genes, as well as general miRNA biogenesis and function genes. Thirty-six suppressors were identified after screening 48,000 haploid genomes, and were characterized. Five alleles of the heterochronic gene *lin-66* were identified, validating this approach as a way to identify new heterochronic mutations. miR-48 levels were measured among the suppressor strains and there were

remarkable differences, indicating that different classes of suppressors were isolated. Four suppressors were mapped to chromosomes and subjected to whole-genome sequencing, though the relevant suppressor mutations have yet to be identified.

INTRODUCTION

Investigation of the heterochronic gene *lin-4* in *C. elegans* led to the finding of the first miRNA in 1993 (CHALFIE *et al.* 1981; AMBROS and HORVITZ 1984; LEE *et al.* 1993). *lin-4* was discovered to encode a 21 nt regulatory RNA that was predicted to bind multiple partially complementary sites in the *lin-14* 3'UTR and regulate its expression (ARASU *et al.* 1991; WIGHTMAN *et al.* 1991; LEE *et al.* 1993; WIGHTMAN *et al.* 1993). The second miRNA identified, Let-7, was also discovered in *C. elegans*, and remarkably, is perfectly conserved from nematodes to humans (PASQUINELLI *et al.* 2000; LAGOS-QUINTANA *et al.* 2003). Since then, countless miRNAs have been identified throughout the animal kingdom as regulators of virtually every biological process (FRIEDMAN *et al.* 2009). miRNAs are transcribed as long primary transcripts, that go through a series of processing steps before they reach their active form (WINTER *et al.* 2009). The primary transcript is cleaved in the nucleus by the Microprocessor complex into a ~60nt hairpin structure that is exported from the nucleus into the cytoplasm where it is cleaved by Dicer into a ~21nt heteroduplex. The duplex is loaded onto the multi-protein miRNA Induced Silencing Complex (miRISC) and one

strand is released. miRISC then interacts through a partial complementary RNA-RNA interaction between the miRNA and the 3'UTR of the targeted mRNA to inhibit translation or decrease mRNA stability. The accumulation of a mature miRNA is tightly regulated and regulation can occur at any step throughout this process.

C. elegans undergoes four larval stages (L1-L4) before reaching adulthood. Lin-4 functions early in development to promote the L1-to-L2 transition, while Let-7 functions in the L3 stage to promote developmental progression (LEE *et al.* 1993; REINHART *et al.* 2000; VADLA *et al.* 2012). If these miRNAs do not function properly, developmental programs are reiterated, causing subsequent stages to occur too late or not at all. These defects can be easily observed in the hypodermal seam cells, which terminally differentiate at the larval-to-adult transition. *let-7(lf)* mutants reiterate a larval stage which causes hypodermal seem cells to terminally differentiate after five larval stages (REINHART *et al.* 2000).

Let-7 belongs to a family of miRNAs that share a common seed region (Figure 1A) (ABBOTT *et al.* 2005). The seed region, which consists of nucleotides 2-8 of the mature miRNA, plays an important role in binding to the target mRNA (LAU *et al.* 2001; LIM *et al.* 2003). Three of the let-7 family members, *mir-48*, *mir-241*, and *mir-84*, function redundantly to promote the transition from the L2-to-L3 transition in *C. elegans* (ABBOTT *et al.* 2005). If one of the *let-7* family miRNAs is disrupted there is little to no phenotype, however in the absence of all three

miRNAs, animals have a strong retarded defect that results in a reiteration of the L2 stage seam cell programs (Figure 1B). At the L2 stage in wild type, seam cells undergo a proliferative division, thus *mir-48*/*241*/*84(0)* animals reach adulthood with more seam cells than wild-type animals. Additionally, seam cells in these mutants terminally differentiate one stage later than in wild-type animals. miR-48/241/84 collectively regulate the L2-to-L3 transition by down-regulating their targets *hbl-1, daf-12,* and perhaps *lin-28* (ABBOTT *et al.* 2005). Further investigation of miR-48/241/84 may uncover other targets that have yet to be identified.

Because *mir-48/241/84* function redundantly, an analysis of their loss-of-function is complicated. However, gain-of-function studies can provide additional insight. For example, two *mir-48(gf)* mutations were isolated from a *lin-4(lf)* suppressor screen (ABRAHANTE *et al.* 1998; LI *et al.* 2005). When analyzed on their own *mir-48(ve33)* and *mir-48(ve12)* were found to cause precocious expression of miR-48, which is sufficient to cause a precocious alae phenotype (ABRAHANTE *et al.* 1998). Both alleles have mutations in a GC-rich element upstream of the *mir-48* coding sequence, which suggested that this region was important to regulate *mir-48* expression (LI *et al.* 2005). Over-expression of wild-type *mir-48* from a multi-copy array causes a precocious alae phenotype, as well as a mild **EG**g **L**aying defective (Egl) phenotype. However, over-expression of *mir-48(ve33)* from a multi-copy array (*veIs48*) causes a stronger precocious defect and a fully penetrant Egl defect, which is advantageous for a suppressor

screen. Hermaphrodites are unable to lay eggs, therefore the progeny hatching inside the parent resulting in a “bag of worms.” Animals with the *veIs48* array were mutagenized and suppressors identified by searching for non-Egl animals using a dissecting microscope.

Suppressor/enhancer screens are powerful methods to study a gene of interest because they can identify genes that biologically interact with a gene. A *mir-48* over-expression suppressor screen could identify mutations in new heterochronic genes that function with *mir-48* in the heterochronic pathway, mutations in miR-48 targets that would prevent miR-48 from binding to and downregulating its expression, and finally, mutations in genes that regulate miRNA expression and function would reduce the level of miR-48 accumulation or impede the ability of miR-48 to regulate its target. New heterochronic mutations could lead to the identification of a new member of the heterochronic gene pathway or a new interaction between *mir-48* and an existing member of the heterochronic pathway. Identification of genes involved in miRNA biogenesis and function could expand upon our limited understanding of such a highly regulated process. Finally, identification of a mirR-48 target would be exciting since only a few have been confirmed.

MATERIALS & METHODS

Strains and Nematode Maintenance. *C. elegans* were grown and maintained at 20°C on NGM plates seeded with *E. coli* OP50 as previously described (BRENNER 1974).

qPCR. Eggs were isolated by hypochlorite treatment of 6-10 plates of gravid adults and then hatched overnight in M9 buffer in 15ml tube. To harvest worms, debris was allowed to settle for 10-15 minutes. Actively swimming worms in the top 5-6 ml were transferred to a new tube. The original tube was filled up to 10ml with M9, and the harvest process was repeated 3 more times. Harvested animals were pelleted and flash frozen in liquid nitrogen. Worm samples were processed through three rounds of freezing in liquid nitrogen, thawing at 37˚, and vortexing with 425-600 µm glass beads (Sigma) to disrupt the cuticle. RNA was extracted using Trizol (Life Technologies) according to manufactures instructions. RNA (7.5 ng) was reverse transcribed with the TaqMan miRNA reverse transcription kit using specific miRNA primers. Quantitative Taqman miRNA assays were performed as directed (Life Technologies Part# 4427975 AssayID# 208(miR-48), 249(miR-241), 236(miR-84)). qPCR reactions were run in triplicate on an Eppendorf Realplex Thermocycler using 96-well plates, and data were collected and analyzed using Realplex 2.0 software. Each time-point was normalized to miR-1 or miR-58 as an internal control. Data was analyzed using the ΔΔCt method in Microsoft Excel (LIVAK and SCHMITTGEN 2001).

SNP mapping. SNP mapping was performed using the method described in Davis et al. (2005). Briefly, mutated strains were backcrossed to the CB4856 strain, single heterozygous F1 progeny allowed to self-fertilize, producing F2 animals that were sorted into Egl and non-Egl pools to generate lysates. Selected SNPs were PCR amplified using the primers designed in Davis et al. (2005) and digested with DraI. Products were run out on an agarose gel to compare digest patterns between the Egl and non-Egl samples. Since crossing the *veIs48* array into the the CB4856 strain background enhanced the egg-laying phenotype, mapping was performed using strains that contained only one chromosome from the CB4856 strain. WE5236 was used for Ch I, WE5238 for Ch III, WE5241 for Ch IV. When a single chromosome was SNP mapped, primers were used for only that chromosome.

DNA prep for Whole Genome Sequencing. Suppressed strains were backcrossed 4 times to remove any background mutations. To prep genomic DNA, animals were washed in M9 2-3 times, soaked in M9 on a rocker for 2 hours to remove food from gut, and then washed in M9 2-3 times. Worms were pelleted and flash frozen. DNA was prepped using the Gentra Puregene kit (Qiagen). Pellet was thawed and resuspended in 3 ml of Cell Lysis Solution (Qiagen Gentra Puregene kit). Worms were disrupted with a sonicator at amplitude 20% for 30 sec (pulse: ON for 1 sec, OFF for 1 sec). 15µl of Proteinase K (20mg/ml) was added and incubated at 55˚ for 3 hrs, during which the tube was inverted periodically. The lysate was cooled to room temp, 15µl

RNase A added, and incubated on rocker for 1 hr. The lysate was cooled for 3 min on ice, 1ml Protein Precipitate Solution (Qiagen Gentra Puregene kit) was added and samples vortexed for 20 sec. at high speed. Lysates were centrifuged for 10 min at 2000xg, supernatant was transferred to a new tube, 3 ml isopropanol was added and inverted to mix. Samples were centrifuged for 3 min again at 2000xg, supernatant transferred to a new tube, 3 ml of 70% ethanol was added and inverted to wash pellet. Sample was centrifuged for 3 min at 2000xg, supernatant removed, and pellet air dried. 150 µl of DNA Hydration Solution was added to the pellet, resuspended by pipetted up/down, and transferred to clean eppendorf tube. Incubated at 65˚ for at least 30 min until DNA was dissolved.

Whole Genome Sequencing. Whole-genome sequencing was conducted on Illumina GAIIx instrument according to the manufacturer's instructions by the UMN genomics core facility. The sequencing depth coverages were 41 (RG1383), 34 (RG1385), 52 (RG1437), 49 (RG1440). Data were analyzed using MAQGene (BIGELOW *et al.* 2009).

RESULTS

mir-48 over-expression suppressor screen

A *mir-48* loss-of-function screen is not feasible, since *mir-48* functions redundantly with *mir-241* and *mir-84*. *mir-48(gf)* mutants identified in a *lin-4(lf)* suppressor screen have precocious heterochronic defects, therefore a *mir-48*

over-expression screen was performed to identify genes that interact with *mir-48* *(ABRAHANTE et al. 1998)*. miR-48 was over-expressed from an integrated multi-copy array (*veIs48*) that expressed *mir-48(ve33)* and *sur-5::gfp (LI et al. 2005)*. Animals containing the *veIs48* array display a fully penetrant Egl phenotype (Figure 1C) since vulval precursor cells divide precociously and cause abnormal vulval development. Animals were mutagenized with EMS and F2 progeny screened for restoration of egg-laying, with the expectation that the other heterochronic defects would also be restored. 48,000 haploid genomes were screened, and 36 suppressors were isolated.

Suppressors that silence multi-copy arrays are one class of suppressors that are likely to come out of this screen; however, they would not provide useful information about *mir-48* and need to be avoided. A *sur-5::gfp* reporter, which is expressed at high levels and in all cell nuclei, was included in the multi-copy array to monitor expression from the array. Since multi-copy array suppressors should decrease expression of *mir-48* as well as the *sur-5::gfp* array marker, selecting arrays with high levels of GFP expression can exclude suppressors that silence multi-copy arrays.

Animals were analyzed by eye on a Nikon fluorescent dissecting microscope and given a score from 0-5 in a non-blind analysis, 5 being as bright as the original *veIs48* array and 0 being undetectable GFP. *ve119*, *ve126*, *ve135*, and *ve152* are some of the suppressors that were classified as multi-copy array suppressors, whereas *ve145* and *ve132* expressed GFP at a level

indistinguishable from *veIs48* (Table 1). Suppressors with high GFP levels are of most interest.

Prioritization of suppressors

Suppressors identified in the *mir-48* over-expression screen were initially scored for penetrance of the suppression, since strong suppressors are more likely to lead to the identification of genes that play more central roles. Five strains were analyzed for penetrance of Egl suppression, *ve151* had the highest level of penetrance since 95% of animals laid eggs, whereas *ve150* had the lowest penetrance at 35%. The other three suppressors found in strains *ve145*, *ve148*, and *ve153* were about 70% penetrant (Table 2). These results demonstrate that some of the suppressors are not fully penetrant, and that *ve151* is a suppressor that should be prioritized because it most efficiently suppresses the *mir-48* over-expression defect.

Outcross suppressed strains.

Animal strains with bright, uniform GFP expression and high penetrance of Egl suppression were outcrossed to analyze two key properties: whether the suppressor displayed a phenotype in the absence of the *mir-48* array and whether the suppressor had a maternal effect.

Phenotypic analysis of suppressors. Suppressors that come out of this screen may have recognizable phenotypes when separated from the original

veIs48 array. Many heterochronic mutants have vulval defects that cause animals to have a protruding vulva (P-vul), explode from the vulva, or have an Egl defect. Mutations in genes that are essential for embryogenesis or larval development will cause an embryonic or larval lethal defect, and many known heterochronic mutants have a semi dumpy (Dpy) appearance. These and other morphological defects were scored among isolated suppressor alleles.

Animals from suppressed strains were backcrossed to N2 males, individual F1s were allowed to self fertilize, and non-fluorescent F2 progeny were analyzed for obvious phenotypes. A very distinct vulval bursting phenotype was observed when five strains were outcrossed: RG1365, RG1366, RG1386, RG1424, and RG1436. These animals burst at approximately the L4 stage, the gonad protruded from the vulva, and animals continued to crawl around the plate. This phenotype appeared similar to mutants of *lin-66,* a gene previously reported to function in parallel with *mir-48* in the heterochronic gene pathway to promote the L2-to-L3 transition (MORITA and HAN 2006). *lin-66* was sequenced in each of these strains and five new alleles of *lin-66* were revealed (Figure 2). These results are exciting because they validate the screening approach to identify heterochronic mutations.

The remaining suppressors that had uniform and bright GFP expression were outcrossed and their *gfp*- F2 progeny were scored as described above for egg-laying, protruding vulva, or lethal phenotypes, however no other phenotypes were found. It is possible that some of the suppressors caused mild phenotypes

that were not observed through analyses performed using a dissecting microscope.

Analysis of maternal effect. When strains were backcrossed with N2 males, bright green (GFP+/+) F2s were single-picked and scored for egg-laying to determine if suppressors had a maternal effect, which requires a hermaphrodite to be homozygous for the suppressor in order to produce non-Egl animals. If the F2 was Egl, all the F3 offspring were raised on the original F2 plate and the entire plate was scored for ANY egg-layers, and again the plate was scored for any F4 egg-layers. If a recessive suppressor has no maternal effect, 25% of the F2s will lay eggs, 75% of the plates will produce F3 egg layers, and the same 75% of the plates will have F4 egg layers. If there is a maternal effect, 0% of the F2s will lay eggs, 25% of the plates will have F3 egg-layers, and 75% of the plates will have F4 egg-layers. The *ve145* allele behaved like a non-maternal effect gene, while *ve142* and *ve143* behaved like suppressors with a maternal effect (Figure 3). Results were a bit different with the *ve150* suppressor. None of the *ve150* plates had F2 or F3 egg-layers, but 77% contained F4 egg-layers. These results could be explained because of the low penetrance of suppression observed with this suppressor in addition to having a maternal effect.

Measurement of miR-48 levels using qPCR

Heterochronic suppressors that function in parallel to or downstream of *mir-48* in the pathway are expected to have little effect on miR-48 levels, whereas suppressors that contribute to miRNA processing or stability are likely to cause a reduction in miR-48 levels. As a way to analyze and prioritize *mir-48* over-expression suppressors, miR-48 levels were measured in each suppressed strain.

miR-48 levels were measured using Taqman qPCR assays on RNA isolated from arrested L1 animals and miRNA levels were normalized to N2. The original array used in the mutagenesis screen *veIs48* (strain RG995) expressed levels of miR-48 that were more than 700 fold greater than wild type levels (Figure 4). Levels of miR-48 in the suppressed strains range from 32 to 665 fold greater than wild type, demonstrating that there is diversity in the types of suppressors isolated. Some suppressors appear to function by reducing miR-48 levels, while others function through a different mechanism. One *lin-66* allele (*ve130*) was included in the analysis; this allele did not affect miR-48 levels. These results are consistent with *lin-66* functioning downstream or in parallel to *mir-48* in the heterochronic gene pathway.

As a way to classify and prioritize the *mir-48* over-expression suppressors, GFP expression levels were analyzed together with miR-48 levels (Figure 4). Suppressor strains can then be sorted into three classifications: (1) high GFP and high miR-48, (2) high GFP and reduced miR-48, and (3) low GFP and low miR-48. As expected, strains with low GFP expression (≤ 3) also have low miR-

48 levels compared to *veIs48* and are likely multi-copy array suppressors. Suppressors with GFP levels ≥4 have a wide range of miR-48 levels. Identification of any of these could provide valuable information as to the function of *mir-48*. Those that are affecting miR-48 levels could be functioning upstream either in the heterochronic pathway or in the miR processing pathway of *mir-48*, while suppressors that do not affect miR-48 levels are likely functioning in parallel to *mir-48* in the heterochronic pathway or as a downstream target of miR-48.

Linkage mapping places three suppressors on Ch X and one suppressor on Ch 3

To map and identify the gene disrupted, linkage mapping was used to narrow down the location of the prioritized suppressors. Genetic markers were chosen that were located near the center of each chromosome and had an obvious phenotype that did not interfere with the *mir-48* over-expression Egl phenotype. In addition to the Egl phenotype, *veIs48* animals had a mild Dpy phenotype which made markers with subtle Dpy phenotypes unusable. Lon phenotypes did not work because they were often masked by the *veIs48* Dpy phenotype. Generally, Unc markers did not work well since Egl animals often had difficulties moving as they filled with embryos. Fortunately, a few markers worked well. Strong Dpy phenotypes worked the best to map *veIs48* suppressors. Markers that were used include *dpy-5(e61)* for Ch I, *dpy-17(e164)* for ChIII, *unc-5(e152)* for Ch IV, and *dpy-11(e224)* was used for Ch V. The *unc-5*

phenotype was successful in this analysis because it was a distinct coiling phenotype that could be scored during early larval stages. When suppressors were tested for linkage, the *ve150* was linked to *dpy-5* on Ch I, *ve148* was linked to *dpy-17* on Ch III, and *ve143* was linked to *unc-5* on Ch IV (Table 3). The *veIs48* array is integrated on the X-chromosome, making it rather simple to test for linkage to the X-chromosome. Suppressors that were linked to *veIs48* were *ve129*, *ve132*, *ve151* (Table 3).

SNP mapping

SNP mapping was used in an attempt to obtain more fine-scale mapping information. Single Nucleotide Polymorphism (SNP) mapping is an efficient way to narrow down the location of an unknown mutation, and has made significant contributions to genetic mapping studies since its inception to *C. elegans* in 2001 (WICKS *et al.* 2001). In this approach, an N2 (Bristol) strain with an unknown mutation is backcrossed to the CB4856 (Hawaiian) strain, single heterozygous F1 animals are allowed to self-fertilize, and then F2 animals are segregated into two pools: mutants and non-mutants. Forty-eight sets of primers located across the *C. elegans* genome are used for PCR. Each set of primers flanks a SNP that specifically alters affects a DraI restriction site (DAVIS *et al.* 2005) so that an amplified fragment from either the Bristol genome or the Hawaiian genome will be cleaved by DraI, but not both (DAVIS *et al.* 2005). PCR products are digested, run on a gel, and digest patterns compared between the mutant and non-mutant

pools. Since the mutation originated in the Bristol strain, the region near the mutation will be enriched for Bristol SNPs.

SNP mapping was performed with two suppressed strains, RG1437 and RG1432, using the *Davis et al.* method summarized above (DAVIS *et al.* 2005). Selecting for suppression required the *mir-48* over-expression array to be homozygous. Animals homozygous for the *veIs48* array (GFP+/+) had brighter *sur-5::gfp* expression than their heterozygous siblings, so bright GFP animals were single picked as young larvae. Once the animals were young adults, they were sorted into Egl and non-Egl pools. As recessive mutations, 25% of the GFP+/+ progeny from a heterozygous parent are expected to be non-Egl. When this was done with two suppressed strains, RG1437 and RG1432, 150 GFP+/+ animals were selected and all of the GFP+/+ animals were Egl making it impossible to continue the experiment. This experiment was repeated with RG1437, to make sure the first results were not an anomaly. 48 GFP+/+ animals were selected and again all were Egl. The next generation was scored for egg laying. Of the 48 F2s selected, 2 produced F3 animals that laid eggs. The remaining 46 F2s all produced Egl progeny. These results strongly suggest that there are SNPs/mutations in the CB4856 background that enhance the Egl phenotype. The enhancement is not specific to one suppressor since this occurred with both RG1437 and RG1432 suppressed strains, and the enhancement is not a maternal effect since the enhancement also occurred in the F3 generation. Additionally, for the enhancement to be so highly penetrant

there are likely to be multiple *veIs48* Egl enhancers in the CB4856 background. This approach was discontinued.

Since the CB4856 strain is enhancing the egg laying defect, another approach to doing SNP mapping is to use strains that contain a single chromosome from the CB4856 strain while the remaining chromosomes are from the Bristol N2 strain. This approach may work if the enhancer is limited to a certain number of chromosomes. Linkage mapping results (see above) linked the RG1437 suppressor to Ch III. So I performed SNP mapping of the RG1437 suppressor using strain WE5238 that only contained the CB4856 Ch III. *nhl-2*, a known suppressor of *veIs48*, is located on Ch III, so the strain VT1520 (*nhl-2(ok818); veIs48*) was used as a positive control, and since *lin-66* is located on Ch IV it was used as a negative control. The strains were crossed to WE5238, heterozygous F1s were single picked, and then GFP(+/+) F2s were single picked. From the VT1520 cross, 36 F2s were non-Egl (21%) and 139 F2s were Egl (79%). From the RG1437 cross, 39 F2s were non-Egl (21%) and 148 were Egl (79%). These results show that the Hawaiian Ch III does not enhance the Egl phenotype and allowed SNP mapping of Ch III to proceed. Eight sets of PCR primers that amplified regions of Ch III were used for PCR of RG1437 (DAVIS *et al.* 2005), and the products digested with DraI and analyzed on an agarose gel. Suppressed animals were compared to unsuppressed animals, and suppressed animals were highly enriched for the N2 digest pattern at the -12, -7, and -1 m.u.

position on Ch III (Figure 5). These results indicate that *ve148* is on the left arm of Ch III.

While SNP mapping one chromosome at a time is laborious, suppressors that have been linked to a chromosome from the linkage mapping results are good candidates for this approach. One of these candidates is RG1439 which was linked to Ch I. RG1439 was crossed to WE5236 which contains Hawaiian Ch I and the remaining chromosomes are from the Bristol N2 strain. GFP+/+ F2s were picked from a heterozygous F1, however only 13% of the F2s were Egl. These results were very surprising since most of the effect from the Hawaiian background was an enhancement of the Egl phenotype. Since the non-Egl pool of worm contained animals that were not homozygous for the suppressor and would confound the results, this experiment was not completed. In conclusion, the Hawaiian Ch I is not a good candidate to SNP map the *mir-48* over-expression suppressors.

SNP mapping was performed with RG1386 (*lin-66(ve130); veIs48*) to test whether Ch IV is a chromosome that can be mapped with SNP mapping, using WE5241. Egl and non-Egl pools of worms were isolated, and SNP mapping results showed strong linkage to 1, 8, 12, and 14 m.u. Primers at the 12 m.u. position showed the strongest linkage, consistent with *lin-66* being located on Ch IV at the 10.9 m.u. position. These results demonstrate that the Hawaiian Ch IV can be used to SNP map *mir-48* over-expression suppressors.

Whole Genome Sequencing

Since the cost of Whole Genome Sequencing continues to decrease, it has become a more widely used tool for mutation identification. For optimal results, it is useful to have multiple alleles of an unknown gene. Although it was not known whether multiple *veIs48* alleles have been isolated, three mapped to Ch X (*ve129, ve127, ve151*) and another was positioned within a small interval on Ch III (*ve148*), so I decided to sequence their genomes as a way to identify the suppressors. Importantly, each strain can serve as a control for polymorphisms in the RG995 strain relative to the *C. elegans* reference genome. The four strains were sequenced using the Illumina GAIIx sequencing machine and data analyzed using MAQGene (BIGELOW *et al.* 2009). The list of mutations for each strain was filtered for gene mutations on the chromosome of interest and not present in all four strains (Table 4).

Unfortunately, none of the mutations found in the X-linked strain were in the same gene. From the list of suppressor candidates, the most likely suppressors were *ncr-1* and *ain-1* based on known function. *ncr-1* is a transmembrane glycoprotein with a patched-like domain that is homologous to human NPC1 (LI *et al.* 2004). *ncr-1* and *ncr-2* function redundantly to suppress *daf-12,* and *daf-12* promotes expression of *mir-42/241/84. ain-1* is a member of RISC which functions with the mature miRNA to downregulate mRNA targets (DING *et al.* 2005). *ain-1* and *ncr-1* were tested for their abilities to suppress *veIs48* using existing alleles *ain-1(tm3681)* and *ncr-1(nr2023)*. Since these

mutations are linked to *veIs48,* it was first necessary to select recombinants. The alleles were crossed to *veIs48*. >200 GFP+/+ animals were selected from heterozygous animals, and non-egl animals were never found among their progeny. Either the alleles did not suppress *veIs48* or the mutations are tightly linked to *veIs48*. In addition, RNAi was performed with both *ncr-1* and *ain-1* in *veIs48* mutants and there was no suppression of the Egl phenotype. It is important to note, however, that there are a few caveats with RNAi. The *eri-1* and *rrf-3* RNAi sensitization mutations suppress the *veIs48* egl phenotype, so RNAi was done without these mutations. Additionally, two known suppressors of *veIs48*, *nhl-2* and *lin-66,* do not suppress the Egl phenotype through RNAi by feeding, so RNAi is not the best approach to test for *veIs48* suppression.

The list of *ve148* candidate suppressors on Ch III was tested by using existing alleles. Suppression was tested with *lon-1(e185)*, *nekl-4(tm4910)*, *wht-6(ok882)* , *sdz-15(tm2209)*, *rbf-1* (js232) and *mua-3(rh195)* but none of these alleles tested suppressed *veIs48*.

DISCUSSION & CONCLUSIONS

This work set out to identify genes that function with *mir-48*. Identifying five alleles of *lin-66* as suppressors of *mir-48* over-expression validated the screen as a way to identify new heterochronic mutations. *lin-66* is a novel protein that is composed of 627 amino acids. *lin-66* mutants have a retarded heterochronic defect, which contrasts with the precocious defect observed in *mir-*

48 over-expression mutants. *lin-66* was reported to repress *lin-28* post transcriptionally (MORITA and HAN 2006). *lin-28* is a heterochronic gene that promotes L2 larval program, and may be targeted by *let-7* family miRNAs (MORITA and HAN 2006). One possible model for the suppressive relationship between *lin-66* and *mir-48* is that they act in parallel to suppress *lin-28*. If high levels of *mir-48* are down-regulating *lin-28*, then loss of *lin-66* activity could relieve some repression on *lin-28* and restore L2 programs through *hbl-1*. Since *lin-66* has already been genetically characterized and placed in the heterochronic pathway, further work was not pursued.

Measurement of miR-48 levels within the suppressor strains has provided useful information into the types of suppressors identified. *lin-66(ve130)* had little to no effect on miR-48 levels, which supports the model that they function in parallel to promote the L2-to-L3 transition. Of the unidentified suppressors analyzed, *ve127* had the highest level of miR-48 expression. The difference in miR-48 levels between *ve127* and *lin-66(ve130)* is within the error of the qPCR assay and is likely not significant. Since *ve127* does not greatly affect miR-48 levels it could be a heterochronic gene that functions in parallel or downstream of *mir-48*. One gene that functions downstream of *mir-48* and is a target of miR-48 in the heterochronic pathway is *hbl-1*. *ve127* was mapped to the X-chromosome, which is the same chromosome *hbl-1* is linked to, however whole genome sequencing results did not identify an *hbl-1* mutation in RG1385. Of the list of candidate suppressors identified from sequencing, a mutation was found in *ain-1*,

which is a member of RISC and necessary to downregulate the targeted mRNA. Unfortunately, when an existing allele of *ain-1* was tested it did not suppress the *veIs48* egg-laying defect.

Interestingly some of the suppressors have about a 50% reduction in miR-48 levels, including *ve131, ve136, ve149, ve129*, whereas others have a more drastic reduction including *ve150, ve145, ve148, ve151, ve132*. Suppressors that affect miR-48 levels could function upstream of *mir-48* in the hetereochronic pathway, for example they could be alleles of *daf-12*. *daf-12* encodes a nuclear hormone receptor which promotes *mir-48* expression during continuous development. Another type of suppressor that could affect miR-48 levels includes miRNA processing or stabilizing genes. Since miRNA biogenesis is a multi-step process, there are many ways to regulated miRNA accumulation. One regulator of miRNA biogenesis is the heterochronic gene *lin-28,* which inhibits *let-7* processing by binding to the precursor miRNA (LEHRBACH *et al.* 2009). Surprisingly, most of the suppressors analyzed showed a reduction in miR-48 levels and therefore could include multiple regulators of miRNA biogenesis.

miR-48 levels were found to be reduced by 50% in the *nhl-2(ok818); veIs48* strain. *nhl-2* is a known suppressor of *veIs48* and is cofactor of miRNA function (HAMMELL *et al.* 2009b). This was rather surprising because *nhl-2(0)* does not function to regulate miRNA expression or processing. Rather, NHL-2 interacts with the miRISC complex after it is bound to the targeted mRNA and has been shown to further silence the targeted mRNA. However, my results

suggest that miR-48 stability is affected when *nhl-2* is disrupted. It is possible that the some of the suppressors that reduce miR-48 levels are involved in miRNA stability.

While a lot of information was gathered about the suppressors, a few challenges made it difficult to identify the mutations responsible. Since most of the suppressors did not have an obvious phenotype the *veIs48* array was required to be in the background, and unfortunately the *veIs48* background was sensitive to enhancement or suppression. SNP mapping is a powerful way to narrow down the location of an unknown mutation because it is quick and inexpensive. Unfortunately, since mutations in the Hawaiian background suppressed and enhanced the Egl phenotype, the way in which SNP mapping could be used was limited. Mapping could be done by using a single Hawaiian chromosome or linkage mapping but it was a less efficient process and limited the number of suppressors that could be mapped.

Identifying genes that interact with and regulate miR-48 remains important to study because of it can lead to the identification of new heterochronic mutations, genes involved in miRNA regulation and function, and *let-7* family miRNA targets. Further analysis of the *veIs48* suppressors may still lead to the identification of a suppressor. Alternatively, other approaches may be beneficial to explore. Since mapping suppressors was a challenge, a suppressor screen that could lead to the identification of other genes more quickly may be useful. Unfortunately, an RNAi screen is not a reasonable alternative since known

veIs48 suppressors do not suppress through RNAi, and the RNAi sensitization mutations suppress the Egl defect. Another option that could be explored is a transposon-mediated insertional mutagenesis screen. The technique has been developed using a Mos1 transposon that originated in flies and is not endogenous to *C. elegans* (BOULIN and BESSEREAU 2007; FROKJAER-JENSEN *et al.* 2014). Rather than using EMS to cause single nucleotide changes, genes are disrupted when a transposon inserts itself in the middle of the gene or an important regulatory region. While the mutation rate is less efficient, it drastically cuts down on the time it takes to identify the suppressor, since the gene can be sequenced using a primer within the transposon.

Through analysis of the suppressors, we've learned that *mir-48* over-expression can be suppressed through parallel or downstream mechanisms without affecting miR-48 levels, like with *lin-66*, or through upstream mechanisms that affect miR-48 levels. Future work investigating *mir-48* is promising and could further our understanding of how miRNAs are regulated and processed.

Figure 1: *mir-48* is a *let-7* family microRNA and functions redundantly with *mir-241* and *mir-84*. A. The Let-7 family of miRNAs share a common seed region (highlighted region), which consists of the first 8 nucleotides. Let-7 is perfectly conserved between *C. elegans* and humans. **B.** Seam cell lineages of wild type and *let-7* family mutants. *mir-48(lf)* mutants do not have a seam cell defect. *mir-48/241/84(0)* mutants reiterate the L2 stage proliferative division. Over-expression of *mir-48* (*mir-48(+++)*) causes a precocious defect because the L2 proliferative division is often skipped. **C.** Animals that over-express *mir-48* on an integrated multicopy array (*veIs48*) with *sur-5::gfp* have an EGg-Laying defective (Egl) phenotype. *veIs48* worms were mutagenized with EMS, and suppressors (*veIs48;sup*) identified based on restoration of egg-laying (Non-egl).

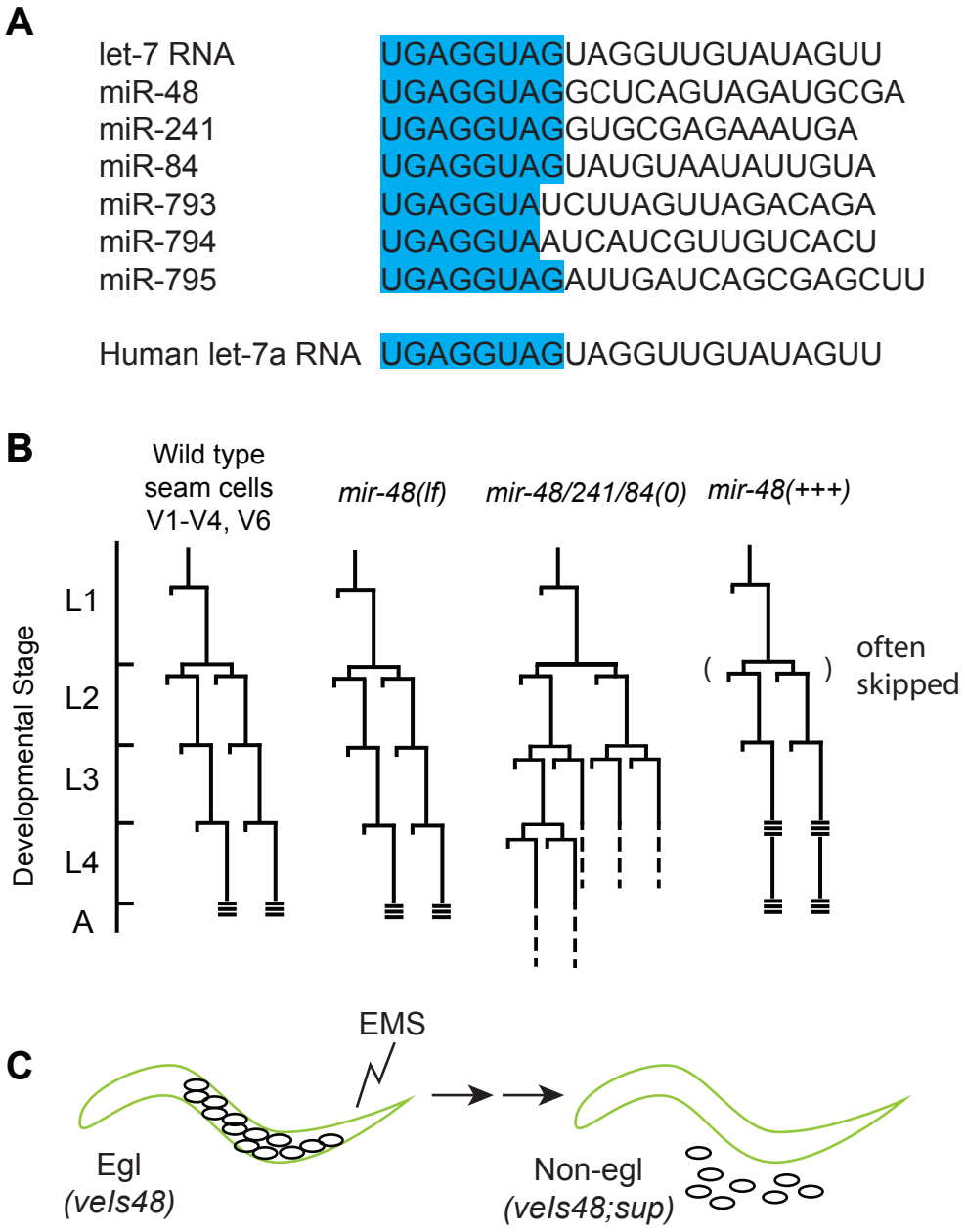
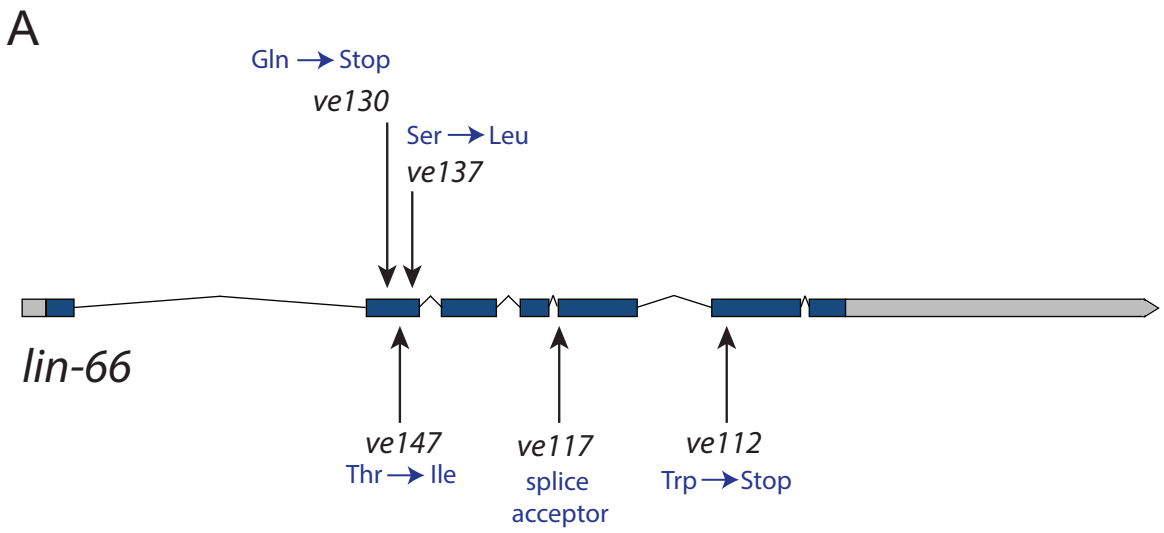

A
let-7 RNA UGAGGUAGUAGGUUGUAUAGUU
miR-48 UGAGGUAGGCUCAGUAGAUGCGA
miR-241 UGAGGUAGGUGCGAGAAAUGA
miR-84 UGAGGUAGUAUGUAAUAUUGUA
miR-793 UGAGGUAUCUUAGUUAGACAGA
miR-794 UGAGGUAAUCAUCGUUGUCACU
miR-795 UGAGGUAGAUUGAUCAGCGAGCUU
Human let-7a RNA UGAGGUAGUAGGUUGUAUAGUU
B
Wild type seam cells V1-V4, V6
mir-48(lf)
mir-48/241/84(0)
mir-48(+++)
Developmental Stage
L1
L2
L3
L4
A
( )
often skipped
C
EMS
Egl
(vels48)
Non-egl
(vels48;sup)

Figure 1

Figure 2. Five *mir-48* over-expression suppressors were identified as alleles of *lin-66*. A. *lin-66* alleles include missense mutations, premature stop codons, and a splice acceptor mutation. **B.** *lin-66* functions in the heterochronic pathway to inhibit *lin-28* expression. *lin-66* and *mir-48* function in parallel promote the L2-to-L3 transition.


A
Gln → Stop
ve130
Ser → Leu
ve137
lin-66
ve147
Thr → Ile
ve117
splice
acceptor
ve112
Trp → Stop

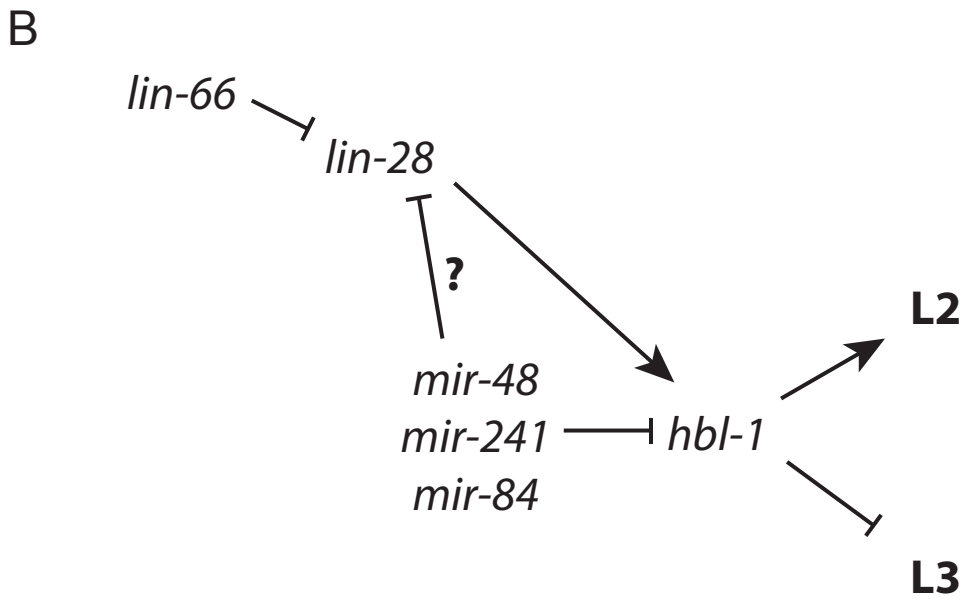
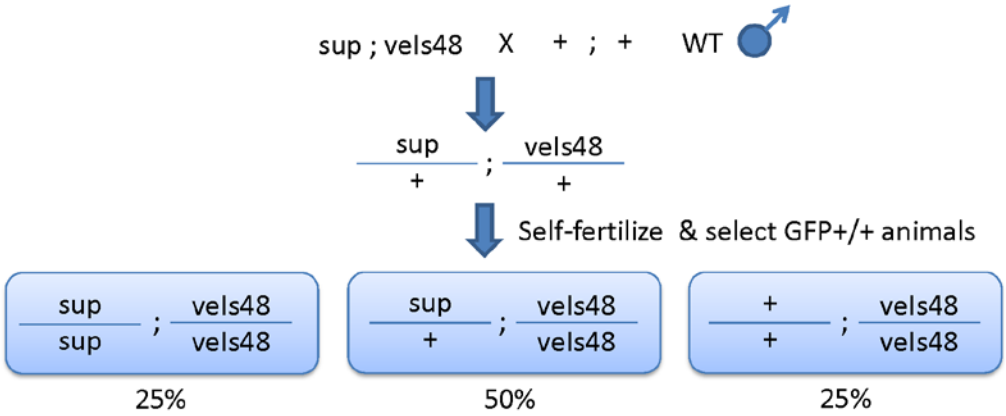

B
lin-66
lin-28
?
mir-48
mir-241
mir-84
hbl-1
L2
L3

Figure 2

Figure 3. Maternal effect analysis identifies three suppressors with a maternal effect. A. Suppressed strains were crossed to WT males, heterozygous F1 animals were allowed to self-fertilize, and individual bright green F2 animals that were homozygous for the *veIs48* array were selected. F2 animals were scored for egg-laying. The entire F3 and F4 generations were raised on the same plate as the F2 animal and collectively scored for egg laying. If a gene has a maternal effect, an animal will only be non-Egl if the parent is homozygous for the suppressor. 25% of F2 animal will be homozygous for the suppressor and have an Egl phenotype, but produce non-Egl progeny. 50% of F2 animals will be heterozygous for the suppressor and have an Egl phenotype, but produce F3 animals that are homozygous for the suppressor. Those F3 animals will be Egl but produce non-Egl F4 animals. Finally, 25% of F2 animal will not contain the suppressor and will never produce non-egl progeny. **B.** The percentage of egg-laying scored at each generation. Suppression is expected to been seen in a non-maternal gene from 25% of the F2s and 75% of the plates will have F3 and F4 egg-layers. Suppression from a maternal gene will be expected to be seen in 0% of the F2 generation, 25% of the F3 generation, and 75% of the F4 generation.



	Expected	Maternal	RG1431	RG1432	RG1434	RG1439
F2	25%	0%	0%	0%	25%	0%
F3	75%	25%	20%	33%	75%	0%
F4	75%	75%	87%	73%	81%	77%
n			15	15	16	13

Figure 3

Figure 4. Evaluation of GFP and miR-48 expression among the *mir-48* overexpression suppressors. All strains contain the *veIs48* array which expresses *mir-48* and *sur-5::gfp* except for wild type. Fold change miR-48 was measured by Taqman qPCR (see materials and methods). GFP levels were scored by eye on a GFP dissecting microscope on a scale of 0 to 5 (0=no visible GFP, 5=GFP expression as bright as the strain containing *veIs48* (RG995)).

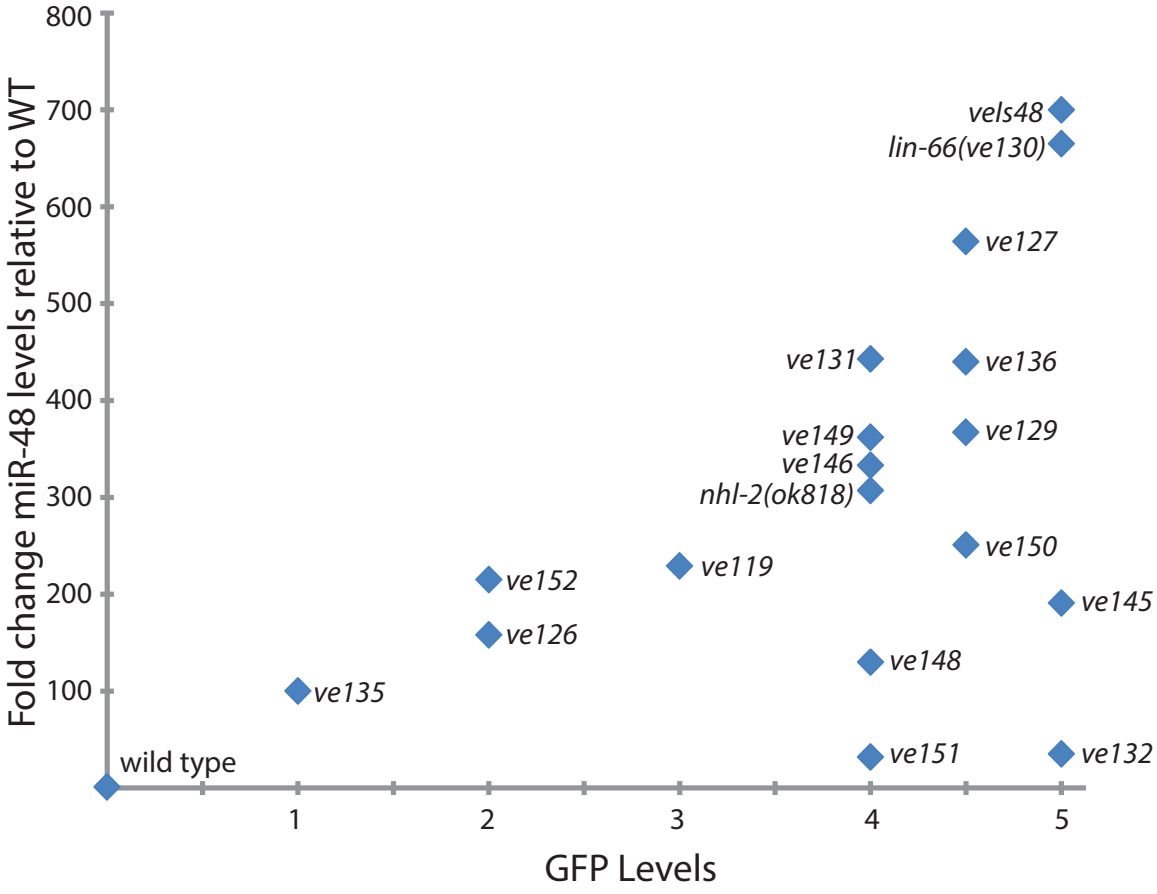

Fold change miR-48 levels relative to WT
800
700
600
500
400
300
200
100
wild type
ve135
ve152
ve126
ve119
ve131
ve149
ve146
nhl-2(ok818)
vels48
lin-66(ve130)
ve127
ve136
ve129
ve150
ve145
ve148
ve151
ve132
1
2
3
4
5
GFP Levels

Figure 4

Figure 5. SNP Mapping places *ve148* on Ch III. Regions flanking a SNP were amplified by PCR, digested with Dra I, and analyzed on an agarose gel pictured here. Chromosome III positions analyzed with SNP mapping are indicated at the top by map unit (M.U.) position (1 M.U. = 275 kb). Samples are compared between Egl (E) or Non-egl (N) pools of worms. Genotype for each strain analyzed are listed to the left of each gel image. Linkage is indicated when the Bristol product (green star) is more enriched than the Hawaiian product (orange square) in the non-Egl sample. The strain with *ve148* is enriched for the Bristol product at M.U. positions -12, -7, -1. *nhl-2 (ok818)* is known to be located at position -2.4 M.U, and as expected SNP mapping results show the strongest Bristol enrichment at the -1 M.U. position. The *lin-66(ve130)* mutation is on Ch IV, and SNP mapping results show no Bristol enrichment on Ch III. DNA fragments that associate with either Bristol (B) or Hawaiian (H) are indicated on the *lin-66(ve130); veIs48* gel picture. * indicates a fragment that associates with both Bristol and Hawaiian.

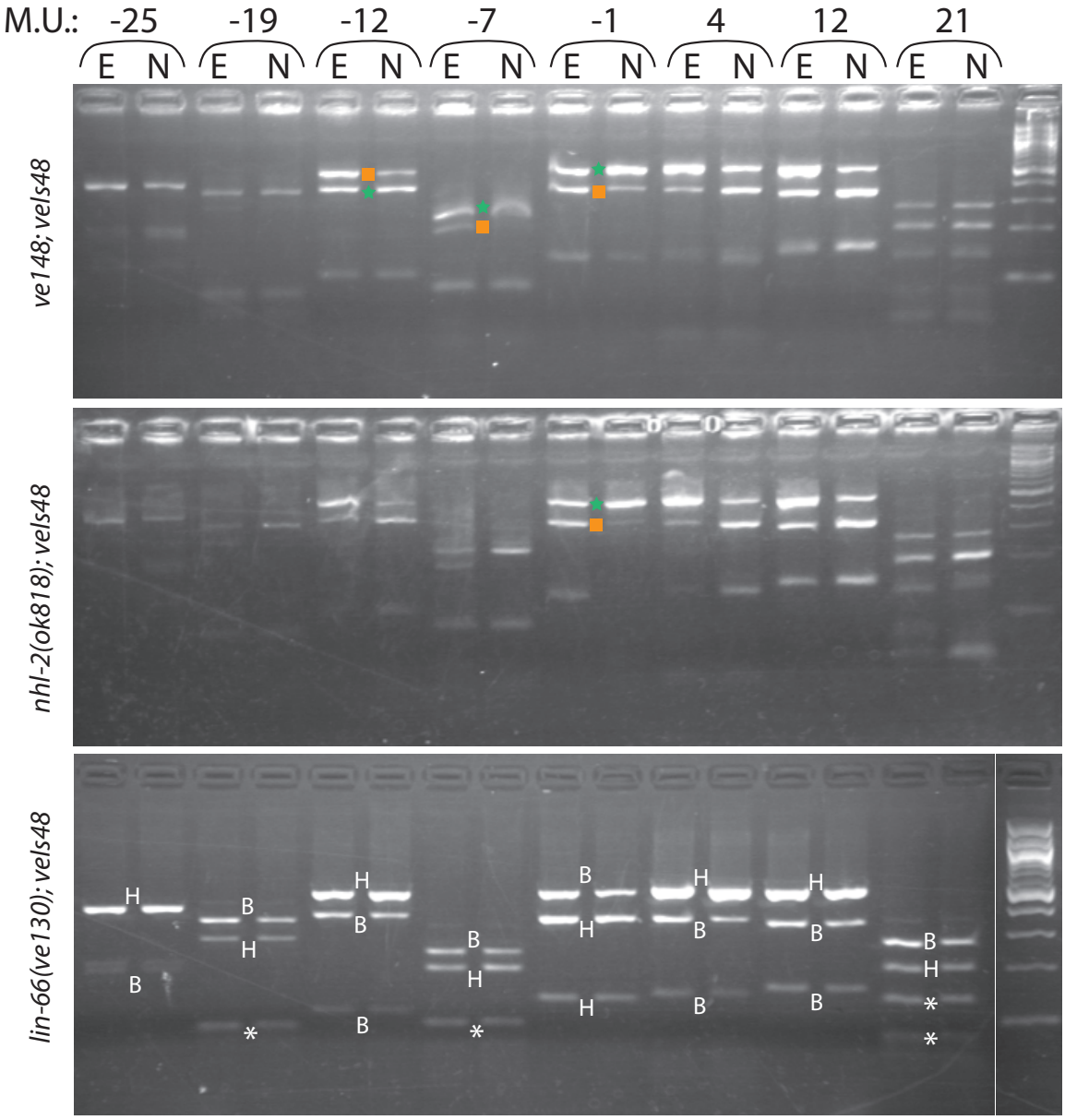

M.U.: -25 -19 -12 -7 -1 4 12 21
E N E N E N E N E N E N E N E N
ve148; vels48
nhl-2(ok818); vels48
lin-66(ve130); vels48
H
B
*

Figure 5

Table 1: *vels48* suppressor strains

Strain Name[1]	Suppressor		GFP expression[3]
	allele	gene[2]	
N2			0
RG995			5
VT1520	*ok818*	*nhl-2*	4
RG1364	*ve111*		
RG1365	*ve112*	*lin-66*	
RG1366	*ve117*	*lin-66*	
RG1367	*ve118*		
RG1368	*ve113*		
RG1369	*ve114*		
RG1370	*ve115*		
RG1371	*ve116*		
RG1372	*ve119*		3
RG1383	*ve129*		4.5
RG1384	*ve126*		2
RG1385	*ve127*		4.5
RG1386	*ve130*	*lin-66*	5
RG1387	*ve128*		
RG1388	*ve131*		4
RG1389	*ve132*		5
RG1420	*ve133*		
RG1421	*ve134*		
RG1422	*ve135*		1
RG1423	*ve136*		4.5
RG1424	*ve137*	*lin-66*	
RG1425	*ve138*		
RG1428	*ve139*		
RG1429	*ve140*		
RG1430	*ve141*		
RG1431	*ve142*		
RG1432	*ve143*		
RG1433	*ve144*		
RG1434	*ve145*		5
RG1435	*ve146*		4
RG1436	*ve147*	*lin-66*	
RG1437	*ve148*		4
RG1438	*ve149*		4
RG1439	*ve150*		4.5
RG1440	*ve151*		4
RG1441	*ve152*		
RG1442	*ve153*		
RG1443	*ve154*		
RG1444	*ve152*		3
RG1445	*ve151*		

[1]Every strain listed includes vels48, except N2.
[2]Gene affected by suppressor mutations
[3]GFP expression analyzed by eye on a 0-5 scale.

Table 2: Penetrance of the suppressors

	non-Egl (%)	n
RG995	0	20
RG1434	72	18
RG1437	70	37
RG1439	37	19
RG1440	95	38
RG1442	73	15

Table 3: Two-factor mapping links suppressors to chromosomes

	Chromosome					
Genotype	I	II	III	IV	V	X
lin-66(ve130); vels48	No		No			
ve129; vels48						Yes
ve132; vels48						Yes
ve142; vels48	No		No			No
ve143; vels48			No	Yes	No	No
ve145; vels48	No		No		No	
ve148; vels48			Yes	No	No	
ve150; vels48	Yes		No	No		No
ve151; vels48			No			Yes
ve153; vels48			No		No	

Table 4: Suppressor candidates identified from Whole Genome Sequencing

RG1383 (Ch X)	RG1385 (Ch X)	RG1440 (Ch X)	RG1437 (Ch III)
T04G9.6	mir-271	F40F4.6	lon-1 *
ncr-1*	sax-3	ZK563.5	nekl-4*
mec-2	syx-3	C31H2.3	wht-6 *
C15H9.7	gln-1	mec-2	sdz-15*
pept-1	C15H9.5.1	F31B12.2a	rdf-1 *
apy-1	C15H9.9	lpr-4	C48B4.1
C18A11.6	C34D10.2	hum-4	mua-3*
M02D8.3	**fln-2**	W06D11.5	
git-1	ZK899.7	F42F12.4	
alr-1	M79.2	**C49F8.1**	
C35C5.6	Y62H9A.9	nhr-71	
F55G7.1	F02C12.4	**phat-6**	
pqn-18	bra-1	bus-17	
hda-4	K11E4.3	F01G12.6	
crb-1	mrp-5		
	T10B10.4b		
	F53H4.2		
	ain-1 *		
	F52G3.3		

Bold = Pre-mature stop
* = Genetic analysis

CHAPTER IV: PROSPECTUS

SUMMARY

The research described in my thesis investigates the role of *lin-42* in *C. elegans* development. I show that the *lin-42(0)* mutants have a precocious heterochronic and molting defect, which is more penetrant than hypomorphic alleles. I also show that over-expression of *lin-42a* or *lin-42b* strongly rescues *lin-42(0)* defects, while over-expression of *lin-42c* mildly rescues *lin-42(0)* defects. These results indicate that all three *lin-42* transcripts provide some level of *lin-42* function. I also compare expression patterns of the three *lin-42* transcripts and show that all three transcripts cycle in unison during each larval stage. In this chapter I discuss that significance of this work and propose future directions.

PROSPECTUS AND FUTURE DIRECTIONS

Future investigations with the *lin-42(0)*.

Now that a null allele of *lin-42* exists and has been characterized there are many ways it can be used to more thoroughly investigate the role *lin-42* plays to regulate developmental progression in both the heterochronic and molting pathways. Previous to my work, *lin-42* phenotypic and rescue analysis was performed with hypomophic alleles that left one LIN-42 isoform intact (Figure 1A). I show that the *lin-42(0)* is more severe than either of the deletion mutants, *lin-42(n1089)* and *lin-42(ok2385)*, which demonstrates that the intact isoform

present in either mutant performs some level of *lin-42* function (Figure 1A). The null provides the premier tool necessary to investigate the function of *lin-42*.

The *lin-42(0)* is the long awaited tool needed to perform genetic epistatic experiments to help place *lin-42* in the heterochronic pathway. *lin-42* has been shown to play an important role in regulating expression of primary miRNA expression (McCulloch and Rougvie 2014; Perales *et al.* 2014; Van Wynsberghe *et al.* 2014). Dr. McCulloch, while in the Rougvie lab, used the *lin-42(0)* to perform genetic epistatic analysis with the *let-7* family miRNAs and demonstrated that *lin-42* is epistatic to *let-7* and *mir-48* (McCulloch and Rougvie 2014). *lin-42* has been suggested to play a more global role in regulating miRNA and mRNA expression, and the *lin-42(0)* can be used to test those interactions by measuring expression levels of proposed targets in the absence of *lin-42* and investigating genetic relationships between *lin-42* and possible targets.

The phenotypic analysis that I performed in my thesis demonstrates that the *lin-42(0)* mutants have a precocious heterochronic defect, as well as a molting defect. *lin-42* has been shown to genetically interact with two genes that function to regulate molting, *nhr-23* and *nhr-25 (Kostrouchova et al. 1998; Gissendanner and Sluder 2000; Kostrouchova et al. 2001; Frand et al. 2005; Kouns et al. 2011)*. One possibility is that *nhr-23* and *nhr-25* regulate *lin-42* expression because *nhr-23* and *nhr-25* are nuclear hormone receptors with an oscillating expression pattern like *lin-42 (Kostrouchova et al. 2001;*

GISSENDANNER et al. 2004). Genetic epistasis analysis will be important to determine the relationship between *lin-42*, *nhr-23* and *nhr-25* and how they work together to regulate molting. Furthermore, measurement of *lin-42* levels in *nhr-23/25* mutants using the qPCR assays I designed in Chapter 2 of my thesis, can determine whether *nhr-23/25* affects *lin-42* expression.

In addition to genetic epistasis analysis, *lin-42(0)* can be used to determine which tissues are important for *lin-42* function. I performed a tissue specific rescue experiment to determine whether *lin-42* functions in the hypodermis by over-expressing *lin-42a* or *lin-42b* under control the *dpy-7* promoter. The *dpy-7* promoter stimulates expression predominantly in the main body hypodermis (hyp7) and occasionally in the hypodermal seam. Results from this experiment suggested that expression in the hypodermal tissues is sufficient to rescue the precocious heterochronic defect, whereas the molting defect was not rescued. However, since these results were generated, a caveat was found. We observed that over-expression of the *Pscm::GFP* and *Pmlt-10::GFP* transcriptional reporters suppressed the precocious alae defect observed in the *lin-42(0)* (see Chapter 2). Both of these promoters promote expression in hypodermal tissues: the *Pscm::GFP* in the seam cells and *Pmlt-10::GFP* in the main body hypodermal syncytium. It is possible that these promoters are titrating out an important factor required for precocious alae formation. These results leave open the possibility that the *dpy-7* promoter, which controls expression in the hypodermis, may also be sufficient to suppress the precocious alae defect. A

follow-up experiment needs to be done to test whether the *dpy-7* promoter alone rescues the *lin-42(0)* precocious alae defect. These results will determine whether the tissue specific rescue of the alae defect is indeed because of *lin-42* or because of an effect of the *dpy-7* promoter. If the *dpy-7* promoter has an effect of the alae phenotype, then another promoter needs to be used to test for hypodermal rescue. The *Pelt-5::his-24::mCherry* reporter which is expressed in the seam and hypodermis has been shown to not affect the *lin-42(0)* precocious alae defect, thus *elt-5* promoter could be used to test for rescue.

Over-expression of *lin-42a* or *lin-42b* in the hypodermis did not rescue the molting defect, therefore further investigation into the tissue responsible for control of molting is needed. One possibility is the neuroendocrine system because it has been shown to regulate lethargy and ecdysis (VAN BUSKIRK and STERNBERG 2007; SINGH *et al.* 2011). The ALA interneurons and ciliated sensory neurons are neural cells important for regulating molting, and would be good candidates to test for rescue of the molting defect (VAN BUSKIRK and STERNBERG 2007; SINGH *et al.* 2011). The *lin-42* molting defect is observed only when the downstream region containing the SYQ and LT domains is deleted, as in *lin-42(ok2385)* and *lin-42(ox461)* mutants, and suggests this region plays an important role in molting. It will be important to test both *lin-42a* and *lin-42b* for rescue of the molting defect in neurons by single copy rescue.

Overall, experiments using the *lin-42(0)* will be significant in investigating the role *lin-42* plays in the heterochronic and molting pathways through epistatic analysis and rescue analysis.

Structure function analysis of LIN-42

LIN-42 Protein Expression Analysis. Not only is the genetic locus of *lin-42* complex in that it expresses two non-overlapping isoforms, but the *lin-42* expression pattern is complex because it cycles with each larval stage. This cycling is similar to the way *period* cycles with the light-dark cycle. In my thesis work, I investigated the expression pattern of each *lin-42* transcript. Since *lin-42a* and *lin-42b* are expressed from two different promoters there is a possibility that their expression patterns differ. There was also evidence in the field that suggested one transcript was expressed earlier than the other, however two separate reports disagreed about which one (MONSALVE *et al.* 2011; PERALES *et al.* 2014). My thesis work examined the expression pattern of each transcript using qPCR and provided a detailed analysis that demonstrated that each transcript cycles in unison before each larval molt. Since PERIOD is known to regulate its own transcription through a negative feedback loop, I also looked at *lin-42* expression patterns in hypomorphic alleles and showed that *lin-42a* and *lin-42c* expression do not require expression of the other transcripts. While we now have a detailed report of *lin-42* mRNA expression, less is known about the expression of LIN-42 protein.

Whole mount antibody staining using an antibody that recognizes both LIN-42A and LIN-42B demonstrated that they cycle with each molt, similar to the *lin-42* transcripts (TENNESSEN *et al.* 2006). Antibody staining also showed that LIN-42A and LIN-42B are nuclearly localized and expressed in many tissues, including the hypodermis, vulva, somatic gonad, distal tip cell, intestine, muscle and neurons. The information currently available describing LIN-42 expression does not provide a comparison between the different isoforms. Using CRISPR, fluorescent tags can be introduced to different locations in the *lin-42* locus as a way to label different isoforms.

Analysis of individually tagged isoforms, will allow temporal and spatial expression of each isoform to be compared. Tags that would be useful include an N-terminal tag that label the PAS containing isoforms, LIN-42B and LIN-42C, as well as an N-terminal tag that labels the non-PAS containing isoform LIN-42A. I introduced an N-terminal tag to label LIN-42B and LIN-42C. Sequence analysis confirmed it was inserted correctly into the genome, however fluorescence was not visible, possibly because the tag is being masked or cleaved. Future work introducing a tag at the C-terminus of LIN-42C and the N-terminus of LIN-42A could differentiate expression of LIN-42C and LIN-42A (Figure 1B). Additionally, a tag to specifically label LIN-42B could be placed in the large intron between exon 5 and 6 and engineered to be spliced with LIN-42B.

I began work to introduce a C-terminal tag which would label LIN-42A and LIN-42B, since C-terminal tags have been used successfully in over-expression

studies (TENNESSEN *et al.* 2006). While a C-terminal tag will not distinguish the expression patterns between LIN-42A and LIN-42B, it can determine if endogenous expression can be visualized with a GFP tag and used in subsequent targeted mutagenesis to ensure protein stability is not affected. With the help of others in the lab, a C- terminal GFP tag was successfully introduced into the genome. The tagged protein displayed a noticeable expression pattern, similar to what was observed from antibody staining, and demonstrated that endogenous LIN-42 expression can be visualized. Mutations can now be made to the endogenous locus using CRISPR and the fluorescent expression can be used to verify that protein function is affected and not just being degraded.

lin-42 targeted mutagenesis. Since the development of the CRISPR/Cas-9 system in *C. elegans*, it has become rather straightforward to make specific and targeted modifications to the genome (DICKINSON *et al.* 2013; FRIEDLAND *et al.* 2013). The process requires the Cas9 nuclease to be guided to a specific genomic locus by a single guide RNA (sgRNA). The sgRNA can be modified to target virtually any location in the genome that has an NGG motif. Specific changes can be made by providing a repair template that can be used to repair the double-stranded break (DSB) created by the Cas-9 cleavage enzyme.

Before CRISPR-Cas9 was developed, specific mutations were often tested by over-expressing mutated genes at levels much higher than what is found in nature. In the case of *lin-42*, over-expression of *lin-42a* rescued both upstream and downstream mutations, in addition to the null allele. However, at

endogenous levels *lin-42a* is not sufficient, since mutations in the upstream locus cause a phenotype (TENNESSEN *et al.* 2006). These results demonstrate the importance of analyzing *lin-42* at endogenous levels. CRISPR-Cas9 technology was developed towards the end of my thesis work and is why my work investigated *lin-42* through over-expression studies. CRISPR-Cas9 will be a powerful tool to use in the future to analyze specific mutations. Since the *lin-42* locus has recently been fluorescently tagged (as discussed above) it can be used to introduce mutations to the genetic locus and ensure LIN-42 stability has not been affected. The fluorescent tag can also indicate whether mutations affect protein localization.

Surprisingly, not one LIN-42 missense allele has been isolated that has a phenotype. These results have been rather perplexing since missense mutations cause phenotypes in PERIOD. In future experiments, multiple bases should be mutated when making targeted changes. LIN-42 is a homolog of PERIOD and they share several homology domains, including the PAS, SYQ and LT domains, which are excellent candidates for targeted mutagenesis.

Since dimerization through the PAS domain is essential for PERIOD function (HENNIG *et al.* 2009), it is important to establish whether dimerization is also essential for LIN-42 function. A single conserved tryptophan in PERIOD has been shown to be essential for dimerization. Disruption of the same Trp residue in LIN-42(W258E) has been shown to prevent dimerization *in vitro* (personal communication, C. Partch USC). To test whether that residue is essential in *C.*

elegans, I made the W258E muation using CRISPR. Surprisingly, no phenotype was observed which suggest that LIN-42 dimerization is not necessary for function. To investigate this further, it is important to test the W258E mutation in a sensitized background, for example scoring *lin-42(W258E)/lin-42(0)* trans-heterozygotes. Since *lin-42a* is not affected in *lin-42(W258E)* mutations, a reduction in endogenous levels of *lin-42a* may reveal a phenotype. The conservation of this residue, as well as the PAS domain, indicates that they serve important functions. If the W258E mutation does not have an effect on LIN-42 function, it is possible that other residues in the PAS domain contribute to function. Other highly conserved residues in the PAS domain could be simultaneously mutated to alanines, or an in-frame deletion could be introduced. The CRISPR sgRNA designed when the W258E mutation was introduced has also been used to introduce small deletions with the goal of identifying an in-frame deletion, however only frame-shift deletions have been identified. If frameshift mutations are not identified using the sgRNA, a targeted deletion could be made using a repair template. Identification of an in-frame deletion that affects the PAS domain would help determine whether it is necessary for function. Dimerization through the PAS domain is required for PERIOD nuclear localization. With the fluorescent tag, we could determine whether LIN-42 protein localization is also affected in the PAS mutant.

In addition to the PAS domains, the SYQ and LT domains are also conserved between LIN-42 and PERIOD. The SYQ and LT domains are small

conserved domains that can be mutated using CRISPR, by converting the highly conserved residues to alanine. Alternatively, small in-frame deletions could be isolated to determine whether this region is important to LIN-42 function. While regions including the SYQ and LT domains have be implicated as protein interaction domains in flies and mammals, no function has been directly assigned to these domains (CHANG and REPPERT 2003; LEE *et al.* 2004a; SUN *et al.* 2010). Over-expression of *lin-42a* can rescue all *lin-42* alleles, which indicates that *lin-42a* serves an important function (See Chapter 2) (TENNESSEN *et al.* 2006). If mutations/deletions in the SYQ/LT domain cause a phenotype, it would be exciting to compare it to other alleles, such as *lin-42(ve11)* and *lin-42(ok2385)* to determine whether there are other regions in *lin-42a* that are important for function.

There is strong evidence that both PAS domain and the region containing the SYQ/LT domains are important for function. Deletion of either the upstream or the downstream region of the gene results in a strong defect. While the PAS domain is the region of greatest homology between LIN-42 and PERIOD, LIN-42A, which does not contain a PAS domain, can rescue *lin-42(n1089)* and *lin-42(0)* when over-expressed. An interesting question to investigate is whether LIN-42A and LIN-42C function without LIN-42B. We know that *lin-42a* and *lin-42c* can function in the absence of *lin-42b* when overexpressed, since overexpression of *lin-42a* can rescue *lin-42(ok2385)* and overexpression of *lin-42c* can rescue *lin-42(n1089)*. In both cases, *lin-42b* is absent. However, the

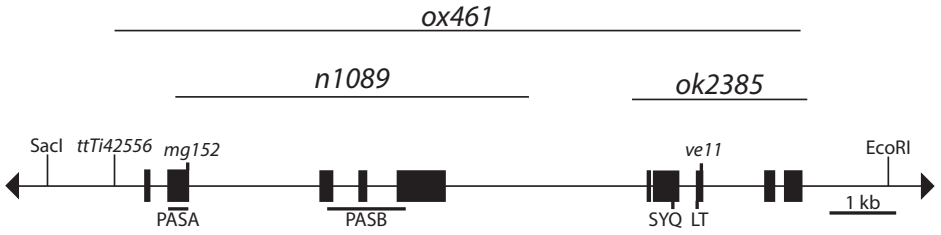
collective function of *lin-42a* and *lin-42c*, in the absence of lin-42b, has not been tested at endogenous levels. To determine if LIN-42A and LIN-42C can function in the absence of LIN-42B, *lin-42a* could be introduced as a single-copy transgene at another genomic location and tested for rescue of the *lin-42(ok2385)* allele (Figure 1C). These results would determine whether *lin-42c* is necessary.

Another question to investigate is whether LIN-42B is sufficient to function without the two non-overlapping isoforms. We know LIN-42B is sufficient when over-expressed (see Chapter 2), however it has not been tested at endogenous levels. To test this, CRISPR can be used to remove the large intron between exons 5 & 6 which removes the *lin-42a* promoter and the second half of exon 5 required for *lin-42c* (Figure 1C). If this deletion rescues, it would indicate that *lin-42b* is sufficient. Since PERIOD is expressed as one protein that contains all conserved domains, my hypothesis is that *lin-42b* is necessary and sufficient.

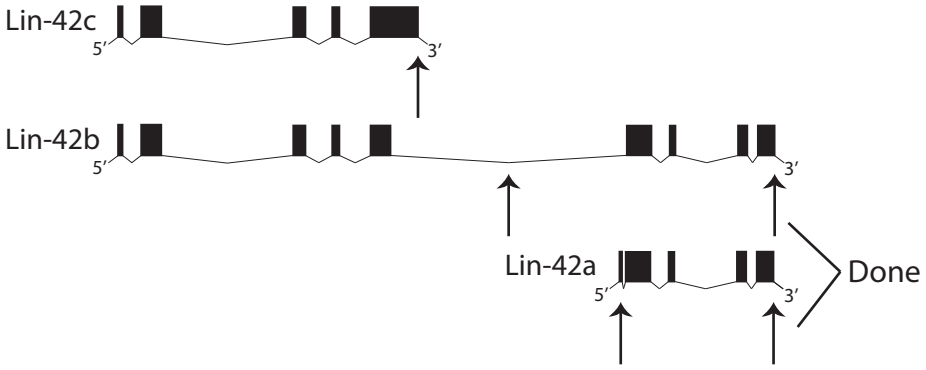
My work investigates the role for *lin-42* in development through analysis of a *lin-42(0)* allele. *lin-42* plays multiple roles to regulate temporal and molting programs, and all three LIN-42 isoforms contribute to LIN-42 function. Further work investigating the conserved domain structures found in each isoform will help elucidate the mechanisms of *lin-42* function.

Figure 1. *lin-42* locus and experimental diagrams. A. *lin-42* locus. *lin-42(ox461)* removes the entire coding region. *lin-42(n1089)* removes the upstream region of the gene that contains the PAS domains. *lin-42(ok2385)* removes the downstream region of the gene that removes the SYQ and LT domains. **B.** Locations to introduce fluorescent tags. LIN-42C can be tagged on the N-terminus. LIN-42B can be tagged by inserting a tag in to the large intron between exons 5 and 6, and engineered to be spliced. LIN-42A can be tagged on the C-terminus. A tag has been introduced on the C-terminus that labels both LIN-42A and LIN-42B. **C.** Design for rescue experiments. *lin-42a* and *lin-42c* can be tested in trans by inserting *lin-42a* into another location of the genome using CRISPR. To determine if *lin-42b* is sufficient for *lin-42* function, the large intron between exons 5 and 6 can been removed.

A lin-42 locus

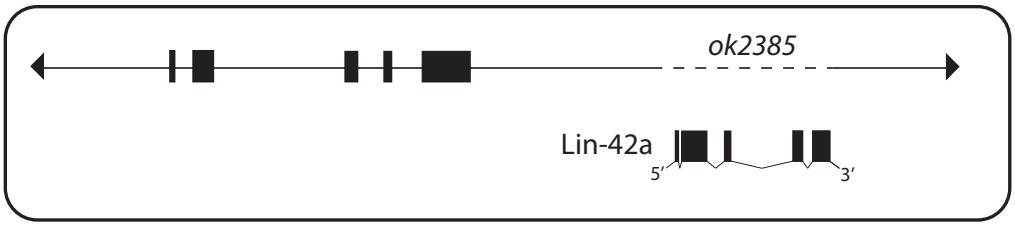


B Fluorescently tag isoforms

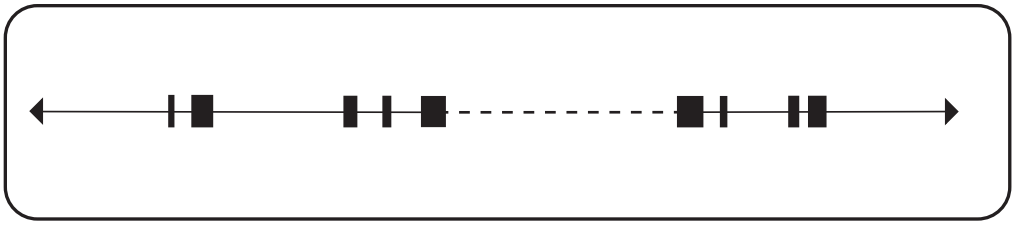


C Test for rescue

Do *lin-42a* and *lin-42c* function in trans?



Is lin-42b sufficient?



CITATIONS

Abbott, A. L., E. Alvarez-Saavedra, E. A. Miska, N. C. Lau, D. P. Bartel *et al.*, 2005 The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. Dev Cell 9**:** 403-414.

Abrahante, J. E., A. L. Daul, M. Li, M. L. Volk, J. M. Tennessen *et al.*, 2003 The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev Cell 4**:** 625-637.

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

Aguinaldo, A. M., J. M. Turbeville, L. S. Linford, M. C. Rivera, J. R. Garey *et al.*, 1997 Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387**:** 489-493.

Akten, B., E. Jauch, G. K. Genova, E. Y. Kim, I. Edery *et al.*, 2003 A role for CK2 in the Drosophila circadian oscillator. Nat Neurosci 6**:** 251-257.

Ambros, V., 1989 A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. Cell 57**:** 49-57.

Ambros, V., and H. R. Horvitz, 1984 Heterochronic mutants of the nematode Caenorhabditis elegans. Science 226**:** 409-416.

Ambros, V., and H. R. Horvitz, 1987 The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. Genes Dev 1**:** 398-414.

Ambros, V., and E. G. Moss, 1994 Heterochronic genes and the temporal control of C. elegans development. Trends Genet 10**:** 123-127.

Antebi, A., J. G. Culotti and E. M. Hedgecock, 1998 daf-12 regulates developmental age and the dauer alternative in Caenorhabditis elegans. Development 125**:** 1191-1205.

Antebi, A., W. H. Yeh, D. Tait, E. M. Hedgecock and D. L. Riddle, 2000 daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in C. elegans. Genes Dev 14**:** 1512-1527.

Arasu, P., B. Wightman and G. Ruvkun, 1991 Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. Genes Dev 5**:** 1825-1833.

Asahina, M., T. Ishihara, M. Jindra, Y. Kohara, I. Katsura *et al.*, 2000 The conserved nuclear receptor Ftz-F1 is required for embryogenesis, moulting and reproduction in Caenorhabditis elegans. Genes Cells 5**:** 711-723.

Ashmore, L. J., and A. Sehgal, 2003 A fly's eye view of circadian entrainment. J Biol Rhythms 18**:** 206-216.

Banerjee, D., A. Kwok, S. Y. Lin and F. J. Slack, 2005 Developmental timing in C. elegans is regulated by kin-20 and tim-1, homologs of core circadian clock genes. Dev Cell 8**:** 287-295.

Barrios, A., R. Ghosh, C. Fang, S. W. Emmons and M. M. Barr, 2012 PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in C. elegans. Nat Neurosci 15**:** 1675-1682.

Bethke, A., N. Fielenbach, Z. Wang, D. J. Mangelsdorf and A. Antebi, 2009 Nuclear hormone receptor regulation of microRNAs controls developmental progression. Science 324**:** 95-98.

Bigelow, H., M. Doitsidou, S. Sarin and O. Hobert, 2009 MAQGene: software to facilitate C. elegans mutant genome sequence analysis, pp. 549 in *Nat Methods*, United States.

Bodicoat, D. H., M. J. Schoemaker, M. E. Jones, E. McFadden, J. Griffin *et al.*, 2014 Timing of pubertal stages and breast cancer risk: the Breakthrough Generations Study. Breast Cancer Res 16**:** R18.

Boulin, T., and J. L. Bessereau, 2007 Mos1-mediated insertional mutagenesis in Caenorhabditis elegans. Nat Protoc 2**:** 1276-1287.

Brenner, S., 1974 The genetics of Caenorhabditis elegans. Genetics 77**:** 71-94.

Brito, V. N., B. B. Mendonca, L. M. Guilhoto, K. C. Freitas, I. J. Arnhold *et al.*, 2006 Allelic variants of the gamma-aminobutyric acid-A receptor alpha1-subunit gene (GABRA1) are not associated with idiopathic gonadotropin-dependent precocious puberty in girls with and without electroencephalographic abnormalities. J Clin Endocrinol Metab 91**:** 2432-2436.

Busza, A., M. Emery-Le, M. Rosbash and P. Emery, 2004 Roles of the two Drosophila CRYPTOCHROME structural domains in circadian photoreception. Science 304**:** 1503-1506.

Chalfie, M., H. R. Horvitz and J. E. Sulston, 1981 Mutations that lead to reiterations in the cell lineages of C. elegans. Cell 24**:** 59-69.

Chan, R. C., A. Chan, M. Jeon, T. F. Wu, D. Pasqualone *et al.*, 2003 Chromosome cohesion is regulated by a clock gene paralogue TIM-1. Nature 423**:** 1002-1009.

Chang, D. C., and S. M. Reppert, 2003 A novel C-terminal domain of drosophila PERIOD inhibits dCLOCK:CYCLE-mediated transcription. Curr Biol 13**:** 758-762.

Chang, H. M., N. J. Martinez, J. E. Thornton, J. P. Hagan, K. D. Nguyen *et al.*, 2012 Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nat Commun 3**:** 923.

Chen, Y. L., R. H. Yuan, W. C. Yang, H. C. Hsu and Y. M. Jeng, 2013 The stem cell E3-ligase Lin-41 promotes liver cancer progression through inhibition of microRNA-mediated gene silencing. J Pathol 229**:** 486-496.

Chitwood, D. J., 1999 Biochemistry and function of nematode steroids. Crit Rev Biochem Mol Biol 34**:** 273-284.

Chu, Y. D., W. C. Wang, S. A. Chen, Y. T. Hsu, M. W. Yeh *et al.*, 2014 RACK-1 regulates let-7 microRNA expression and terminal cell differentiation in Caenorhabditis elegans. Cell Cycle 13**:** 1995-2009.

Darlington, T. K., K. Wager-Smith, M. F. Ceriani, D. Staknis, N. Gekakis *et al.*, 1998 Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. Science 280**:** 1599-1603.

Davis, M. W., M. Hammarlund, T. Harrach, P. Hullett, S. Olsen *et al.*, 2005 Rapid single nucleotide polymorphism mapping in C. elegans. BMC Genomics 6**:** 118.

Denli, A. M., B. B. Tops, R. H. Plasterk, R. F. Ketting and G. J. Hannon, 2004 Processing of primary microRNAs by the Microprocessor complex. Nature 432**:** 231-235.

Dickinson, D. J., J. D. Ward, D. J. Reiner and B. Goldstein, 2013 Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nat Methods 10**:** 1028-1034.

Ding, L., A. Spencer, K. Morita and M. Han, 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. Mol Cell 19**:** 437-447.

Dissel, S., V. Codd, R. Fedic, K. J. Garner, R. Costa *et al.*, 2004 A constitutively active cryptochrome in Drosophila melanogaster. Nat Neurosci 7**:** 834-840.

Elks, C. E., K. K. Ong, R. A. Scott, Y. T. van der Schouw, J. S. Brand *et al.*, 2013 Age at menarche and type 2 diabetes risk: the EPIC-InterAct study. Diabetes Care 36**:** 3526-3534.

Emery, P., W. V. So, M. Kaneko, J. C. Hall and M. Rosbash, 1998 CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. Cell 95**:** 669-679.

Faunes, F., and J. Larrain, 2016 Conservation in the involvement of heterochronic genes and hormones during developmental transitions. Dev Biol 416**:** 3-17.

Fay, D. S., H. M. Stanley, M. Han and W. B. Wood, 1999 A Caenorhabditis elegans homologue of hunchback is required for late stages of development but not early embryonic patterning. Dev Biol 205**:** 240-253.

Frand, A. R., S. Russel and G. Ruvkun, 2005 Functional genomic analysis of C. elegans molting. PLoS Biol 3**:** e312.

Freitas, K. C., G. Ryan, V. N. Brito, Y. X. Tao, E. M. Costa *et al.*, 2007 Molecular analysis of the neuropeptide Y1 receptor gene in human idiopathic gonadotropin-dependent precocious puberty and isolated hypogonadotropic hypogonadism. Fertil Steril 87**:** 627-634.

Friedland, A. E., Y. B. Tzur, K. M. Esvelt, M. P. Colaiacovo, G. M. Church *et al.*, 2013 Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat Methods 10**:** 741-743.

Friedman, R. C., K. K. Farh, C. B. Burge and D. P. Bartel, 2009 Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19**:** 92-105.

Frokjaer-Jensen, C., M. W. Davis, M. Sarov, J. Taylor, S. Flibotte *et al.*, 2014 Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. Nat Methods 11**:** 529-534.

Frøkjær-Jensen, C., M. W. Davis, G. Hollopeter, J. Taylor, T. W. Harris *et al.*, 2010 Targeted gene deletions in C. elegans using transposon excision. Nat Methods 7**:** 451-453.

Gilbert, L. I., R. Rybczynski and J. T. Warren, 2002 Control and biochemical nature of the ecdysteroidogenic pathway. Annu Rev Entomol 47**:** 883-916.

Gissendanner, C. R., K. Crossgrove, K. A. Kraus, C. V. Maina and A. E. Sluder, 2004 Expression and function of conserved nuclear receptor genes in Caenorhabditis elegans. Dev Biol 266**:** 399-416.

Gissendanner, C. R., and A. E. Sluder, 2000 nhr-25, the Caenorhabditis elegans ortholog of ftz-f1, is required for epidermal and somatic gonad development. Dev Biol 221**:** 259-272.

Gregory, R. I., K. P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj *et al.*, 2004 The Microprocessor complex mediates the genesis of microRNAs. Nature 432**:** 235-240.

Hada, K., M. Asahina, H. Hasegawa, Y. Kanaho, F. J. Slack *et al.*, 2010 The nuclear receptor gene nhr-25 plays multiple roles in the Caenorhabditis elegans heterochronic gene network to control the larva-to-adult transition. Dev Biol 344**:** 1100-1109.

Hammell, C. M., X. Karp and V. Ambros, 2009a A feedback circuit involving let-7-family miRNAs and DAF-12 integrates environmental signals and developmental timing in Caenorhabditis elegans. Proc Natl Acad Sci U S A 106**:** 18668-18673.

Hammell, C. M., I. Lubin, P. R. Boag, T. K. Blackwell and V. Ambros, 2009b nhl-2 Modulates microRNA activity in Caenorhabditis elegans. Cell 136**:** 926-938.

Hao, H., D. L. Allen and P. E. Hardin, 1997 A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster. Mol Cell Biol 17**:** 3687-3693.
Hardin, P. E., 2005 The circadian timekeeping system of Drosophila. Curr Biol 15**:** R714-722.
Hardin, P. E., J. C. Hall and M. Rosbash, 1990 Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature 343**:** 536-540.
Hardin, P. E., J. C. Hall and M. Rosbash, 1992 Circadian oscillations in period gene mRNA levels are transcriptionally regulated. Proc Natl Acad Sci U S A 89**:** 11711-11715.
Harrington, J., and M. R. Palmert, 2012 Clinical review: Distinguishing constitutional delay of growth and puberty from isolated hypogonadotropic hypogonadism: critical appraisal of available diagnostic tests. J Clin Endocrinol Metab 97**:** 3056-3067.
Hata, A., and R. Kashima, 2016 Dysregulation of microRNA biogenesis machinery in cancer. Crit Rev Biochem Mol Biol 51**:** 121-134.
Hayes, G. D., A. R. Frand and G. Ruvkun, 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**:** 4631-4641.
Hendriks, G. J., D. Gaidatzis, F. Aeschimann and H. Grosshans, 2014 Extensive oscillatory gene expression during C. elegans larval development. Mol Cell 53**:** 380-392.
Hennig, S., H. M. Strauss, K. Vanselow, O. Yildiz, S. Schulze *et al.*, 2009 Structural and functional analyses of PAS domain interactions of the clock proteins Drosophila PERIOD and mouse PERIOD2. PLoS Biol 7**:** e94.
Heo, I., C. Joo, J. Cho, M. Ha, J. Han *et al.*, 2008 Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell 32**:** 276-284.
Herrero, A., A. Romanowski, E. Meelkop, C. S. Caldart, L. Schoofs *et al.*, 2015 Pigment-dispersing factor signaling in the circadian system of Caenorhabditis elegans. Genes Brain Behav 14**:** 493-501.
Hu, Y. B., C. B. Li, N. Song, Y. Zou, S. D. Chen *et al.*, 2016 Diagnostic Value of microRNA for Alzheimer's Disease: A Systematic Review and Meta-Analysis. Front Aging Neurosci 8**:** 13.
Huang, Z. J., I. Edery and M. Rosbash, 1993 PAS is a dimerization domain common to Drosophila period and several transcription factors. Nature 364**:** 259-262.
Jeon, M., H. F. Gardner, E. A. Miller, J. Deshler and A. E. Rougvie, 1999 Similarity of the C. elegans developmental timing protein LIN-42 to circadian rhythm proteins. Science 286**:** 1141-1146.
Katchy, A., and C. Williams, 2016 Expression Profiles of Estrogen-Regulated MicroRNAs in Breast Cancer Cells. Methods Mol Biol 1366**:** 373-393.
Kawamata, T., H. Seitz and Y. Tomari, 2009 Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol 16**:** 953-960.
Kimble, J., and D. Hirsh, 1979 The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. Dev Biol 70**:** 396-417.
Kippert, F., D. S. Saunders and M. L. Blaxter, 2002 Caenorhabditis elegans has a circadian clock, pp. R47-49 in *Curr Biol*, England.
Kloss, B., J. L. Price, L. Saez, J. Blau, A. Rothenfluh *et al.*, 1998 The Drosophila clock gene double-time encodes a protein closely related to human casein kinase Iepsilon. Cell 94**:** 97-107.

Kloss, B., A. Rothenfluh, M. W. Young and L. Saez, 2001 Phosphorylation of period is influenced by cycling physical associations of double-time, period, and timeless in the Drosophila clock. Neuron 30**:** 699-706.

Konopka, R. J., and S. Benzer, 1971 Clock mutants of Drosophila melanogaster. Proc Natl Acad Sci U S A 68**:** 2112-2116.

Koppen, M., J. S. Simske, P. A. Sims, B. L. Firestein, D. H. Hall *et al.*, 2001 Cooperative regulation of AJM-1 controls junctional integrity in Caenorhabditis elegans epithelia. Nat Cell Biol 3**:** 983-991.

Kostrouchova, M., M. Krause, Z. Kostrouch and J. E. Rall, 1998 CHR3: a Caenorhabditis elegans orphan nuclear hormone receptor required for proper epidermal development and molting. Development 125**:** 1617-1626.

Kostrouchova, M., M. Krause, Z. Kostrouch and J. E. Rall, 2001 Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode Caenorhabditis elegans. Proc Natl Acad Sci U S A 98**:** 7360-7365.

Kouns, N. A., J. Nakielna, F. Behensky, M. W. Krause, Z. Kostrouch *et al.*, 2011 NHR-23 dependent collagen and hedgehog-related genes required for molting. Biochem Biophys Res Commun 413**:** 515-520.

Kucera, N., I. Schmalen, S. Hennig, R. Ollinger, H. M. Strauss *et al.*, 2012 Unwinding the differences of the mammalian PERIOD clock proteins from crystal structure to cellular function. Proc Natl Acad Sci U S A 109**:** 3311-3316.

Lagos-Quintana, M., R. Rauhut, J. Meyer, A. Borkhardt and T. Tuschl, 2003 New microRNAs from mouse and human. Rna 9**:** 175-179.

Lau, N. C., L. P. Lim, E. G. Weinstein and D. P. Bartel, 2001 An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294**:** 858-862.

Lee, C., K. Bae and I. Edery, 1999 PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. Mol Cell Biol 19**:** 5316-5325.

Lee, C., D. R. Weaver and S. M. Reppert, 2004a Direct association between mouse PERIOD and CKIepsilon is critical for a functioning circadian clock. Mol Cell Biol 24**:** 584-594.

Lee, R. C., R. L. Feinbaum and V. Ambros, 1993 The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75**:** 843-854.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee *et al.*, 2004b MicroRNA genes are transcribed by RNA polymerase II. Embo j 23**:** 4051-4060.

Lehrbach, N. J., J. Armisen, H. L. Lightfoot, K. J. Murfitt, A. Bugaut *et al.*, 2009 LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat Struct Mol Biol 16**:** 1016-1020.

Li, J., G. Brown, M. Ailion, S. Lee and J. H. Thomas, 2004 NCR-1 and NCR-2, the C. elegans homologs of the human Niemann-Pick type C1 disease protein, function upstream of DAF-9 in the dauer formation pathways. Development 131**:** 5741-5752.

Li, M., M. W. Jones-Rhoades, N. C. Lau, D. P. Bartel and A. E. Rougvie, 2005 Regulatory mutations of mir-48, a C. elegans let-7 family MicroRNA, cause developmental timing defects. Dev Cell 9**:** 415-422.

Lim, L. P., N. C. Lau, E. G. Weinstein, A. Abdelhakim, S. Yekta *et al.*, 2003 The microRNAs of Caenorhabditis elegans. Genes Dev 17**:** 991-1008.

Lin, F. J., W. Song, E. Meyer-Bernstein, N. Naidoo and A. Sehgal, 2001 Photic signaling by cryptochrome in the Drosophila circadian system. Mol Cell Biol 21**:** 7287-7294.
Lin, S., and R. I. Gregory, 2015 MicroRNA biogenesis pathways in cancer. Nat Rev Cancer 15**:** 321-333.
Lin, S. Y., S. M. Johnson, M. Abraham, M. C. Vella, A. Pasquinelli *et al.*, 2003 The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev Cell 4**:** 639-650.
Liu, X., F. Long, H. Peng, S. J. Aerni, M. Jiang *et al.*, 2009 Analysis of cell fate from single-cell gene expression profiles in C. elegans. Cell 139**:** 623-633.
Liu, Z., 1990 Genetic control of stage-specific developmental events in *C. elegans*., pp. in *The Department of Cellular and Developmental Biology*. Harvard University.
Livak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25**:** 402-408.
Loedige, I., D. Gaidatzis, R. Sack, G. Meister and W. Filipowicz, 2013 The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Res 41**:** 518-532.
Ludewig, A. H., C. Kober-Eisermann, C. Weitzel, A. Bethke, K. Neubert *et al.*, 2004 A novel nuclear receptor/coregulator complex controls C. elegans lipid metabolism, larval development, and aging. Genes Dev 18**:** 2120-2133.
Martinek, S., S. Inonog, A. S. Manoukian and M. W. Young, 2001 A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock. Cell 105**:** 769-779.
McCulloch, K. A., and A. E. Rougvie, 2014 Caenorhabditis elegans period homolog lin-42 regulates the timing of heterochronic miRNA expression. Proc Natl Acad Sci U S A 111**:** 15450-15455.
McDonald, M. J., M. Rosbash and P. Emery, 2001 Wild-type circadian rhythmicity is dependent on closely spaced E boxes in the Drosophila timeless promoter. Mol Cell Biol 21**:** 1207-1217.
Meelkop, E., L. Temmerman, T. Janssen, N. Suetens, I. Beets *et al.*, 2012 PDF receptor signaling in Caenorhabditis elegans modulates locomotion and egg-laying. Mol Cell Endocrinol 361**:** 232-240.
Meli, V. S., B. Osuna, G. Ruvkun and A. R. Frand, 2010 MLT-10 defines a family of DUF644 and proline-rich repeat proteins involved in the molting cycle of Caenorhabditis elegans. Mol Biol Cell 21**:** 1648-1661.
Mohler, W. A., J. S. Simske, E. M. Williams-Masson, J. D. Hardin and J. G. White, 1998 Dynamics and ultrastructure of developmental cell fusions in the Caenorhabditis elegans hypodermis. Curr Biol 8**:** 1087-1090.
Monsalve, G. C., C. Van Buskirk and A. R. Frand, 2011 LIN-42/PERIOD controls cyclical and developmental progression of C. elegans molts. Curr Biol 21**:** 2033-2045.
Morita, K., and M. Han, 2006 Multiple mechanisms are involved in regulating the expression of the developmental timing regulator lin-28 in Caenorhabditis elegans. Embo j 25**:** 5794-5804.
Moss, E. G., R. C. Lee and V. Ambros, 1997 The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell 88**:** 637-646.
Nawathean, P., and M. Rosbash, 2004 The doubletime and CKII kinases collaborate to potentiate Drosophila PER transcriptional repressor activity. Mol Cell 13**:** 213-223.

Newman, M. A., J. M. Thomson and S. M. Hammond, 2008 Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. Rna 14**:** 1539-1549.

O'Reilly, S., 2016 MicroRNAs in fibrosis: opportunities and challenges. Arthritis Res Ther 18**:** 11.

Olmedo, M., J. S. O'Neill, R. S. Edgar, U. K. Valekunja, A. B. Reddy *et al.*, 2012 Circadian regulation of olfaction and an evolutionarily conserved, nontranscriptional marker in Caenorhabditis elegans. Proc Natl Acad Sci U S A 109**:** 20479-20484.

Paix, A., Y. Wang, H. E. Smith, C. Y. Lee, D. Calidas *et al.*, 2014 Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in Caenorhabditis elegans. Genetics 198**:** 1347-1356.

Park, J. E., I. Heo, Y. Tian, D. K. Simanshu, H. Chang *et al.*, 2011 Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature 475**:** 201-205.

Pasquinelli, A. E., B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda *et al.*, 2000 Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408**:** 86-89.

Perales, R., D. M. King, C. Aguirre-Chen and C. M. Hammell, 2014 LIN-42, the Caenorhabditis elegans PERIOD homolog, negatively regulates microRNA transcription. PLoS Genet 10**:** e1004486.

Prentice, P., and R. M. Viner, 2013 Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. Int J Obes (Lond) 37**:** 1036-1043.

Price, J. L., J. Blau, A. Rothenfluh, M. Abodeely, B. Kloss *et al.*, 1998 double-time is a novel Drosophila clock gene that regulates PERIOD protein accumulation. Cell 94**:** 83-95.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger *et al.*, 2000 The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403**:** 901-906.

Renn, S. C., J. H. Park, M. Rosbash, J. C. Hall and P. H. Taghert, 1999 A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in Drosophila. Cell 99**:** 791-802.

Resnick, T. D., K. A. McCulloch and A. E. Rougvie, 2010 miRNAs give worms the time of their lives: small RNAs and temporal control in Caenorhabditis elegans. Dev Dyn 239**:** 1477-1489.

Rosenfield, R. L., R. B. Lipton and M. L. Drum, 2009 Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. Pediatrics 123**:** 84-88.

Rougvie, A. E., and V. Ambros, 1995 The heterochronic gene lin-29 encodes a zinc finger protein that controls a terminal differentiation event in Caenorhabditis elegans. Development 121**:** 2491-2500.

Rougvie, A. E., and E. G. Moss, 2013 Developmental transitions in C. elegans larval stages. Curr Top Dev Biol 105**:** 153-180.

Rybak, A., H. Fuchs, K. Hadian, L. Smirnova, E. A. Wulczyn *et al.*, 2009 The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nat Cell Biol 11**:** 1411-1420.

Rybak, A., H. Fuchs, L. Smirnova, C. Brandt, E. E. Pohl *et al.*, 2008 A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol 10**:** 987-993.

Saez, L., P. Meyer and M. W. Young, 2007 A PER/TIM/DBT interval timer for Drosophila's circadian clock. Cold Spring Harb Symp Quant Biol 72**:** 69-74.

Saigusa, T., S. Ishizaki, S. Watabiki, N. Ishii, A. Tanakadate *et al.*, 2002 Circadian behavioural rhythm in Caenorhabditis elegans, pp. R46-47 in *Curr Biol*, England.

Sehgal, A., A. Rothenfluh-Hilfiker, M. Hunter-Ensor, Y. Chen, M. P. Myers *et al.*, 1995 Rhythmic expression of timeless: a basis for promoting circadian cycles in period gene autoregulation. Science 270**:** 808-810.

Silveira-Neto, A. P., L. F. Leal, A. B. Emerman, K. D. Henderson, E. Piskounova *et al.*, 2012 Absence of functional LIN28B mutations in a large cohort of patients with idiopathic central precocious puberty. Horm Res Paediatr 78**:** 144-150.

Simonetta, S. H., and D. A. Golombek, 2007 An automated tracking system for Caenorhabditis elegans locomotor behavior and circadian studies application. J Neurosci Methods 161**:** 273-280.

Simonetta, S. H., M. L. Migliori, A. Romanowski and D. A. Golombek, 2009 Timing of locomotor activity circadian rhythms in Caenorhabditis elegans. PLoS One 4**:** e7571.

Singh, K., M. Y. Chao, G. A. Somers, H. Komatsu, M. E. Corkins *et al.*, 2011 C. elegans Notch signaling regulates adult chemosensory response and larval molting quiescence. Curr Biol 21**:** 825-834.

Sizonenko, P. C., 1978 Endocrinology in preadolescents and adolescents. I. Hormonal changes during normal puberty. Am J Dis Child 132**:** 704-712.

Skorupskaite, K., J. T. George and R. A. Anderson, 2014 The kisspeptin-GnRH pathway in human reproductive health and disease. Hum Reprod Update 20**:** 485-500.

Slack, F., and G. Ruvkun, 1997 Temporal pattern formation by heterochronic genes. Annu Rev Genet 31**:** 611-634.

Slack, F. J., M. Basson, Z. Liu, V. Ambros, H. R. Horvitz *et al.*, 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. Mol Cell 5**:** 659-669.

Sorensen, K., A. Mouritsen, L. Aksglaede, C. P. Hagen, S. S. Mogensen *et al.*, 2012 Recent secular trends in pubertal timing: implications for evaluation and diagnosis of precocious puberty. Horm Res Paediatr 77**:** 137-145.

Sulston, J. E., 1983 Neuronal cell lineages in the nematode Caenorhabditis elegans. Cold Spring Harb Symp Quant Biol 48 Pt 2**:** 443-452.

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

Sun, W. C., E. H. Jeong, H. J. Jeong, H. W. Ko, I. Edery *et al.*, 2010 Two distinct modes of PERIOD recruitment onto dCLOCK reveal a novel role for TIMELESS in circadian transcription. J Neurosci 30**:** 14458-14469.

Temmerman, L., E. Meelkop, T. Janssen, A. Bogaerts, M. Lindemans *et al.*, 2011 C. elegans homologs of insect clock proteins: a tale of many stories. Ann N Y Acad Sci 1220**:** 137-148.

Tennessen, J. M., H. F. Gardner, M. L. Volk and A. E. Rougvie, 2006 Novel heterochronic functions of the Caenorhabditis elegans period-related protein LIN-42. Dev Biol 289**:** 30-43.

Tennessen, J. M., K. J. Opperman and A. E. Rougvie, 2010 The C. elegans developmental timing protein LIN-42 regulates diapause in response to environmental cues. Development 137**:** 3501-3511.

Terns, R. M., P. Kroll-Conner, J. Zhu, S. Chung and J. H. Rothman, 1997 A deficiency screen for zygotic loci required for establishment and patterning of the epidermis in Caenorhabditis elegans. Genetics 146**:** 185-206.

Thummel, C. S., 1996 Flies on steroids--Drosophila metamorphosis and the mechanisms of steroid hormone action. Trends Genet 12**:** 306-310.

Tommiska, J., K. Sorensen, L. Aksglaede, R. Koivu, L. Puhakka *et al.*, 2011 LIN28B, LIN28A, KISS1, and KISS1R in idiopathic central precocious puberty. BMC Res Notes 4**:** 363.

Tsialikas, J., and J. Romer-Seibert, 2015 LIN28: roles and regulation in development and beyond. Development 142**:** 2397-2404.

Vadla, B., K. Kemper, J. Alaimo, C. Heine and E. G. Moss, 2012 lin-28 controls the succession of cell fate choices via two distinct activities. PLoS Genet 8**:** e1002588.

Vallin, E., J. Gallagher, L. Granger, E. Martin, J. Belougne *et al.*, 2012 A genome-wide collection of Mos1 transposon insertion mutants for the C. elegans research community. PLoS One 7**:** e30482.

Van Buskirk, C., and P. W. Sternberg, 2007 Epidermal growth factor signaling induces behavioral quiescence in Caenorhabditis elegans. Nat Neurosci 10**:** 1300-1307.

van der Linden, A. M., M. Beverly, S. Kadener, J. Rodriguez, S. Wasserman *et al.*, 2010 Genome-wide analysis of light- and temperature-entrained circadian transcripts in Caenorhabditis elegans. PLoS Biol 8**:** e1000503.

Van Wynsberghe, P. M., E. F. Finnegan, T. Stark, E. P. Angelus, K. E. Homan *et al.*, 2014 The Period protein homolog LIN-42 negatively regulates microRNA biogenesis in C. elegans. Dev Biol 390**:** 126-135.

Van Wynsberghe, P. M., Z. S. Kai, K. B. Massirer, V. H. Burton, G. W. Yeo *et al.*, 2011 LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat Struct Mol Biol 18**:** 302-308.

Vella, M. C., E. Y. Choi, S. Y. Lin, K. Reinert and F. J. Slack, 2004 The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes Dev 18**:** 132-137.

Viswanathan, S. R., G. Q. Daley and R. I. Gregory, 2008 Selective blockade of microRNA processing by Lin28. Science 320**:** 97-100.

Wang, G. K., A. Ousley, T. K. Darlington, D. Chen, Y. Chen *et al.*, 2001 Regulation of the cycling of timeless (tim) RNA. J Neurobiol 47**:** 161-175.

Wang, N., Z. Zhou, X. Liao and T. Zhang, 2009 Role of microRNAs in cardiac hypertrophy and heart failure. IUBMB Life 61**:** 566-571.

Wicks, S. R., R. T. Yeh, W. R. Gish, R. H. Waterston and R. H. Plasterk, 2001 Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat Genet 28**:** 160-164.

Wightman, B., T. R. Burglin, J. Gatto, P. Arasu and G. Ruvkun, 1991 Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. Genes Dev 5**:** 1813-1824.

Wightman, B., I. Ha and G. Ruvkun, 1993 Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75**:** 855-862.

Winnier, A. R., J. Y. Meir, J. M. Ross, N. Tavernarakis, M. Driscoll *et al.*, 1999 UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in Caenorhabditis elegans. Genes Dev 13**:** 2774-2786.

Winter, J., S. Jung, S. Keller, R. I. Gregory and S. Diederichs, 2009 Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol 11**:** 228-234.

Yagita, K., S. Yamaguchi, F. Tamanini, G. T. van Der Horst, J. H. Hoeijmakers *et al.*, 2000 Dimerization and nuclear entry of mPER proteins in mammalian cells. Genes Dev 14**:** 1353-1363.

Yildiz, O., M. Doi, I. Yujnovsky, L. Cardone, A. Berndt *et al.*, 2005 Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. Mol Cell 17**:** 69-82.

Yu, W., and P. E. Hardin, 2006 Circadian oscillators of Drosophila and mammals. J Cell Sci 119**:** 4793-4795.