2019-09-23

The Plasticity and Variation in Gene Expression during Development in C. elegans

Io Long Chan
University of Massachusetts Medical School

THE PLASTICITY AND VARIATION IN GENE EXPRESSION DURING DEVELOPMENT IN *C. elegans*

A Dissertation Presented

By

IO LONG CHAN

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences,

Worcester in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 23rd, 2019

Interdisciplinary Gradate Program

THE PLASTICITY AND VARIATION IN GENE EXPRESSION DURING DEVELOPMENT IN *C. elegans*

A Dissertation Presented

By

IO LONG CHAN

This work was undertaken in the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

Under the mentorship of

Oliver Rando, M.D. Ph.D., Thesis Advisor

Cole Haynes, Ph.D., Member of Committee

Sean Ryder, Ph.D., Member of Committee

Craig Peterson, Ph.D., Member of Committee

Eric Greer, Ph.D., External Member of Committee

Marian Walhout, Ph.D., Chair of Committee

Mary Ellen Lane, PhD., Dean of the Graduate School of Biomedical Sciences

September 23rd, 2019

With love

To Mom and Dad, and Winnie

And

To my dearest

Suet Yan

ACKNOWLEDGEMENTS

This work would not have been possible without the help, support and discussion from mentors, colleagues and collaborators within and outside of UMass Medical School. First, I would like to acknowledge and thank my advisor, Oliver Rando, for his patience, support and mentorship. His fearless spirit inspired me to aim higher and reach further, and his creative and rigorous approach to science has dared me to question even the most fundamental assumptions. Furthermore, his personable nature and sage advice helped mold me into a better scientist and thinker. I am lucky and grateful to have the opportunity to do science and learn with him.

I would like to thank past and present Rando lab members for their scientific discussions and support. To all the magical dudes - Ana, Caitlin, Carolina, David, Ebru, Fengyun, Lina, Marina, Morten, Nils, Shweta, Stanley, Tobias, Tiffany, Upasna, Vera: thank you all for making the lab a fun place to work and my home away from home. Special thanks to Colin, who helped me find my way into the fascinating world of worms, my projects would not have been possible without his mentorship and work. I thank the UMMS community for their tremendous help with sharing reagents, protocol and equipment sharing, and experimental ideas, especially Craig Mello lab, Marian Walhout lab, Sean Ryder lab, and Victor Ambros lab. I am indebted to all my past and present committee members Marian, Sean, Cole Haynes, Craig Mello, Amy Walker and Craig Peterson for their excitement and support towards me.

Special thanks to Eric Greer for his contributions as an external committee member.

I would like to thank collaborators whose work and discussions were crucial for the completion of this work. Pranitha Vangala for help in analyzing data for the evolution of gene expression project. Ross Lagoy and Christina Baer for help with imaging and analysis, Alex Artyukhin for in-depth discussions about L1 density. Thanks to all my colleagues, classmates and friends for always being open to discussions, generous with their time and encouragement: Ankit Bhatta, Nate Gioacchini, Jason Fan, Sunil Guharajan, Nick Rice, Sam Ho, Tong Wu, Shang Ye, Dante Lepore, Daniel Gao, Sungwook Choi, Ildar Gainetdinov, Menna Albarqi, Takao Ishidate, Dan Durning, Katya Makeyeva, Greg Dokshin, Aurian Garcia-Gonzalez, Olga Ponomarova, Monika Chitre. Also, I would like to thank the UMMS GSBS staff for their support and advice in navigating grad school.

I thank Mrs. Karen Harrison who supported me when I was in Houston. Mrs. Harrison took me into her home when I had no place to live, I certainly would not have made it this far without her help.

I would like to thank Dad, Mom, and Winnie for their unwavering love and support since day 1. My in-laws, for their support and stimulating conversations about science and life. Finally, I thank my wonderful and loving wife, Suet Yan, for being my number 1 supporter through thick and thin. This work would have been impossible without her care, love and support.

ABSTRACT

Organisms modulate their response to changing environmental conditions through changes in gene expression, and extensive variations in gene expression are prevalent among individuals even within a population. This widespread plasticity and variability of gene expression is thought to play roles in adaptation and drive novel phenotypes in species. Understanding the mechanisms that contribute to such variations requires the analysis of interactions between the genome and its environment and sequence variations within the genome. This work consists of two projects investigating the plasticity and variation of gene expression during post-embryonic development in the nematode *C. elegans*.

In the first study, I examined the response to changes in population density in developmentally arrested L1 larvae. I systematically characterized arrested L1 larvae from low to high densities using single-worm RNA-seq and uncovered that the density of resuspended L1 larvae regulates the expression of hundreds of mRNAs. Further analysis revealed that the physiological response to changes in density is rapid and signaled by a non-canonical *daf-22* ascaroside independent pathway. In the second study, I investigated the evolution of gene expression within species using two genetically divergent *C. elegans* strains (N2 and CB4856). I carried out RNA-seq and allele-specific analysis across six different conditions and four developmental stages, and we examined gene expression divergence using the homozygous parent and F1 hybrid system. This work provides a new experimental model for studying the

evolution of gene expression and a comprehensive view of gene expression variation during development in *C. elegans*.

Chapter 2: Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF FILES

Raw data are available at GEO with accession number GSE112053.

File 2.1 Summary of sequencing from single L1 RNA-seq samples. Mapping statistics for all individual worm RNA-Seq datasets in this study. (.xlsx)

https://www.g3journal.org/highwire/filestream/489363/field_highwire_adjunct_files/3/TableS1.xlsx

File 2.2 *fog-2* dataset. Single-animal RNA-Seq data for fog-2 animals arrested at 1, 5, 20, or 100 eggs/mL, as indicated. (.xlsx)

https://www.g3journal.org/highwire/filestream/489363/field_highwire_adjunct_files/4/TableS2.xlsx

File 2.3 N2 dataset. Single-animal RNA-Seq data for N2 animals arrested at 1, 5, 20, or 100 eggs/mL, as indicated. (.xlsx)

https://www.g3journal.org/highwire/filestream/489363/field_highwire_adjunct_files/5/TableS3.xlsx

PREFACE

The work in Chapter 2 was conducted in the Rando lab. Experiments were designed by Io Long Chan (ILC), Oliver J. Rando (OJR) and Colin C. Conine (CCC). Experiments were performed by ILC. Data was analyzed and figures were generated by ILC and OJR. The manuscript was originally written by ILC and reviewed and edited by OJR and CCC. This work has been published. Some sections of the introduction and the body of Chapter 2 is reprinted and adapted from:

> **Effects of Larval Density on Gene Regulation in Caenorhabditis elegans During Routine L1 Synchronization**
> Io Long Chan, Oliver J. Rando, and Colin C. Conine.
> G3: GENES, GENOMES, GENETICS *May 1, 2018 vol. 8 no. 5 1787-1793;https://doi.org/10.1534/g3.118.200056*

The experimental work in Chapter 3 was conducted in the Rando lab. Experiments were designed and conducted by ILC, with input from OJR. Raw sequencing data processing was performed by Pranitha Vangala (PV), and figures were generated by ILC, OJR and PV. This work is unpublished.

Introduction, the entirety of Chapter 3, and Discussion sections were written by ILC.

Figures 1-1, 3-1, and 4-1 were created with BioRender.com.

Chapter 1: General Introduction

Variation in Phenotypes: Molecular Origins

By breeding pigeons with long beaks to pigeons with longer beaks, or short beaks to shorter beaks, a pigeon fancier can choose to modify certain characters or traits through deliberate selection. An analogous strategy can be applied to how a local environment can selectively "choose" to augment certain characters, for example, a faster hunter is more likely to find prey, survive and reproduce, therefore the "fast" trait is more likely to be passed on to progeny. Regardless of whether selective pressures acted under domestication or in nature, however, traits within a species are generally highly variable and new varieties always appear after successive generations. The observation that species always show "divergence of character" helped form the basis of Darwin's theory of natural selection (Darwin, 1859). But what underlies this variation?

With the revolution in molecular biology and genomics, many genomes have been sequenced and revealed a clear trend: there is great genetic diversity and variation between individuals within and between species. Strictly speaking, there are two sources that contribute to phenotypic diversity. The first source is genetic variation. If we take into consideration an entire genome, for example the human genome, there are over 3 billion base pairs or mutable sites that can potentially give rise to variations in phenotype.

However, not all mutations spawn new features, suggesting an inherent selection towards certain sequences and regions of the genome that can give rise to phenotypic variation.

What are the mutable targets in the genome? If we divide the genome into protein-coding and non-coding regions, mutations in different regions can produce phenotypic changes via different mechanisms. In protein-coding regions, missense mutations change the primary amino acid sequence and give rise to new protein structure and potentially new function. Although any changes in RNA sequence, whether it alters amino acid sequence or not, have the potential to alter RNA secondary structures, stability and translational efficiency affecting protein function through changing intracellular protein concentrations (Ancel and Fontana, 2000; D'Andrea et al., 2019).

In non-coding regions, there are regions that: provide chromosomal structure, such as telomeres, centromeres and scaffold/matrix attachment regions; contain sites for origin of DNA replication; contain transposons, repeats, mobile elements; generate functional non-coding RNA molecules, such as transfer RNAs, ribosomal RNAs, and a universe of regulatory RNAs; harbor gene regulatory regions within genes like introns, or intergenic regions that can affect transcription, such as *cis*- regulatory elements such as promoters, enhancers and silencers. Recent effort using a massive panel of 18,126 fully genotyped F_6 segregants between two *S. cerevisiae* isolates found that the effect-size distributions of genetic variation that affected phenotypic variation overlapped between regions within and outside of open

reading frames (ORFs). Further molecular characterization of variable regions outside of ORFs reveal genetic variation correlated with levels of histone modifications near actively transcribed genes and showed that this had a positive effect on gene expression. This study effectively linked a direct relationship between genetic variation, local chromatin architecture, gene expression and phenotypic variation providing a comprehensive mechanistic view of the genotype-phenotype interactions (Jakobson and Jarosz, 2019).

The second source that contributes to phenotypic variation is the environment. Changes to environmental conditions can occur rapidly and frequently, and organisms can change their phenotypes specific to the environmental stimulus. The basic premise is that the same genome or genotype can produce distinct phenotypes in response to different environmental cues. In general, there are two modes to this type of response: 1) effects that can last for more than one generation without the initial stimulus, suggesting inheritance of information beyond DNA sequence or transgenerational epigenetic inheritance (Bošković and Rando, 2018; Heard and Martienssen, 2014); 2) effects lasting only within the organism's life span and serve as a coping mechanism in response to environmental variation by tuning their phenotype, and this is generally known as phenotypic plasticity (Kelly et al., 2012; Pigliucci, 2001; Rando and Verstrepen, 2007). The response to changes in environment generally affect phenotypes by altering gene expression (Grishkevich and Yanai, 2013). There are numerous studies documenting many different forms of environmental stimuli that can induce changes in gene expression and phenotype, and I will further expand on both

transgenerational epigenetic inheritance and phenotypic plasticity relevant to our system.

In this work, I characterized two sources that contribute to phenotypic variation in *Caenorhabditis elegans*. In Chapter 2, I uncovered a cryptic source for phenotypic plasticity and characterized its effects on gene expression in early stage larvae. In Chapter 3, I implemented a novel system to investigate how natural genetic variation within *C. elegans* species can contribute to gene expression variation.

Feeling the World Through the Worm

Often found in compost and soil, *C. elegans* is a free-living nematode worm that was first domesticated for genetic experiments by Sydney Brenner in the 1960s (Brenner, 1974). Since then, the worm has been a key model for studying the genetic basis for many biological processes. One of the most remarkable aspects of the worm is their repertoire of plastic responses to changing environments.

One prominent example of phenotypic plasticity in *C. elegans* is the dauer/non-dauer decision during early development (Cassada and Russell, 1975; Hu, 2007; Klass and Hirsh, 1976). The decision is based on three environmental cues: population density, food supply and temperature. Based on these cues, worms can remodel their developmental trajectory and enter a highly stress resistant, developmentally arrested state called dauer that can survive without food for three months (Cassada and Russell, 1975; Golden and Riddle, 1982, 1984; Hu, 2007; Klass and Hirsh, 1976).

Two recent curious cases in *C. elegans* demonstrated highly complex plasticity phenotypes to shifting environments. First, the Kuhara lab characterized the necessary cell types and natural variations in a cold acclimation response where worms grown at 20 or 25ºC is unable to survive when placed directly at 2ºC, but is able to display almost 100% survival rate if first cultivated at 15ºC for 3 hours (Ohta et al., 2014; Okahata et al., 2016; Sonoda et al., 2016). This study suggests a mechanism in which gradual

changes in environment can reprogram an organism's physiology (perhaps transcriptomic changes) to help them cope with changing environments, in this case temperature. This provides a clear example of how the same genotype can produce distinct phenotypes under varying environments. Second case is related to pathogenic avoidance behavior. It's known that *C. elegans* can learn to avoid the pathogenic bacteria *Pseudomonas aeruginosa* after brief exposure (4 hours) (Zhang et al., 2005). Recently, work from the Murphy lab showed that long (24 h) but not brief (4 h) exposure to *P. aeruginosa* allowed not only the parents, but also their progeny to avoid *P. aeruginosa* despite progeny never being exposed to *P. aeruginosa* (Moore et al., 2019). This avoidance behavior lasted for four generations, and was specific to *P. aeruginosa* but not another pathogen *Serratia marcescens* even though the avoidance behavior was observed in parents exposed to *S. marcescens* (Moore et al., 2019), suggesting inherent differences within a similar plasticity response. This example also demonstrates that some phenotypic plasticity and adaptive responses activated in ancestors can extend beyond an organism's life span, providing a source of phenotypic variability within species.

What are the mechanisms for sensing environmental cues that can lead to such coordinate changes in physiology and morphology? Without visual- or audio-sensory capabilities, worm make sense of the world through chemosensory, mechanosensory, and thermosensory mechanisms. As ectotherms, they depend on surrounding temperatures to regulate their body heat, and it is known that temperature can alter developmental rate (Byerly et

al., 1976) and body size (Van Voorhies, 1996). With only 30 (hermaphrodites) or 52 (males) mechanoreceptor neurons, worms can display a variety of complex behavior through mechanical stimuli (Goodman, 2006). But perhaps the most amazing of these sensory mechanisms is chemosensation in *C. elegans*, with more than 5% of proteins dedicated to interacting with environmental signals (Bargmann, 2006).

One important use for chemosensation is to sense pheromones. A diverse family of small molecule pheromones termed ascarosides are a modular chemical language used for inter-nematode communication (Ludewig , A and Schroeder, F, 2013). Various classes of ascaroside molecules and pathways have been extensively studied in a multitude of biological functions, including dauer formation and exit (Butcher et al., 2009a; Golden and Riddle, 1982), aggregation (MacOsko et al., 2009; Srinivasan et al., 2012) ,olfactory plasticity (Yamada et al., 2010), mate attraction (Srinivasan et al., 2008), mate competition (Shi et al., 2017), and even signaling to other organisms (Hsueh et al., 2013; Manosalva et al., 2015; Zhao et al., 2016).

To generate ascarosides, a thiolase – DAF-22 – is necessary for catalyzing the final step in a β-oxidation cycle for producing ascarosides (Butcher et al., 2009b; Von Reuss et al., 2012), and *daf-22* mutant worms or extracts are commonly used as negative controls when investigating ascarosides and inter-nematode signaling (Artyukhin et al., 2018; Golden and Riddle, 1985; Maures et al., 2014; Shi et al., 2017; Srinivasan et al., 2008).

Curiously, few studies document or hinted at *daf-22* independent non-ascaroside based signaling in *C. elegans* (Aprison and Ruvinsky, 2016; Artyukhin et al., 2013; Ludewig et al., 2017, 2019; Maures et al., 2014). Of great interest to the work in Chapter 2, the experiments performed in Artyukhin et al. (2013) was the first to characterize the effects of population density during L1 (larval stage 1) starvation (Artyukhin et al., 2013). They found that starved L1s at high densities survived longer than starved L1s at lower densities. They also characterized a broad range of density-dependent survival rates in different *Caenorhabditis* species. Further, they found that this response to density was independent of *daf-22* (Artyukhin et al., 2013). These results demonstrated a highly plastic and genetically adjustable modes of inter-nematode communication through an elusive *daf-22* independent pathway within *C. elegans* species.

Paternal Effects on Phenotype

As mentioned above, one mode of environmentally induced phenotypic variation is via transgenerational epigenetic inheritance (TEI). Compared to mitotic epigenetic inheritance, i.e. the inheritance of cell states, TEI is unique in a sense that environmental information is transmitted through the germline of multicellular organisms from one generation to the next. TEI has been investigated extensively in both plants and animals in a large number of paradigms. The transmission of ancestral environmental conditions from parent to offspring is, of course, available regardless of parental sex. Here, I will focus on paternal transmission of environmental conditions in mammals that motivated the study in Chapter 2 and relevant evidence in *C. elegans*, rather than attempting to provide a comprehensive list of TEI studies. Following listed are excellent reviews on TEI and other paternal effects paradigms (Bošković and Rando, 2018; Ferguson-Smith, 2011; Heard and Martienssen, 2014; Hollick, 2016; Perez and Lehner, 2019; Rando, 2012; Tucci et al., 2019).

There is particular interest in studying paternal effects as it is easier, at least conceptually, to separate epigenetic effects and maternal effects by limiting the exposure of males to their partners and offspring (Rando, 2012). Generally, the experiments in paternal effects studies involve using inbred colonies or isogenic populations and exposing males to different environmental conditions, and changes observed in offspring are therefore an

effect from changes beyond DNA sequence. Experimental work has shown mainly three types of environmental stimuli fathers experience that can affect offspring phenotype in animals.

Firstly, nutrition availability during a father's lifetime can alter offspring metabolism. Using mice, work from our lab and others have shown males placed on different diets sire offspring exhibiting altered gene expression and metabolism (Carone et al., 2010; Chen et al., 2015b; Fullston et al., 2013; Ng et al., 2010; Rando and Simmons, 2015; Sharma et al., 2015; Wei et al., 2014). Dietary paradigms include low-protein diets (Carone et al., 2010; Sharma et al., 2015), high-fat diets (Chen et al., 2015b; Ng et al., 2010), and caloric restriction (Anderson et al., 2006; Jimenez-Chillaron et al., 2009; Radford et al., 2014). In *Drosophila melanogaster*, acute increase in sugar intake in males induced an obesity phenotype in F1 offspring (Öst et al., 2014). These studies all showed a common trend of alternations in glucose and lipid metabolism in offspring. Secondly, paternal stress paradigms such as exposure to maternal separation, social defeat, and chronic stress, resulted in progeny that showed altered corticosterone levels, glucose metabolism and blood-brain barrier permeability in mice (Bale, 2015; Gapp et al., 2014; Rodgers et al., 2015). Thirdly, paternal exposure to toxic compounds and drugs, such as endocrine disruptors which led to F1 male infertility (Anway et al., 2005), nicotine exposure led to increased F1 male specific survival to lethal doses of nicotine in mice (Vallaster et al., 2017), or cocaine exposure led to a F1 male specific decrease in reinforcing effectiveness of cocaine (Vassoler et al., 2013), and carbon tetrachloride

exposure resulted in an increase in hepatic wound-healing in rats (Zeybel et al., 2012). These results indicate that phenotypic responses in F1 offspring are related to the conditions similar to the paternal environment. This suggest that there is specificity in the pathways that form the “memory” of paternal environment that is then transmitted to offspring, and that the memory of paternal environmental conditions is not just a general response to stressful conditions.

In *C. elegans,* there are many cases of parental transmission of environmental conditions, these include exposure to osmotic stress, olfactory stimuli, dietary restriction, high glucose exposure, starvation, maternal age, heat shock, pathogen exposure, and heavy metal exposure (Frazier and Roth, 2009; Jobson et al., 2015; Kishimoto et al., 2017; Klosin et al., 2017; Miersch and Döring, 2012; Moore et al., 2019; Perez and Lehner, 2019; Perez et al., 2017; Remy, 2010; Webster et al., 2018). However, only a handful of studies report the transmission of paternal environments. Miersch et al. reported that paternal dietary restriction (DR) affected fat content in progeny (Miersch and Döring, 2012). In this study, they exposed male worms to various DR regimes ranking from moderate to strong restriction. They found an inverse U-shaped correlation between the dose of paternal DR and offspring fat content, i.e. fathers exposed to the most moderate and most extreme DR regimes sired progeny that have the most reduced fat content, while progeny sired by fathers exposed to intermediate DR showed the highest fat content (Miersch and Döring, 2012), revealing a curious metabolic plasticity phenotype to paternal nutrition status. In a study by Klosin et al, they reported that a

memory of parental exposure to high temperatures can be transmitted to future generations (Klosin et al., 2017). Using an integrated multicopy array containing *daf-21* promoter driven fluorescence reporter transgene, they showed that transgene expression was different between descendants from ancestors that were exposed high temperatures vs. controls. Furthermore, they showed that this difference in expression can be inherited through both the male and hermaphrodite germline (Klosin et al., 2017). Although the physiological relevance of the memory of heat stress remains unclear, this finding suggests equal potential for males to transmit epigenetic memories across generations. In the same year, Kishimoto et al reported that males exposed to heavy metal (arsenite), osmotic stress (NaCl), or fasting sired offspring that showed a mild increase in resistance to oxidative stress and lifespan (Kishimoto et al., 2017). In the recent study from Moore et al. mentioned in the previous section, the authors showed that *P. aeruginosa* avoidance behaviors can be passed through both the male and female germline for one generation. The authors also showed transmission of avoidance behavior through the hermaphrodite germline which lasted for four generations (Moore et al., 2019). It would be interesting to characterize the transmission dynamics of pathogenic avoidance behavior through the male and female germline for at least four generations. This can help determine whether any potential differences exist between the male and female germline in carrying the memory of ancestral pathogenic exposure that can alter offspring phenotypes.

Studies on paternal inheritance of environmental effects have been investigated at a much lesser extent in *C. elegans* compared to mammals. This is due to a key experimental challenge, i.e. majority of the population in *C. elegans* are hermaphrodites with a very low incidence of males, making it difficult to distinguish between paternal and maternal effects. However, this can be overcome by using mutant strains that either have higher incidence of males (*him* mutants) or forced to be gonochoristic by feminization of the hermaphrodite germline (*fog* mutants). The use of *fog* mutants will be further discussed in the following sections. With the wide range of functional genetic tools available, studying paternal effects in *C. elegans* present an opportunity to dissect the genetic pathways that underlie TEI.

The Evolution of Gene Expression

Changes in gene function or gene expression are thought to contribute to changes in phenotype between and within species. Although species-specific genes may contribute to unique phenotypes, many pathways between species are conserved yet divergent traits are frequently observed. It is known that gene expression can vary greatly between species, populations, and even between individuals for the same biological process. This widespread variation is often thought to contribute to adaptation and emergence of novel traits (For detailed reviews, see Signor and Nuzhdin, 2018; Wittkopp, 2013). In essence, to study the evolution of gene expression is to analyze the mechanisms that drive gene expression divergence between and within species.

In the past decade, genome-wide investigations into DNA sequence and the transcriptome has been made possible by the advances and democratization of next generation sequencing technologies. One approach to studying the evolution of gene expression is to perform comparative genomics studies to draw evolutionary models that acted to shape and change a phenotype of interest. For example, one recent study profiled the transcriptomes of seven major organs across multiple development stages of five mammalian species and observed that transcriptomes are most similar early in development and become gradually more distinct (Cardoso-Moreira et al., 2019). Further analysis of gene expression patterns over developmental

time between human, rat, mouse, rabbit and opossum revealed genes that evolved new developmental trajectories in human compared to other species, and these genes may explain the unique organ phenotypes between species (Cardoso-Moreira et al., 2019). This type of comparative genomics approach is powerful for observing global trends in the transcriptome and identifying correlative relationships that link gene expression and phenotypic changes (Brawand et al., 2011; Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019; Stern and Crandall, 2018). However, it is difficult to dissect the underlying genetic sequence contributions of gene regulatory mechanisms that had evolved to change gene expression which ultimately lead to variation in phenotype using a comparative genomics approach.

To analyze the genetic basis of the evolution of gene regulatory mechanisms that lead to variation in gene expression, there are three main approaches: the expression quantitative trait loci (eQTL), the common reference design, and the parent/F1 hybrid system (**Figure 1-1**, **Figure 1-2**). In all approaches, the first step is to generate hybrids between or within species. These approaches make use of the natural genetic variation that exists between or within species that can potentially affect gene expression outcomes. For each approach, the samples assayed are generated using different methods and are discussed below. Finally, gene expression is measured as a quantitative trait and the aim is to resolve the genetic loci that contribute to gene expression variation.

The goal for these experiments is to compartmentalize genetic loci that contribute to variation in gene expression into *cis* and *trans* components. Local effects that affect gene expression such as linked polymorphisms in gene regulatory sequences are considered as *cis*. Upstream effects from factors that are linked or unlinked (such as transcription factors) that can act on the gene affected are considered *trans*. For *cis* regulatory effects, the expectation is that the effects are local or allele specific. For *trans* regulatory differences, at least in diploids, the expectation is that the *trans*-acting factors affect both alleles equally without *cis* regulatory differences (Fear et al., 2016; Tirosh et al., 2009).

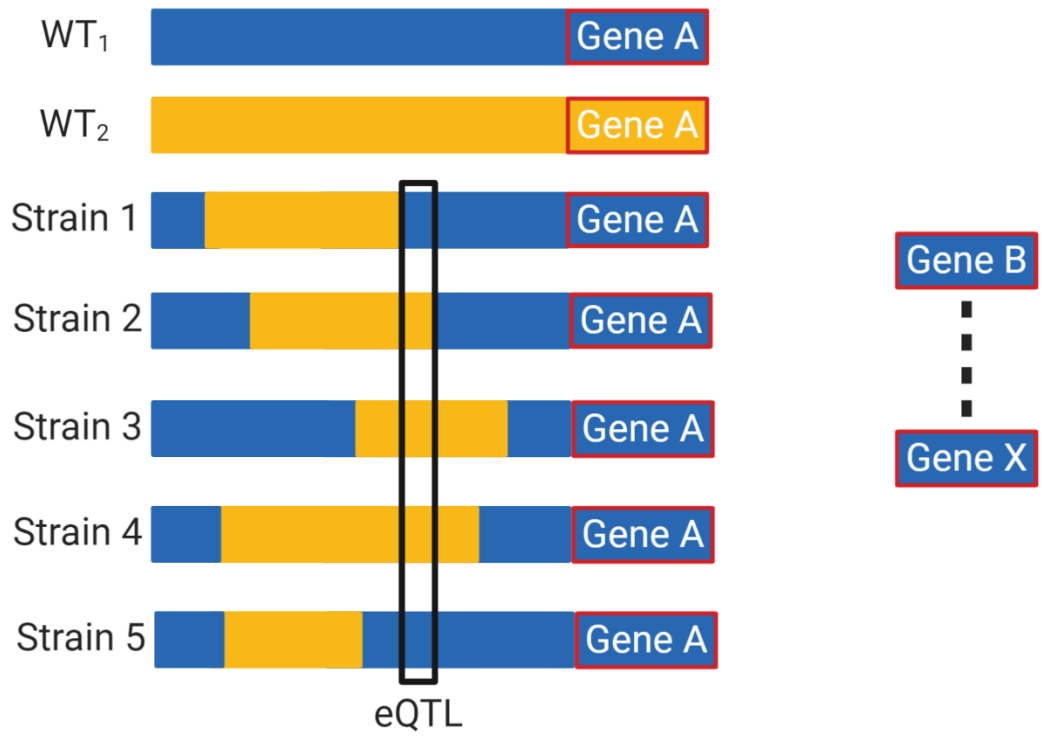
In the eQTL approach, multiple genetic variants (typically hundreds) are generated for testing by crossing two pure breeds or isogenic strains of different backgrounds to generate heterozygous F1 progeny, and crossing the F1 progeny to generate F2 hybrids where some regions would be homozygous. The resulting F2 hybrids can be used directly for eQTL studies or further crossed with its siblings or selfing for multiple generations to generate recombinant inbred lines (Broman, 2005). As a result of recombination, the strains selected for experimentation provide a combination of “blocks” of the parental genomes (as illustrated by blue and orange blocks in **Figure 1-1**). These blocks are distinguished by strain specific polymorphisms which contain variations in genetic sequence that potentially affect gene expression. Both *cis*- and *trans*-effect loci can be mapped using the eQTL approach, and it is important to note that *cis*- and *trans*-eQTL maps the physical proximity to the genetic loci that is causing variation. When an

eQTL is mapped, the region may contain one or multiple mutations that can affect gene expression (Brem et al., 2002; Francesconi and Lehner, 2014; Gilad et al., 2008; Hansen et al., 2008; Jansen and Nap, 2001; Nica and Dermitzakis, 2013; Rockman and Kruglyak, 2006).

Using genome-wide gene expression information, this approach is especially powerful for detecting single regions that affect the expression of many genes or '*trans* hotspots' that indicate the existence and location of single or multiple gene regulators. However, the resolution for mapping the genetic variation (SNPs, indels, copy number) that causes expression variation is limited by the frequency and ability of recombination events occurring at every variable loci (**Figure 3-1**; Francesconi and Lehner, 2014; Martin and Orgogozo, 2013; Tian et al., 2016; Yang and Wittkopp, 2017).

Figure 1-1

A



B

Strain	Gene A	Gene B	...Gene X
WT_1	100	100	WT-levels
1	100	100	WT-levels
2	5	100	WT-levels
3	5	100	WT-levels
4	5	100	WT-levels
5	100	100	WT-levels

C

Strain	Gene A	Gene B	...Gene X
WT_1	100	100	WT-levels
1	100	100	WT-levels
2	5	5	Altered
3	5	5	Altered
4	5	5	Altered
5	100	100	WT-levels

Figure 1-1. Illustration of cis and trans expression quantitative trait loci (eQTL).

A) The rectangle represents a chromosomal region that contains gene A and genetically linked upstream regions. Gene B to Gene X represents unlinked or distal genes. Two hypothetical parental strains blue (WT_1) and orange (WT_2), and a panel of 5 segregant strains generated from WT_1 and WT_2 crosses (Strains 1-5) are shown. The rectangular blocks in segregants represent different combinations of genetic sequences originating from each parent as a result of recombination events. The black rectangle indicates a region that contains an eQTL.

B) Hypothetical gene expression results if the eQTL region is a *cis*-eQTL. The numbers represent expression level. In this case, only the expression of gene A is affected when the linked locus contains sequence variations, therefore representing a *cis*-eQTL since it acts only on the linked gene A and not unlinked genes B to X.

C) Hypothetical gene expression results if the eQTL is a *trans*-eQTL. In this case, the expression of linked and unlinked genes (A, B to X) are affected when this region contains sequence variations, therefore this is a *trans*-eQTL and it is a hotspot since it affects the expression of multiple genes.

The common reference design and the parent/F1 hybrid approaches are similar in that they are both based on directly assaying gene expression in F1 offspring as compared to segregants that have undergone recombination in eQTL approaches. In the common reference design approach, a tester strain is crossed with different variant strains generating a panel of heterozygous F1 progeny where half their genomes are identical. Next, these heterozygous F1 are assayed for gene expression and each variant allele is compared relative to the tester allele. Within each experiment, *cis* effects are identified by comparing the relative expression ratios between the tester allele and the variant allele, and *trans* effects are identified by comparing the tester allele across different experimental groups (Fear et al., 2016; Nuzhdin et al., 2012) (**Figure 1-2A**).

In the parent/F1 hybrid approach, gene expression in F1 hybrids are compared against its two homozygous parents. To uncover *cis* and *trans* effects, the basic analysis is to compare the ratio of expression between parents and between each allele within the F1 hybrid. Depending on experimental design, there can be at least three distinct nuclear environments between individual groups, i.e. the nucleus that harbors the genome of each parent (P1 or P2) and the nucleus that harbors the hybrid genome (F1) (**Figure 1-2B**). It is reasonable to perform analysis only on F1 hybrids generated by crossing from one direction, e.g. male P1 crossed with female P2, however the nuclear environments generated male P2 crossed with female P1 should be considered as distinct from the reciprocal cross. This is due to the potential for parent-of-origin and genomic imprinting effects that

may provide a distinct *trans* environment between hybrids generated from reciprocal crosses. In this system, *cis* effects are identified when the expression ratio is maintained in the hybrid as in parents, indicating that allele-specific genetic architecture alone is sufficient to drive the expression of a target gene at the observed levels irrespective to different nuclear environments. For *trans* effects, it is when expression in the hybrids deviate from the parental range, indicating *trans* regulatory differences of a target gene by factors beyond genetic architecture (i.e. *trans*-acting factors) (Landry et al., 2005; Li and Fay, 2017; McManus et al., 2010; Tirosh et al., 2009; Wittkopp et al., 2004) (**Figure 1-2B**).

In both common reference design and the parent/F1 hybrid systems, knowing allele-specific expression is necessary for decomposing *cis* and *trans* effects, therefore the power in these analyses are restricted to only genes that contain polymorphisms. Nonetheless, the parent/F1 hybrid approach is most powerful in providing insight into compensatory *cis-trans* interactions, a type of interaction where gene expression is regulated by *cis* and *trans* mutations that act in opposing directions (Goncalves et al., 2012). These compensatory *cis-trans* interactions are thought to stabilize overall gene expression variation within individuals (Signor and Nuzhdin, 2018). These compensatory interactions are perhaps the result of selection acting to minimize pleiotropy which could lead to extreme or deleterious phenotypes within or between populations.

Figure 1-2

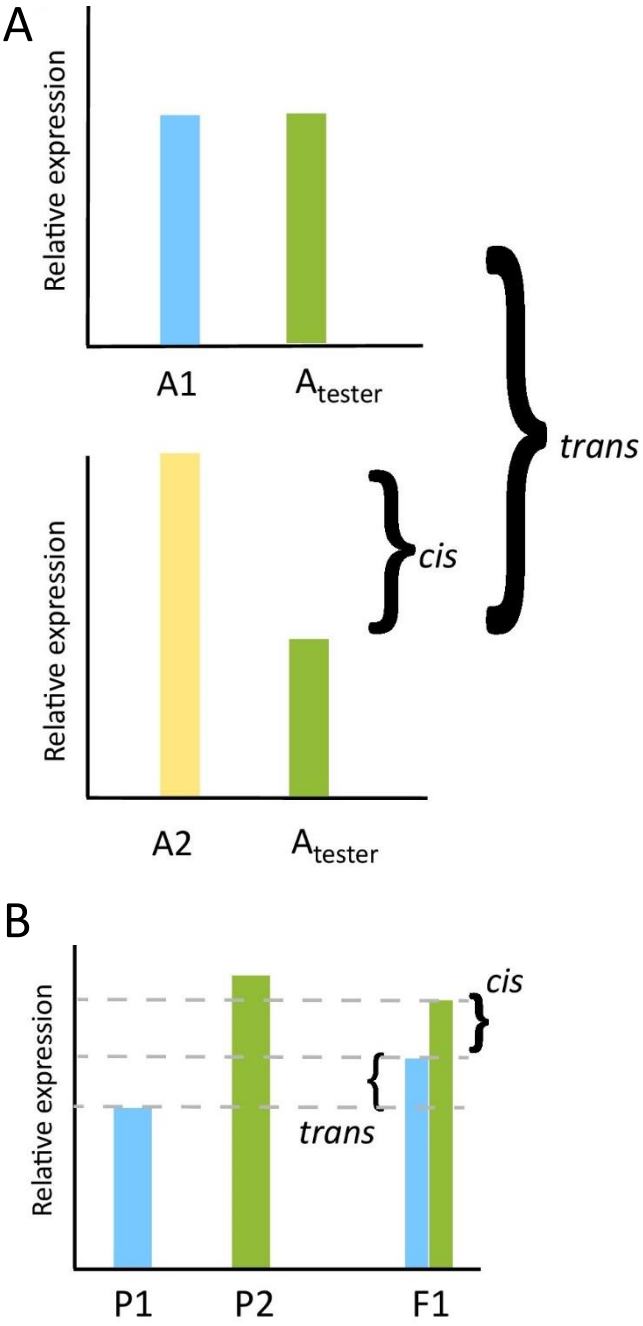


Figure 1-2. *cis* and *trans* effects in gene regulation in common reference design and parent/hybrid F1 approaches.

A) In the common reference design, a panel of individuals from a population sample is crossed to a single 'tester' strain (shown in green as A_{tester}). Differences between the expression of the 'tester' allele and the population

alleles (A1 and A2–blue and yellow, respectively) within an individual are *cis* effects, while differences in the expression of the tester allele between individuals are *trans* effects. In the top panel, there is no *cis* effect as the A_{tester} and A1 allele are expressed at the same level. In the bottom panel, there is a *cis* effect between A2 and A_{tester}, and there is a *trans* effect because the A_{tester} allele is expressed at a different level in different individuals (between the A1 and A2 background). Note that this does not measure all *trans* effects, only those originating from the A1 and A2 backgrounds but not the A_{tester} background. This section of the figure is based off of a figure from Fear *et al*.

B) The earliest approach to *cis*–*trans* decomposition was to characterize expression of parental alleles in an F1 hybrid, often using techniques such as pyrosequencing. In the F1 hybrid differences in the expression of only one allele relative to the other allele is a *cis* effect. *trans* effects can be detected by comparing the expression ratio of each allele in hybrids to the ratio of expression between parents – if the ratio of expression is different in the F1 hybrid, this is a *trans* effect. Here the bracket labeled *cis* indicates the differences between the two alleles in the F1. The *tran*s bracket measures the amount of expression which is altered relative to the expression of the parental alleles. *From (Signor and Nuzhdin, 2018). Reprinted with permission from Elsevier.*

Extensive number of studies have characterized the natural genetic variation in *cis* and *trans* regulatory elements and their contributions to variation in gene expression within and between species in yeast (Brem et al., 2002; Li and Fay, 2017; Tirosh et al., 2009), *Drosophila* (Fear et al., 2016; Landry et al., 2005; Wittkopp et al., 2004), mice (Goncalves et al., 2012; Mack et al., 2016; Shen et al., 2014), plants (Rhoné et al., 2017; Shi et al., 2012), birds (Wang et al., 2017), insects (Wang et al., 2016), fish (Verta and Jones, 2019). However, very few studies on evolution of gene expression have been performed in *C. elegans* (Denver et al., 2005; Snoek et al., 2017), as *C. elegans* lack a sister species that can generate viable hybrids for evolutionary studies. Majority of linkage mapping and quantitative trait loci studies depend on within species analysis, especially with the Hawaiian strain CB4856 (Burga et al., 2019; Rockman and Kruglyak, 2009). In Chapter 3, I analyze the evolution of gene expression within species using a common lab reference strain N2 and the Hawaiian strain CB4856. The experimental strategy used will be described in detail below.

Key experimental models and tools:

C. elegans has been an indispensable model for developmental biology due to its robust and invariable developmental trajectory (Kimble and Hirsh, 1979; Sulston and Horvitz, 1977; Sulston et al., 1983) and the swath of molecular tools and resources perfectly suited for genetic and molecular analysis of organismal development.

Post-embryonic Development, L1 Developmental Arrest and Population Synchronization

Under replete conditions, *C. elegans* undergo a life cycle trajectory of four larval stages starting from L1, L2, L3 and L4 before reaching adulthood (**Figure 1-3**). The life cycle is typically 3 days starting from fertilization to adulthood at 20°C. From the heroic effort completed on cell lineage tracing, early *C. elegans* biologists traced the developmental trajectory of every cell starting from the one-cell embryo to the full adult and found that each cell follows an invariant developmental path (Kimble and Hirsh, 1979; Sulston and Horvitz, 1977; Sulston et al., 1983).

Developing *C. elegans* larvae can enter developmental arrest under unfavorable environmental conditions. For example, starvation-induced developmental arrest is reversible and can occur during all stages of larval development (Baugh, 2013; Hu, 2007; Schindler et al., 2014; Seidel and Kimble, 2011). Investigations into the regulation of developmental arrest have

led to key insights into characterizing metabolic signaling pathways, such as insulin-like signaling (Murphy and Hu, 2013), and TGF-β signaling (Gumienny and Savage-Dunn, 2013).

While various developmental arrest paradigms serve as key models for studying a multitude of biological processes, *C. elegans* researchers also make use of such developmental arrest protocols to generate synchronous populations for experiments (Lewis and Fleming, 1995). For example, rather than progressing through development, embryos that hatch as first larval stage (L1) worms in the absence of food will instead enter L1 developmental arrest. A routine method performed by *C. elegans* researchers makes use of this L1 arrest phenomenon to generate developmentally-synchronized populations for experiments. The general protocol is to bleach gravid adult animals to obtain embryos, followed by hatching and a short period of starvation in buffer to generate arrested L1s; arrested L1s can then resume development synchronously upon being provided with food (Emmons et al., 1979; Lewis and Fleming, 1995).

Figure 1-3

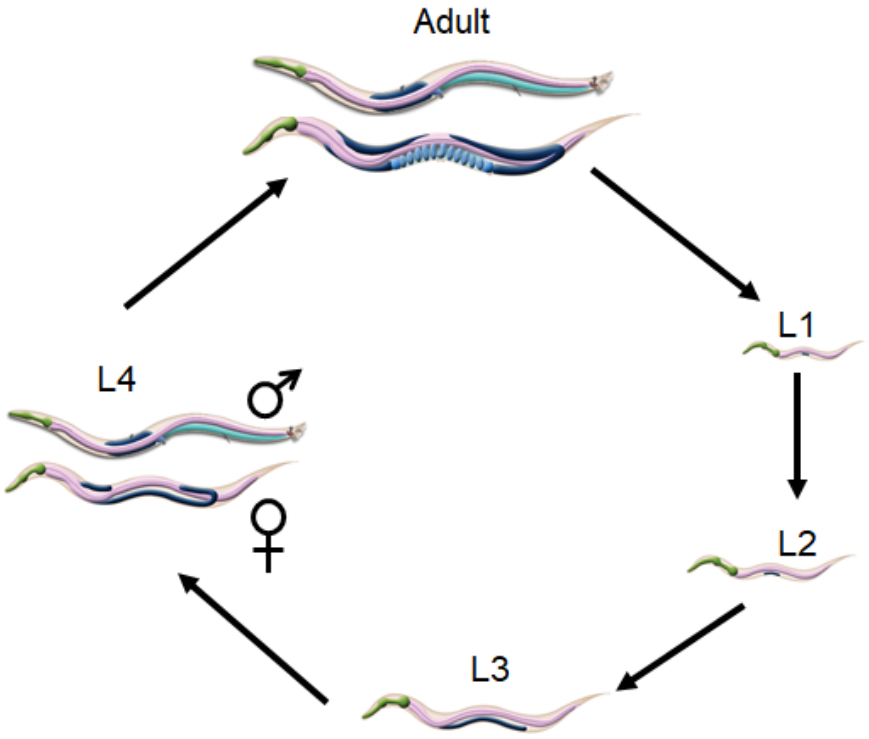


Figure 1-3. The life cycle and post-embryonic development of *C. elegans* under standard conditions

Starting from first larval stage (L1), *C. elegans* complete post-embryonic development through stages L2, L3, L4. In this diagram, the two different sexes at L4 larval stage were separately labelled to emphasize that sex is most easily distinguished at this stage under a dissecting microscope. Note that sex is determined at birth. This is a key factor for experimental work. Images were adapted from WormAtlas.org.

The *fog-2* genetic background

Wild-type *C. elegans* are diploid and sex is determined by the number of X chromosomes, with hermaphrodites having two X chromosomes (XX) and males having one X chromosome (XO). Having two sexes provide a powerful system for genetic analysis because it allows outcrossing. However, in a typical wild-type population, males occur only at <0.2% making it challenging for large-scale screens that require a population of purely cross-fertilized animals (Corsi et al., 2015).

The *fog-2* (feminization of the germline) alleles and strains were first isolated in 1988 by Schedl and Kimble while mapping pathways that regulate sex determination in the germline (Schedl and Kimble, 1988). During the L4 stage, both the hermaphrodite and male germline undergo spermatogenesis. While males continuously produce sperm from L4 to the rest of adulthood, hermaphrodites switch from spermatogenesis to oogenesis at the onset of adulthood (L'Hernault, 2006). Loss-of-function mutations in *fog-2* only affects spermatogenesis in hermaphrodites and not males (Ellis and Schedl, 2007; Schedl and Kimble, 1988). In this work, both studies make use of the *fog-2* genetic background in which hermaphrodites do not generate sperm, therefore reproduction relies on cross-fertilization by males. The strain is maintained as a gonochoristic male/female population, thus providing a convenient system for generating purely cross-fertilized hybrids for genetic and paternal/maternal effects analysis.

The Hawaiian strain (CB4856)

The Hawaiian strain, CB4856, was first isolated from a pineapple field in Hawaii by L. Hollen in 1972 (Hodgkin and Doniach, 1997). In stark geographical contrast, N2, the common reference strain, was derived from an isolate found in mushroom compost near Bristol, England (Nicholas et al., 1959). Apart from their ecological histories, N2 and CB4856 exhibit many phenotypic differences as listed below, and worm researchers take advantage of their highly divergent genomes to study the genetic mechanisms of biological processes. This is achieved by generating recombinant inbred lines and recombinant inbred advanced intercross lines by crossing N2 and CB4856, and these strains have been used for mapping mutations, linkage mapping and identifying quantitative trait loci (Burga et al., 2019; Davis and Hammarlund, 2006; Doroszuk et al., 2009; Rockman and Kruglyak, 2009; Shelton, 2006; Wicks et al., 2001). This approach has shed light on the genetics of vast number of biological phenomena, including but not limited to, aggregation and foraging behavior, unique mating features, temperature tolerance, germline RNAi, defense against parasites, phoretic behavior, drug resistance, and genetic incompatibility (De Bono and Bargmann, 1998; Brady et al., 2019; Burga et al., 2019; Kammenga et al., 2007; Lee et al., 2017; Okahata et al., 2016; Palopoli et al., 2008; Schulenburg and Müller, 2004; Seidel et al., 2008, 2011; Tijsterman et al., 2002).

CB4856 has been and continues to be one of the most genetically divergent strains compared to N2 (Andersen et al., 2012; Kim et al., 2019; Thompson et al., 2015). Recently, the CB4856 genome was re-sequenced

with Pacific Biosciences (PacBio) RSII long-read sequencing platform (Kim et al., 2019). The results reveal multiple new insights into the genomic differences between CB4856 and N2 that were invisible to short-read sequencing technologies that were used to generate previous genome assemblies (Thompson et al., 2015; Vergara et al., 2014). First, the final assembled genome is 103 Mb compared to the last assembled genome at 98.2 Mb (Thompson et al., 2015). Second, the CB4856 genome contained 1 SNP in 500 bp compared to N2. Third, there are substantial structural variations in CB4856 that affect 2694 genes, and the authors identified 300 new genes between N2 and CB4856 totaling more than 600 strain-specific genes in addition to previous findings (Maydan et al., 2007, 2010). Fourth, numerous rearrangements (translocations, inversions) in the range of 10 to 100 kb were revealed on multiple chromosomes, totaling more than 4.95 Mb. Fifth, at more than half of chromosome ends, new subtelomeric regions ranging from 11k to >200k bp were discovered. In all, this study showed that considerable structural variations in genomes can be tolerated within *C. elegans* species, and provide a view of extraordinary genetic diversity within a species.

In Chapter 3, I take advantage of the great genetic diversity between CB4856 and N2 and implement a strategy to investigate the evolution of gene expression within a species.

Chapter 2: Effects of Larval Density on Gene Regulation in *C. elegans* During Routine L1 Synchronization

Introduction

Bleaching gravid *C. elegans* followed by a short period of starvation of the L1 larvae is a routine method performed by worm researchers for generating synchronous populations for experiments. However, details of the culture conditions used for population synchronization by L1 arrest – such as, e.g., worm density – are not commonly reported in the literature, and may have effects on organismal physiology. In the course of investigating paternal dietary effects on gene regulation in *C. elegans*, we uncovered massive batch effects in an L1-stage RNA-Seq dataset. As the density of arrested L1 animals has been reported to affect starvation survival in an experimental setting that is similar to the routine population synchronization method (Artyukhin et al., 2013), we hypothesized that failure to control L1 arrest density in our initial studies resulted in the RNA-Seq batch effects.

In this Chapter, we set out to systematically characterize genome-wide effects of L1 density on gene expression using single worm RNA-seq. We characterized mRNA abundance genome-wide in two genetic backgrounds –

wild-type N2 animals as well as the gonochoristic *fog-2* mutant – arrested in L1 at 1, 5, 20, or 100 eggs/μL. We identified 53 genes, primarily encoding various metabolic enzymes and potential signaling molecules, that were robustly affected by larval density in both strains. Specifically, a number of genes related to metabolism and signalling are highly expressed in worms arrested at low density, but are repressed at higher arrest densities. We confirmed these findings in detail for the uncharacterized *lips-15* gene, generating a promoter *Plips-15::gfp* reporter strain as a sensor for L1 density. We showed that conditioned media from high density L1 cultures was able to downregulate *lips-15* even in L1 animals arrested at low density. Finally, results from conditioned media experiments reveal that a *daf-22* ascaroside-independent pathway mediates this chemical communication between L1s, and suggest the possibility of an unknown density-sensing system in *C. elegans*.

Together, our data implicate a soluble signalling molecule in density sensing by L1 stage *C. elegans*. We reveal a robust molecular phenotype that correlates with larval density during L1 developmental arrest, demonstrating that this variable must be carefully controlled when performing L1 arrest for molecular studies and provide guidance for design of experiments focused on early developmental gene regulation.

Results

Effects of arrest density on gene expression during L1 starvation

To systematically identify density-regulated genes in arrested L1 larvae, we used low-input RNA-Seq (Ramsköld et al., 2012; Trombetta et al., 2014) to characterize the transcriptome of individual L1 animals arrested at different densities (**Figure 2-1A, Figure 2-2**). Briefly, we generated arrested L1s for experiments by bleaching gravid adults to collect embryos, and resuspended the asynchronous embryos at four different densities (1, 5, 20, 100 eggs/μL) in M9 media supplemented with polyethylene glycol 3350 (0.5%, w/v) to prevent animals from sticking to tips and tubes lacking food. Embryos hatched in M9 enter L1 arrest due to absence of food, and were collected as single L1s 22 hours later for RNA-seq (**Figure 2-1A**, **Methods**). As our initial studies were performed in the gonochoristic *fog-2* mutant background – in which hermaphrodites do not make sperm and are therefore phenotypically female (Schedl and Kimble, 1988) – we repeated the dilution series and RNA-Seq in wild-type N2 animals to ensure that the results do not reflect physiological quirks of the *fog-2* mutant (**Figure 2-3, File 2.1**).

We first compared RNA-Seq profiles for L1s hatched at the highest and lowest densities (100/μL vs. 1/μL). As results in the *fog-2* and N2 backgrounds were essentially identical (**File 2.1 and 2.2**), we merged these datasets and identified a core set of 53 significantly density-regulated genes (**Figure 2-1B**, **Table 1**). A small number of genes were upregulated under

high-density growth arrest, including two collagen genes (many other collagen-related genes were also upregulated in these conditions but were not individually significant). In contrast to the moderate changes in gene expression for density-activated genes, several of the density-repressed genes were affected ~10-100-fold by arrest density. Most notably, we find that four relatively uncharacterized genes located in a cluster on chromosome II – *nspe-1*, *nspe-2*, *lips-15*, and *lips-16*, encoding two predicted lipases and two potential signaling peptides – were highly-transcribed in low-density conditions, but were either undetectably expressed or expressed at low levels at high density. Other density-repressed genes encoded signaling molecules (*daf-5*, *ins-26*, *sgk-1*, *gpa-12*, *dmd-7*), metabolic enzymes (*sams-1*, *pyk-2*, *ahcy-1*, *fat-2*, *cah-4*, *cyp34-A4*), and a handful of uncharacterized genes (*F35E8.13*, etc.). Although we do not further characterize the physiology of animals maintained at high density, we predict that lipid metabolism – in particular, phosphatidylcholine metabolism (Walker et al., 2011) – is likely to be substantially altered in these animals.

Figure 2-1

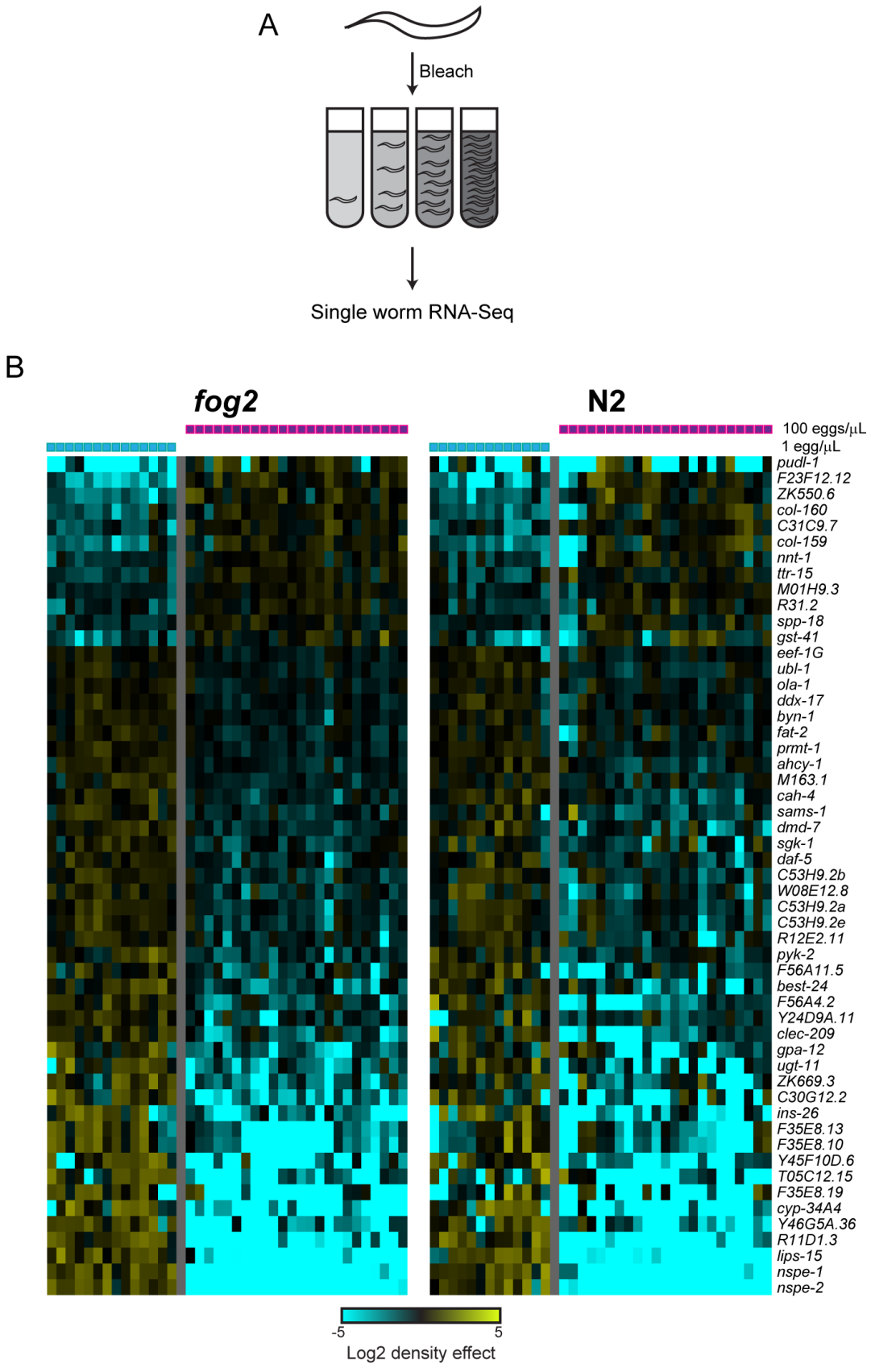

A
Bleach
Single worm RNA-Seq
B
fog2
N2
100 eggs/µL
1 egg/µL
pudl-1
F23F12.12
ZK550.6
col-160
C31C9.7
col-159
nnt-1
ttr-15
M01H9.3
R31.2
spp-18
gst-41
eef-1G
ubl-1
ola-1
ddx-17
byn-1
fat-2
prmt-1
ahcy-1
M163.1
cah-4
sams-1
dmd-7
sgk-1
daf-5
C53H9.2b
W08E12.8
C53H9.2a
C53H9.2e
R12E2.11
pyk-2
F56A11.5
best-24
F56A4.2
Y24D9A.11
clec-209
gpa-12
ugt-11
ZK669.3
C30G12.2
ins-26
F35E8.13
F35E8.10
Y45F10D.6
T05C12.15
F35E8.19
cyp-34A4
Y46G5A.36
R11D1.3
lips-15
nspe-1
nspe-2
-5
5
Log2 density effect

Figure 2-1. Effects of L1 arrest density on larval gene expression.

A) Experimental schematic. Gravid *C. elegans* adults are bleached to recover a mixed population of early embryos. Embryos are then washed in M9 buffer containing PEG, counted, and resuspended at the indicated densities in 3 mL of M9. 22 hours postbleaching, individual L1 animals were picked into separate tubes and processed for single-worm RNA-Seq.

B) Significant effects of L1 arrest density on mRNA abundance. Heatmap shows clustered RNA abundance data for 53 genes significantly differentially-expressed ($p_{adj} < 0.05$) between 1 and 100 egg/μL cultures. Data are shown separately for replicate experiments using N2 hermaphrodites, or matings between *fog-2* mutant males and “females” (genotypically XX hermaphrodites which are incapable of making sperm and therefore phenotypically female), as indicated. In both datasets, mRNA abundance for individual L1s is expressed as the $\log_2$ ratio relative to the average for that gene across all individuals in the relevant strain background.

Figure 2-2

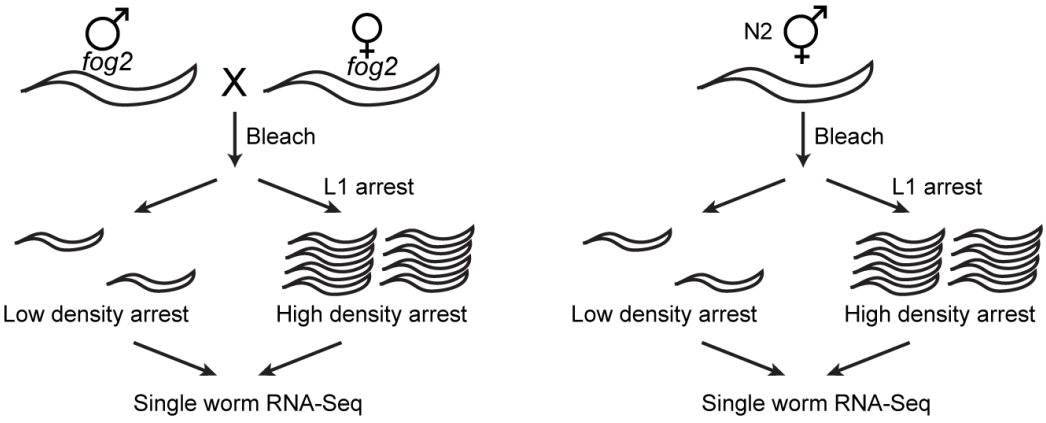


Figure 2-2. Experimental scheme.

Comparison of N2 and *fog-2* background density experiments.

Figure 2-3

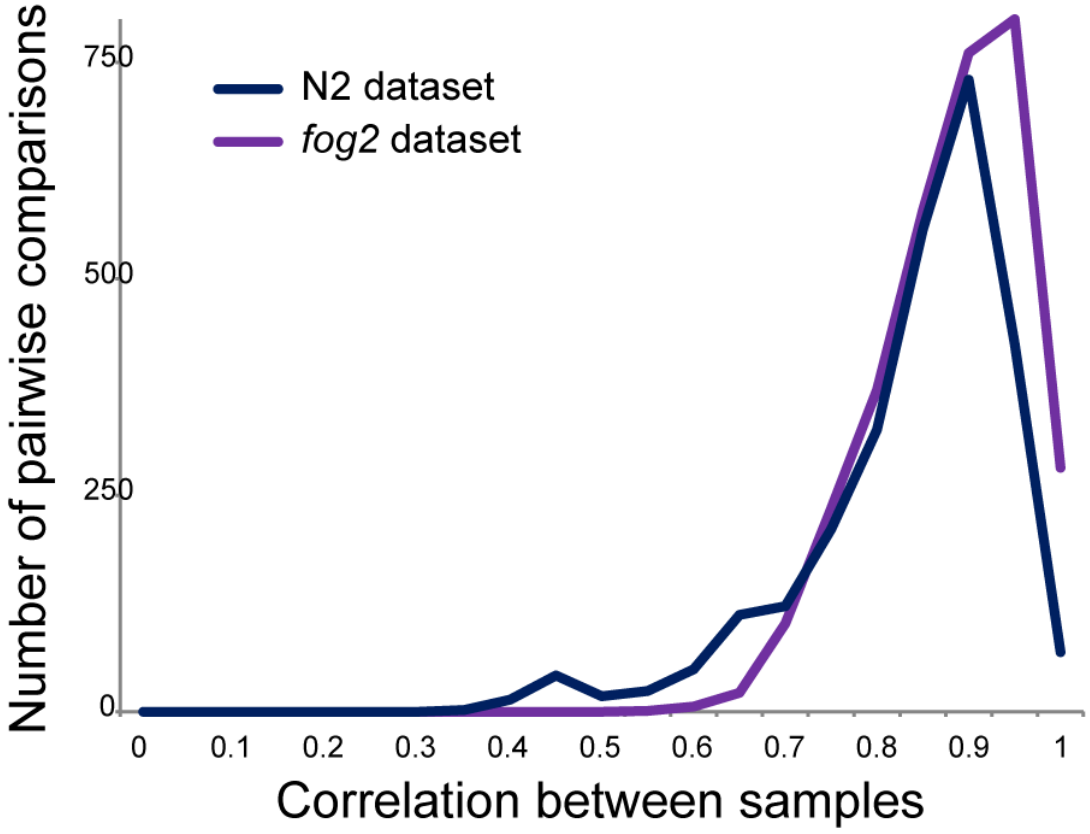


Figure 2-3. Deep sequencing reproducibility.

For each experiment, we calculated pairwise correlation coefficients between individual sequencing libraries for all pairwise combinations of animals. Histograms show the distribution of correlation coefficients for each dataset, as indicated. Note the presence of a small number of outliers in the N2 dataset.

An L1 density reporter

To further validate the dramatic effects of arrest density on larval gene expression, we sought to develop a GFP reporter for L1 density. We focused on *lips-15*, which is highly-expressed under low density conditions (>1,000 tpm) and was nearly undetectable (median expression of 0 tpm, average of ~20 tpm) in worms arrested at high densities. We therefore generated transgenic strains by integrating a promoter *Plips-15::GFP* fusion construct into the genome on chromosome II using MosSCI (Frøkjaer-Jensen et al., 2008) (**Figure 2-4A**), again creating reporter lines in the N2 and *fog-2* backgrounds. We assessed the effects of L1 arrest density on reporter gene expression, resuspending reporter embryos at 1 or 100 eggs/μL. Consistent with the dramatic repression of *lips-15* mRNA abundance observed in high density L1 cultures, we confirmed that *Plips-15::*GFP expression was strongly reduced in high density L1s (**Figure 2-4B-D**). Density-dependent regulation of this reporter was also observed in animals hatched into food, demonstrating that effects of density are not confined to conditions of starvation (**Figure 2-5**).

Interestingly, the *Plips-15::GFP* reporter exhibits a surprisingly specific localization pattern at low density, with robust GFP expression confined to the excretory cell (**Figure 2-4B**). This cell serves a function analogous to the kidney of mammals (Buechner et al., 1999; Nelson and Riddle, 1984), providing a further hint that the high density-repressed genes are involved in organismal metabolism. These results provide independent validation of our

RNA-Seq dataset, and provide a robust single-worm reporter for L1 arrest density.

Figure 2-4

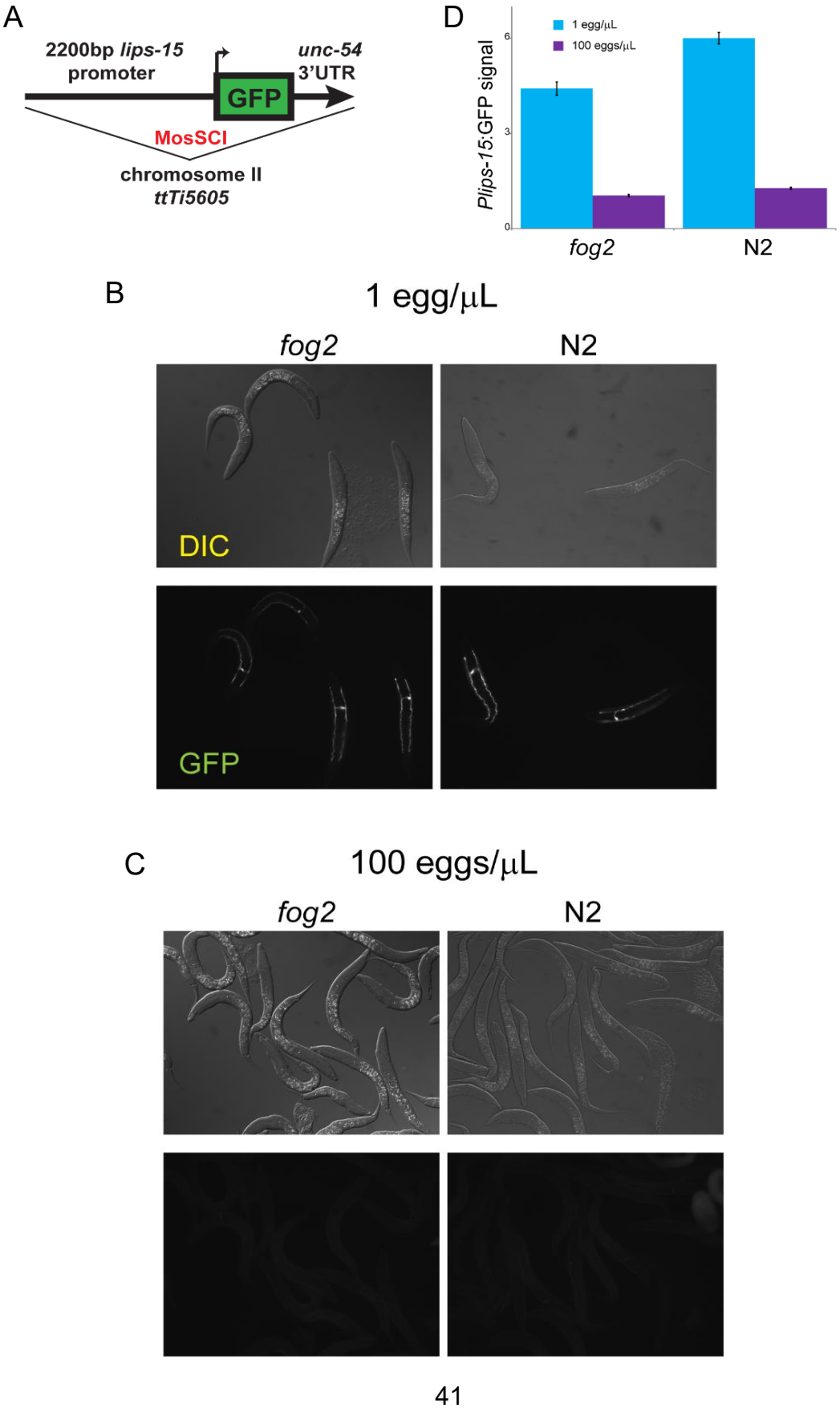

A
2200bp lips-15 promoter
unc-54 3'UTR
GFP
MosSCI
chromosome II
ttTi5605
D
1 egg/µL
100 eggs/µL
Plips-15:GFP signal
6
3
0
fog2
N2
B
1 egg/µL
fog2
N2
DIC
GFP
C
100 eggs/µL
fog2
N2

Figure 2-4. A GFP reporter for L1 density sensing.

A) Schematic of the reporter for *lips-15* promoter activity.

B) Typical images of *Plips-15::gfp* reporter expression in animals (of the indicated genetic background) arrested at low density. GFP expression is confined to a single large cell with long lateral projections, characteristic of the *C. elegans* excretory cell.

C) As in (B), for reporter animals arrested at 100 eggs/μL.

D) Quantitation of GFP expression at two densities in two strain backgrounds – data are shown as mean +/- s.e.m. for low and high density *fog-2* animals (n=206 and 230, respectively), and low and high density N2 animals (n=181 and 304, respectively).

Figure 2-5

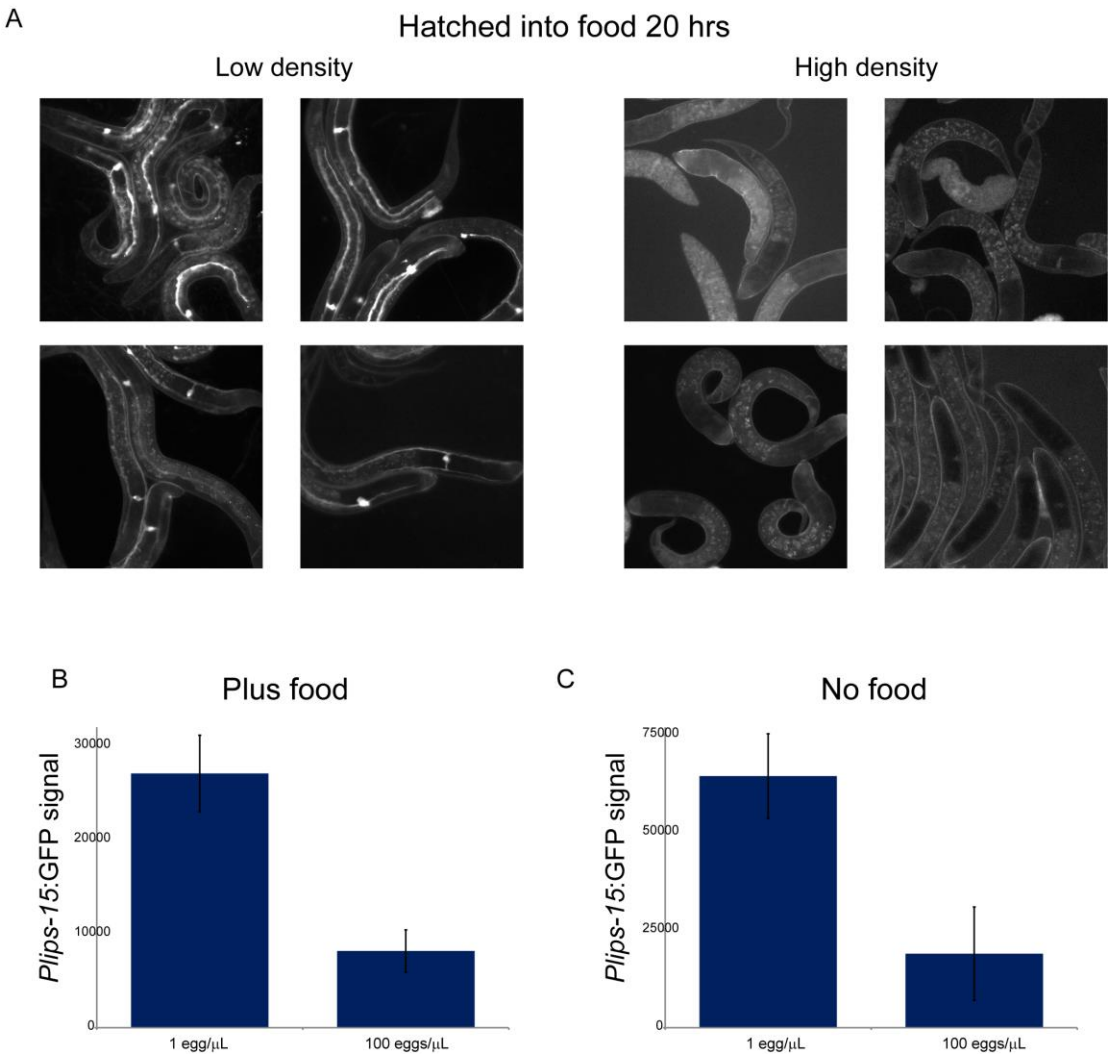


Figure 2-5. Density-dependent lips-15 repression is not starvation-dependent.

A) GFP images for *Plips-15::gfp* animals at 20 hours post-bleaching at low density (1 egg/µL) or high density (100 eggs/µL) in M9 containing OP50 *E. coli* as food.

B-C) Quantification of GFP expression at low (1 egg/µL) vs. high (100 eggs/µL) density, with food (B) or without food (C) as control.

Data show mean +/- standard deviation of the background-corrected fluorescence intensity values for three biological replicate experiments, with 50-100 worms quantitated for each biological replicate.

Density signaling is mediated by a soluble factor

What is the nature of the signal that mediates density signaling? To further characterize the nature of the density-dependent signal, we examined the expression of density-regulated genes in animals arrested at intermediate L1 densities (5 and 20 eggs/μL). In general, we find that expression of these genes changes continuously across the four densities used, rather than exhibiting a switch-like transition at some density (**Figure 2-6**). Because our data are single-L1 RNA-Seq data rather than being an ensemble measure from pooled animals, the continuous changes in gene expression reveal a tunable response being mounted in individual animals, rather than changes in the frequency of phenotypic subpopulations.

Figure 2-6

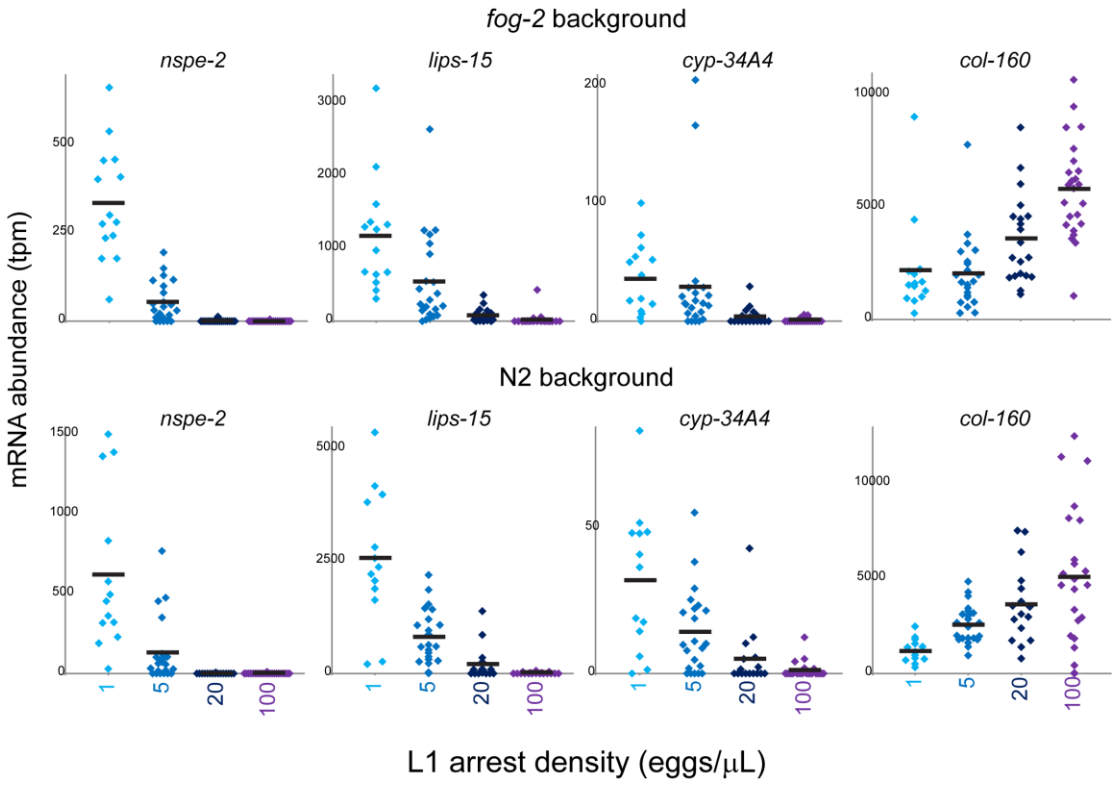


Figure 2-6. Continuous effects of arrest density on target gene expression.

Dot plots show data for the indicated genes in individual L1s arrested at 4 different densities.

Is density signaling mediated by physical contact between animals, or by a soluble factor produced (or consumed) by worms? A previous study demonstrated that conditioned media generated from L1 larvae arrested at high density, provided to low-density L1s, is sufficient to phenocopy a long-term survival phenotype of L1s arrested at high density (Artyukhin et al., 2013). We therefore assayed *Plips-15::GFP* expression in animals bleached and maintained at low density in either control media or in conditioned media produced by L1-arrested animals cultured at high density (**Figure 2-7A**). As before, we observed robust expression of the reporter in animals maintained at low density in control media (**Figure 2-7B**). However, this expression was extinguished when L1s were arrested in high density conditioned media (**Figure 2-7C**). Repression by conditioned media was reversible, as 1) animals maintained at high density for 22 hours were capable of inducing *lips-15* expression when diluted 100X into control media conditions, and 2) animals hatched at low density in conditioned media were able to induce reporter expression after being washed and transferred to control media (**Figure 2-7D**). Activation of GFP occurred within hours in both cases, demonstrating that even under conditions of developmental arrest, L1 worms can sense and rapidly respond to changes in population density. Together with the observation that low density conditioned media could not activate *lips-15* in animals arrested at high density (not shown), our data reveal a density-sensing pathway that is mediated by a soluble factor produced by arrested *C. elegans* L1 larvae. We note that we cannot distinguish in this study between a signal produced only in animals maintained at high density,

and a signal produced constitutively by L1 larvae but which is present in low density cultures at concentrations too low to activate the high density response.

L1 density gene expression effect is mediated by a *daf-22*/ascaroside independent pathway

C. elegans development and behavior are regulated by a variety of small molecule signals. Prominent among these are the ascarosides, a diverse family of ascarylose sugar-based pheromones which regulate an array of phenotypes such as dauer entry and exit (Ludewig , A and Schroeder, F, 2013). To test the hypothesis that the density effects observed here might be mediated by ascaroside signaling, we generated conditioned medium from high density L1 cultures of *daf-22* mutants, which do not produce any wild-type ascarosides (Butcher et al., 2009b; Von Reuss et al., 2012). Conditioned media from *daf-22* mutants robustly repressed *lips-15* expression (**Figure 2-7E**), inconsistent with the hypothesis that ascarosides are responsible for this case of density signaling.

Figure 2-7

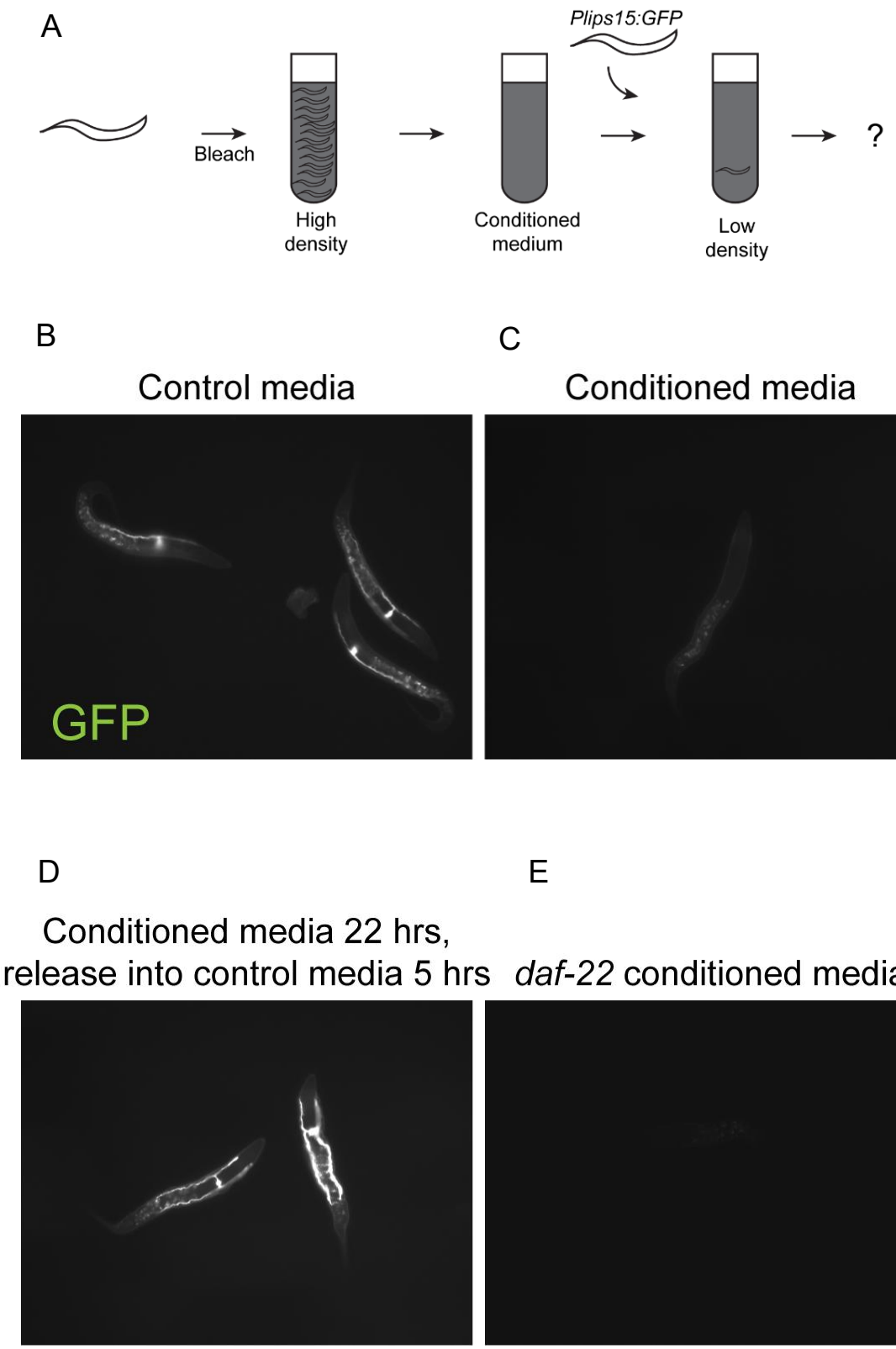

A
Plips15:GFP
Bleach
High density
Conditioned medium
Low density
?
B
C
Control media
Conditioned media
GFP
D
E
Conditioned media 22 hrs, release into control media 5 hrs
daf-22 conditioned media

Figure 2-7. Density effects are mediated by a soluble, ascaroside-independent signal.

A) Schematic of conditioned media experiments. B-E) GFP images for *Plips-15::gfp* animals maintained at 1 egg/μL in control media (B), high density L1-conditioned media (C), and conditioned media from high-density arrest cultures of *daf-22* mutant animals (E). In (D), reporter animals were arrested in conditioned media as in (C), then washed and suspended at low density in fresh control media for five hours.

Conclusions

Taken together, our data reveal that the density of starved *C. elegans* L1-stage larvae significantly affects gene expression, with some genes responding ~10-100-fold to animal density. Further, we identify a highly plastic response to changes in population density that can occur within 5 hours and is driven by a cryptic *daf-22* independent signaling pathway (**Figure 2-7D-E**). These data also demonstrate that arrest density is an important experimental variable to control in molecular studies of L1-stage *C. elegans*. For example, recent pioneering studies of single-cell transcriptomes in developing *C. elegans* were characterized by substantial gene expression variability beyond the expected technical variation, which the authors suggested might be ascribed to developmental timing or preparation of the larvae (Cao et al., 2017).

L1 density signaling in *C. elegans*

Several recent studies have shown that L1 arrest density can alter various phenotypes in *C. elegans*. For example, L1 density plays a role in parent-offspring signaling, as the presence of L1 larvae on plates was recently shown to alter food-related behaviors in adults (Scott et al., 2017). In this case, the relevant signal is lost in *daf-22* mutant, indicating that ascarosides are likely responsible for L1 density signaling. In contrast, an ascaroside-independent signal has been shown to mediate the effects of density on the

survival of L1 worms under prolonged starvation (far longer than the time frame of our study) (Artyukhin et al., 2013). As in our study, the effects of L1 density on starvation survival could be transferred by conditioned media, and were not mediated by ascarosides. L1 arrest density has longer-lasting effects on worm physiology as well, as population density during larval development, but not density experienced as adults, alters later development rates and adult lifespan (Ludewig et al., 2017). As with density effects on L1 starvation survival, the density signal in this study was produced in *daf-22* mutants, again supporting the existence of a non-ascaroside signal produced by L1s.

It seems likely that the target genes identified here respond to this previously-described high-density signal, as we find that *lips-15* expression is suppressed by a density signal present in conditioned media from both wild-type and from *daf-22* mutants. Although we do not identify the signaling molecule(s) produced (or consumed) by L1 animals here, it is interesting to note that two of the most highly density-regulated genes – *nspe-1* and *nspe-2* – correspond to nematode-specific genes encoding predicted neuropeptides, and their response to animal density could reflect negative feedback control of their expression. Alternatively, the prominent regulation of genes related to lipid metabolism here could be related to the previously-reported antagonism between ascarosides and the high density signal (Ludewig et al., 2017), as for example it is plausible that secreted lipases (*lips-15*, etc.) or other enzymes could alter or degrade lipid signals. That said, the high density signal reported by Artyukhin *et al* appears to be a mixture of small (<3 kD) molecules, arguing against this latter hypothesis. Whatever the nature of the signal, the robust

response of the *Plips-15::GFP* strain developed here could provide a convenient reporter for genetic screens to identify the relevant signal and its effector pathway.

Methods

C. elegans strains and maintenance.

All strains were maintained at 20 °C and passaged for at least four generations under non-starvation conditions before experimentation. The Bristol N2 strain was used as wild-type. Alleles used in this study listed by chromosome. LGII *daf-22(m130)* strain DR476*, Plips-15::gfp::unc-54 3'utr* (we are currently awaiting an allele and strain designation for these animals)*;* LGV *fog-2(q71)* strain CB4108. The *Plips-15::gfp* reporter strain was constructed using MosSCI (Frøkjaer-Jensen et al., 2008).

Strain passaging and bleaching.

Animals were passaged by bleaching gravid worms and directly plating 125,000 eggs mixed with *E. coli* OP50 (washed and resuspended in M9 buffer to 0.33 g/mL), therefore worms never experience starvation under these conditions. Strains were maintained on 150 x 15 mm petri dishes that contained standard nematode growth media (NGM) with modification [10g agarose + 7g agar instead of 17g agar]. Gravid animals were washed from plates with M9 buffer, collected in conical tubes and centrifuged at 2000 x g for 30 seconds. Next, the supernatant was removed and worms were treated with 10 mL of bleach solution [41:6:3 ddH_2O, sodium hypochlorite (Fisher Chemical, SS290-1), 5M KOH, and kept on a rocking platform. Worms were spun down after 4 minutes and the supernatant was replaced with fresh

bleach. This procedure was repeated one more time and the suspension was visually checked to ensure all worm bodies were dissolved. The suspension was centrifuged at 2000 x g for 30 seconds, bleach solution was removed, and embryos were washed with M9 buffer supplemented with PEG 3350 (0.5%, w/v) three times before use. In this study, M9 buffer was always supplemented with PEG 3350 (0.5%, w/v) unless noted otherwise. The density of embryos in M9 were determined by manual counting.

For density experiments in the presence of food a single *E.coli* OP50 colony was picked and inoculated in 3 mL of LB media at 37°C for 4 hours, and this was used to inoculate 200 mL of LB media that was grown for 16 hours at 37°C. Next, the bacteria culture was pelleted at 3750 x g for 30 mins at 4°C and resuspended with M9, the final volume was adjusted to 10 mL including the appropriate number of embryos obtained by bleaching. Worms were cultured in 50 mL conical flasks at 20°C shaking at 180 rpm in a temperature-controlled incubator.

Construction of the *Plips-15::GFP* L1 density sensor strain

The L1 density reporter strain was constructed using MosSCI with EG6699 as the starting strain (Frøkjaer-Jensen et al., 2008). The promoter region of *lips-15* was cloned using forward primer ATTTTTCACACAGAATTCCA and reverse primer with overhang into GFP GTGAAAAGTTCTTCTCCTTTACTCATGATGGAGTGAAGATTGTGGAG, and *GFP::unc-54 3'UTR* was cloned from genomic DNA of *Pacdh-1::GFP* (Arda et

al., 2010) using a forward primer with homology to the promoter region of *lips-15* CTCCACAATCTTCACTCCATCATGAGTAAAGGAGAAGAACTTTTCAC, and reverse primer AAACAGTTATGTTTGGTATATTGG. The PCR products were gel purified and used as templates for PCR stitching. Next, the correct sized band was gel purified, inserted into a Zero Blunt TOPO cloning vector (ThermoFisher 450245), and fusion products from positive clones were determined by manual sequencing. The fusion product was then subcloned and inserted into the pCFJ350 plasmid, which was used for transformation by microinjection.

Single L1 RNA-seq.

Single L1 animals were placed in 5 ul of a typical worm lysis buffer (40mM Tris pH7.5, 10mM EDTA, 200mM NaCl, 0.5% SDS, and 0.4ug/ul Proteinase K), incubated at 55° for 10 minutes, and then frozen at -20°C until use for RNA-sequencing. For RNA-seq, a Smart-Seq2 protocol for generating cDNA followed by Nextera tagmentation was used to generate adapters and indexes required for sequencing (Trombetta et al., 2014). 96 samples (animals) were pooled and sequenced at a time on the NextSeq500 system (Illumina). Data were mapped using RSEM after removing rRNA and PCR duplicates.

For *fog-2* single L1 sequencing, 14 animals at 1 egg/μL passed our QC filters, with 22 animals at 5 eggs/μL, 20 animals at 20 eggs/μL, and 24 animals at 100 eggs/μL. For the N2 single L1 sequencing, animal numbers were 13, 22, 16, and 23. Average sequencing depth was 5.5 million

reads/animal across all individuals, with ~1 million reads typically aligning to the transcriptome (RSEM) after removal of PCR duplicates and rRNA reads (**File 2.1**).

L1 conditioned medium experiment.

High density conditioned medium was generated by first bleaching gravid animals grown on plates to obtain embryos, followed by resuspension of those embryos at high density (100/µL) in M9. At 24 hours, the actual density of hatched L1s was confirmed to be between 80-100/µL before processing for collection. To collect high density conditioned medium, we centrifuged the buffer containing L1s at high density at 2000 x g for 30 seconds, and passed the supernatant through a 1 µm glass fiber membrane syringe filter (Pall Life Sciences, Cat. # 4523T) to completely remove L1s and unhatched embryos. Filtered conditioned medium was stored at -80°C in 15 mL conical tubes (Corning #430791). Frozen Conditioned Medium was thawed in a 20°C water bath for at least 30 minutes prior to use.

To perform L1 conditioned medium experiments, embryos were obtained from *Plips-15::GFP* by bleaching gravid animals as described in this above. The density of embryos was counted and diluted to low density (1/ µL) in 3 mL of high density conditioned medium or control (M9) in 15 mL conical tubes. The tubes were placed on a rocking platform for 22 hours, and the density of L1s were confirmed before processing for imaging.

Fluorescence imaging and quantification.

Plips-15::GFP reporter animals were used for all imaging experiments. Animals were first hatched in the appropriate conditions, centrifuged at 2000 x g for 30 seconds then the pellet containing L1 worms was carefully removed and transferred to a PCR tube. Worms were treated with 1:10 volume of 10 mM tetramisole hydrochlorite (Sigma L9756) for 5 minutes before mounted onto a 2% agarose pad with a cover slip for Nomarski and fluorescence imaging using a Zeiss Axioplan2 Microscope.

For initial characterization of the reporter, animals were hatched in either low density (1/µL) or high density (100/µL) conditions in 3 mL of M9. For testing conditioned medium, animals were hatched at low density (1/µL) in either 3 mL of M9 (Control) or high density conditioned medium.

Image analysis was performed using Fiji/ImageJ to calculate background-corrected fluorescence, with the mean intensity of areas without worms or embryos taken as background, and the foreground signal calculated for the excretory cell bulb region.

Gene Ontology analysis

Functional enrichment analysis and identifying statistically significant Gene Ontology (GO) terms for biological pathways (**Table 2**) and KEGG pathways (**Table 3**) were performed using g:Profiler (Reimand et al., 2016).

Acknowledgements

We thank L.R. Baugh, A. Artyukhin, M. Olmedo, and members of the Rando lab for helpful discussions. We thank members of the C. Mello and A.J.M. Walhout laboratory for sharing reagents, imaging equipment, protocols, and many helpful discussions. We thank C. Baer (UMass SCOPE facility), S. Guharajan, and Y. Shang for assistance with microscopy and image data analyses during the revision process. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). This work was supported by NIH R01 HD080224. CCC is a Merck Fellow of the Helen Hay Whitney Foundation.

Table 1. 53 significantly density-regulated genes

Name	Description
ttr-15	Transthyretin-like protein 15 [Source:UniProtKB/Swiss-Prot;Acc:Q22288]
lips-15	LIPaSe related [Source:UniProtKB/TrEMBL;Acc:Q9NAK4]
nspe-1	Nematode Specific Peptide family, group E [Source:UniProtKB/TrEMBL;Acc:Q9NAJ8]
byn-1	Cell adhesion protein byn-1 [Source:UniProtKB/Swiss-Prot;Acc:Q20932]
Y46G5A.36	None
F23F12.12	None
F35E8.10	None
F35E8.13	None
col-159	COLlagen [Source:UniProtKB/TrEMBL;Acc:Q20922]
F35E8.13	None
R11D1.3	None
fat-2	Delta(12) fatty acid desaturase fat-2 [Source:UniProtKB/Swiss-Prot;Acc:G5EGA5]
nspe-2	Nematode Specific Peptide family, group E [Source:UniProtKB/TrEMBL;Acc:Q7YWR0]
M163.1	None
C31C9.7	None
C30G12.2	None
gst-41	Glutathione S-Transferase [Source:UniProtKB/TrEMBL;Acc:Q966G8]
T05C12.15	None
ZK669.3	GILT-like protein ZK669.3 [Source:UniProtKB/Swiss-Prot;Acc:Q23570]
ins-26	INSulin related [Source:UniProtKB/TrEMBL;Acc:Q9XUI9]
F56A4.2	C-type LECtin [Source:UniProtKB/TrEMBL;Acc:G5EBG4]
F56A4.2	C-type LECtin [Source:UniProtKB/TrEMBL;Acc:G5EBG4]
clec-209	C-type LECtin [Source:UniProtKB/TrEMBL;Acc:G5EBG4]
nnt-1	Nicotinamide Nucleotide Transhydrogenase [Source:UniProtKB/TrEMBL;Acc:Q18031]
ZK550.6	Probable phytanoyl-CoA dioxygenase [Source:UniProtKB/Swiss-Prot;Acc:O62515]
Y24D9A.11	None
ugt-11	UDP-GlucuronosylTransferase [Source:UniProtKB/TrEMBL;Acc:O01616]
daf-5	Abnormal DAuer Formation DAF-5, a Ski oncogene homolog involved in a neuronal TGF beta pathway (71.0 kD) (Daf-5) [Source:UniProtKB/TrEMBL;Acc:G5EDM7]
F35E8.19	None
gpa-12	Guanine nucleotide-binding protein alpha-12 subunit [Source:UniProtKB/Swiss-Prot;Acc:Q19572]
best-24	Bestrophin homolog 24 [Source:UniProtKB/Swiss-Prot;Acc:P34672]
R12E2.11	Orotate phosphoribosyltransferase [Source:UniProtKB/Swiss-Prot;Acc:O61790]
pudl-1	PUD-Like protein [Source:UniProtKB/TrEMBL;Acc:Q4R163]

Y45F10D.6	None
C53H9.2	None
cyp-34A4	CYtochrome P450 family [Source:UniProtKB/TrEMBL;Acc:O61935]
C53H9.2	None
col-160	COLlagen [Source:UniProtKB/TrEMBL;Acc:Q20921]
eef-1G	Probable elongation factor 1-gamma [Source:UniProtKB/Swiss-Prot;Acc:P54412]
spp-18	SaPosin-like Protein family [Source:UniProtKB/TrEMBL;Acc:Q19837]
M01H9.3	None
pyk-2	Pyruvate kinase [Source:UniProtKB/TrEMBL;Acc:Q23539]
ubl-1	Ubiquitin-like protein 1-40S ribosomal protein S27a Ubiquitin-like protein 1 40S ribosomal protein S27a [Source:UniProtKB/Swiss-Prot;Acc:P37165]
W08E12.8	None
sams-1	Probable S-adenosylmethionine synthase 1 [Source:UniProtKB/Swiss-Prot;Acc:O17680]
prmt-1	Protein arginine N-methyltransferase 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9U2X0]
ahcy-1	Adenosylhomocysteinase [Source:UniProtKB/Swiss-Prot;Acc:P27604]
ddx-17	DEAD boX helicase homolog [Source:UniProtKB/TrEMBL;Acc:Q9XUW5]
cah-4	Carbonic AnHydrase [Source:UniProtKB/TrEMBL;Acc:Q21614]
ola-1	Obg-like ATPase 1 [Source:UniProtKB/Swiss-Prot;Acc:P91917]
R31.2	None
dmd-7	DM (Doublesex/MAB-3) Domain family [Source:UniProtKB/TrEMBL;Acc:Q8MPU2]
C53H9.2	None
sgk-1	Serine/threonine-protein kinase sgk-1 [Source:UniProtKB/Swiss-Prot;Acc:Q2PJ68]
F56A11.5	None

Table 2. Functional Annotations of Biological Processes

term_name	term_id	adjusted_p_value
metabolic process	GO:0008152	0.000295
cellular metabolic process	GO:0044237	0.000676
small molecule metabolic process	GO:0044281	0.000859
organic substance metabolic process	GO:0071704	0.001789
S-adenosylmethionine metabolic process	GO:0046500	0.017956
biological_process	GO:0008150	0.021967
cellular nitrogen compound metabolic process	GO:0034641	0.035277
one-carbon metabolic process	GO:0006730	0.042206

Table 3. Enrichment of KEGG Pathways

term_name	term_id	adjusted_p_value
Cysteine and methionine metabolism	KEGG:00270	0.015521
FoxO signaling pathway	KEGG:04068	0.027836
Metabolic pathways	KEGG:01100	0.035952
Biosynthesis of amino acids	KEGG:01230	0.0448

Chapter 3: Investigating the Evolution of Gene Expression in *C. elegans*

Overview

In this Chapter, we establish a novel model for studying the evolution of gene expression in *C. elegans* using the common laboratory strain N2 and the genetically divergent isolate CB4856 originally isolated from Hawaii (for the remaining of this Chapter and Discussion, I will use “HW” in place of CB4856). We investigate the genetic mechanism of gene expression divergence within the *C. elegans* species using the parent/F1 hybrid system approach. We characterize genome-wide *cis* and *trans* regulatory effects by analyzing allele-specific expression during distinct stages and conditions throughout *C. elegans* development between the parental strains and the intraspecific hybrids. This work provides a comprehensive view on the extent and nature of gene expression variation between highly divergent populations within the *C. elegans* species.

Results

Implementing the parent and F1 hybrid experimental system in *C. elegans*

To analyze the evolution of gene expression in *C. elegans*, we adopted the parent and F1 hybrid experimental system and collected worms during distinct stages of post-embryonic development as a way for subjecting the two distinct genomes to unique environments for measuring the direction and magnitude of *cis* and *trans* effects. This approach depends on analyzing gene expression between the pure breed homozygous parents and their F1 hybrids, therefore requiring the generation of purely outcrossed progeny. The main mode of reproduction for wild-type *C. elegans* is self-fertilization in hermaphrodites with the rare occurrence of males in the population, and the presence of self-progeny would significantly confound our analysis. To overcome this challenge, we genetically altered the mode of reproduction by using a *fog-2(q71)* mutation in which only hermaphrodites lose the ability to make sperm (Schedl and Kimble, 1988). For simplicity, the self-sterile cross-fertile hermaphrodite will be referred to as female. This ensures the generation of purely cross progeny for the F1 hybrid analysis.

We used the two most genetically divergent *C.* elegans strains as the parents: the popular reference strain N2 and the wild isolate CB4856 (HW). To construct the HW *fog-2(q71)* parental strain, we introgressed the *fog-2(q71)* allele into the HW background. The mating strategy contains four different

groups. The parental crosses: N2 *fog-2* males to N2 *fog-2* females (“NN”), and HW *fog-2* males to HW *fog-2* females (“HH”). The F1 hybrids from reciprocal crosses of each parental genetic background: N2 *fog-2* males to HW *fog-2* females (“NH”), and HW *fog-2* males to N2 *fog-2* females (“HN”). We collected samples for RNA-seq throughout development (L1, L2, L3, L4 larval stages), during developmental arrest (L1 starvation), and when differences between sexes become pronounced (L4 male and L4 female). mRNA enriched libraries were generated from 48 samples (four strains x six conditions x two biological replicates) (**Figure 3-1**).

Figure 3-1

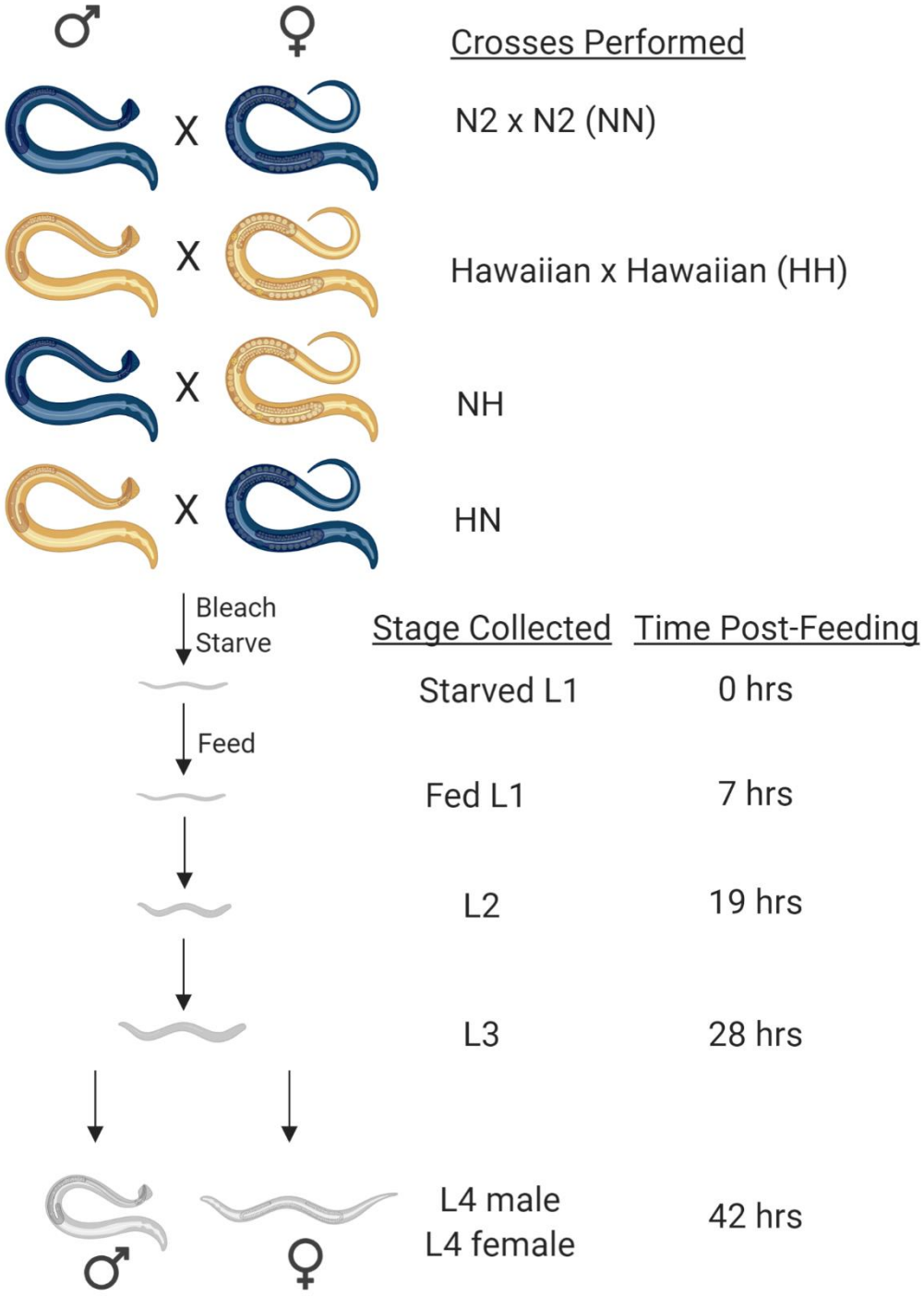

♂
♀
Crosses Performed
X
N2 x N2 (NN)
X
Hawaiian x Hawaiian (HH)
X
NH
X
HN
Bleach
Starve
Feed
Stage Collected
Time Post-Feeding
Starved L1
0 hrs
Fed L1
7 hrs
L2
19 hrs
L3
28 hrs
L4 male
L4 female
42 hrs
♂
♀

Figure 3-1. Schematic of the experimental design.

Worms used for crosses were maintained on plates for at least four generations without starvation before used for experimentation. The experiment begins by extracting embryos from a population of mixed stage worms (N2 *fog-2* and HW *fog-2*) and directly plating the embryos onto plates. Once worms reach the L4 stage (~2 days), the males and females were manually picked onto plates separately to generate hybrid crosses accordingly, and parental crosses were generated from worms left on the original plate. After crossing overnight, worms were washed from plates, bleached to extract embryos, and kept on rocking platform for 20 hours in M9 containing PEG. A portion of the Starved L1s were collected at 20 hours and the rest were plated onto plates to resume development. At the post-feeding time points, worms were visually confirmed to be at the correct stage before collection. A total of four groups (NN, HH, NH, HN) at six conditions (Starved L1, Fed L1, L2, L3, L4 male, L4 female) were collected.

Overview of the sequencing results

Gene expression overall correlated well between biological replicates (**Figure 3-2A**). Biological repeats for each stage were well correlated. Out of the six conditions, we observed clear distinction between Starved L1, L4 male and L4 female, while Fed L1, L2 and L3 were more similar in terms of overall expression of genes (**Figure 3-2A**).

In addition to analyzing expression variation by comparing allele-specific expression between N2 and HW alleles, our experimental design also allows for analysis of parent-of-origin effects on gene expression. As expected, expression of genes from the X-chromosome was biased towards the maternal allele from L1 to L3, and near 100% maternal expression in males (genetically XO thus inheriting only the maternal X chromosome) and ~50% maternal expression in females (XX) (**Figure 3-2B**). These results demonstrate that the overall experimental design is robust and suitable for downstream allele-specific expression analysis.

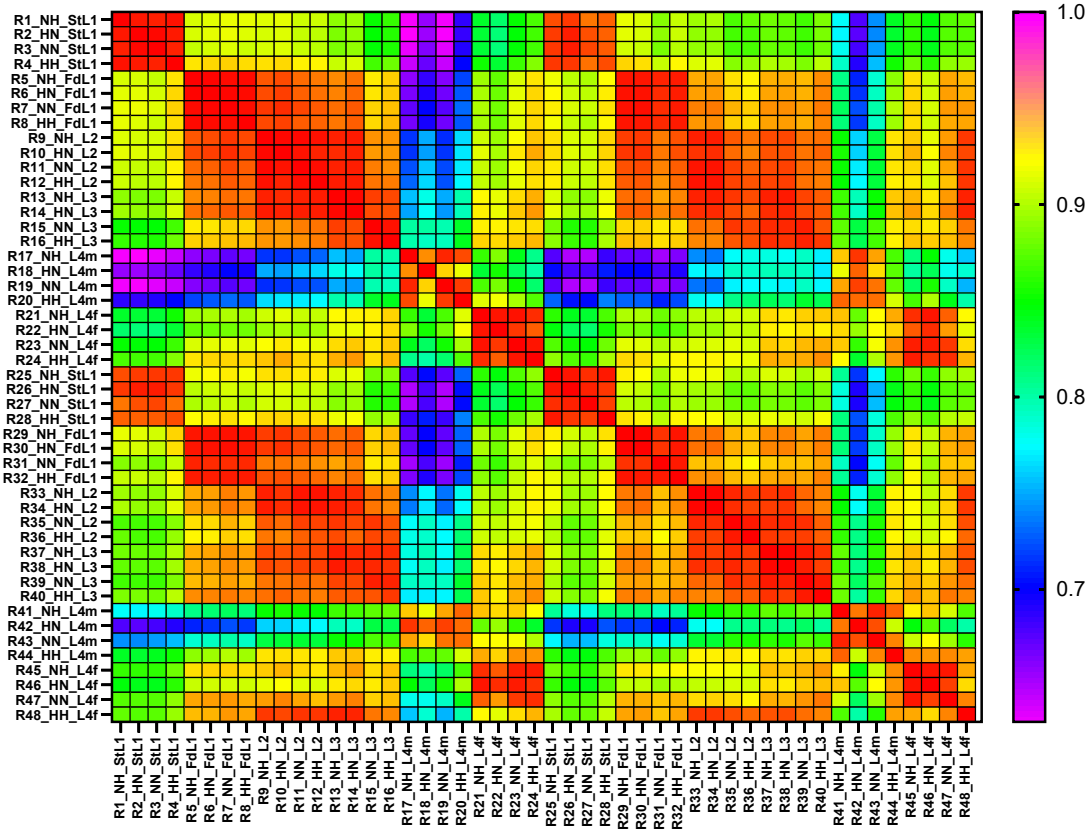
For downstream allele-specific expression analysis, the basic assumption is that the N2 and HW strains both contain the same number of cells without major structural differences between the parents and hybrids, therefore both alleles for each gene are exposed to the same nuclear environment. If differential expression is observed between the two alleles in the hybrid, this suggests functional *cis*- regulatory differences and *cis*-regulatory divergence. Same assumptions were made in previous studies that

implemented similar approaches (Landry et al., 2005; Smith and Kruglyak, 2008; Tirosh et al., 2009; Wittkopp et al., 2004, 2008).

In total, we analyzed 7532 genes that contained genetic polymorphisms. It is worthy to note that we are under-sampling the number of genes that contain polymorphisms that could have been tested. Our analysis revealed large regions in the parental HW *fog-2* genome that retained homozygous N2 sequences, which were found on almost the entire half of right arm of Chromosome V and large portions on left arm of Chromosome I. While *fog-2(q71)* is located on the right end of Chromosome V, this result is unexpected given that ten rounds of backcrossing of the N2 *fog-2(q71)* allele into the HW wild-type strain was performed. These introgressed regions were not considered in downstream analysis.

Figure 3-2

A


R1_NH_StL1
R2_HN_StL1
R3_NN_StL1
R4_HH_StL1
R5_NH_FdL1
R6_HN_FdL1
R7_NN_FdL1
R8_HH_FdL1
R9_NH_L2
R10_HN_L2
R11_NN_L2
R12_HH_L2
R13_NH_L3
R14_HN_L3
R15_NN_L3
R16_HH_L3
R17_NH_L4m
R18_HN_L4m
R19_NN_L4m
R20_HH_L4m
R21_NH_L4f
R22_HN_L4f
R23_NN_L4f
R24_HH_L4f
R25_NH_StL1
R26_HN_StL1
R27_NN_StL1
R28_HH_StL1
R29_NH_FdL1
R30_HN_FdL1
R31_NN_FdL1
R32_HH_FdL1
R33_NH_L2
R34_HN_L2
R35_NN_L2
R36_HH_L2
R37_NH_L3
R38_HN_L3
R39_NN_L3
R40_HH_L3
R41_NH_L4m
R42_HN_L4m
R43_NN_L4m
R44_HH_L4m
R45_NH_L4f
R46_HN_L4f
R47_NN_L4f
R48_HH_L4f
1.0
0.9
0.8
0.7

B

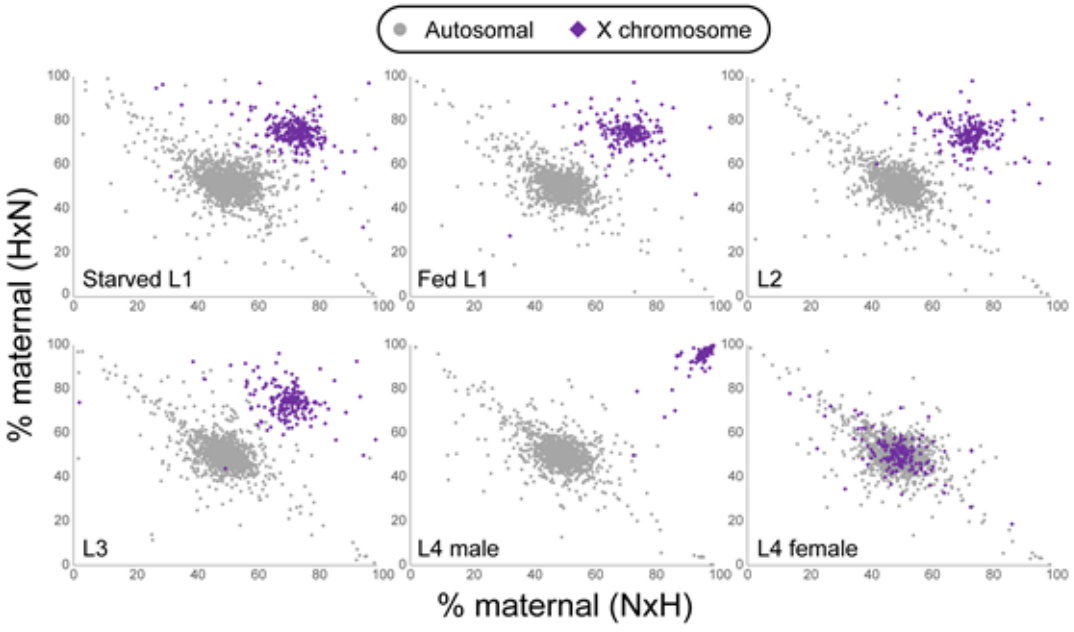

Autosomal
X chromosome
% maternal (HxN)
% maternal (NxH)
Starved L1
Fed L1
L2
L3
L4 male
L4 female
0
20
40
60
80
100

Figure 3-2. Overview of sequencing results.

A) This heatmap represents Pearson correlation of overall genetic expression between parents and hybrids with two biological replicates. Samples R1 to R24 were samples collected on a separate week compared to R25 to R48. Samples are labeled as: SampleName_Cross_Stage. St_L1 = Starved L1; Fd_L1 = Fed L1; L4m = L4 male; L4f = L4 female.

B) Each dot on this plot represents the percent maternal expression of a gene based on the cross. Percent maternal expression is calculated by taking the total number of normalized reads (transcripts per million, TPMs) from the maternal allele and dividing it by the sum of the TPMs from both the maternal and paternal alleles. The crosses are labelled as Paternal x Maternal.

Independent validation of allele-specific expression results

To examine the reproducibility and sensitivity of allele-specific gene expression across biological replicates and whether the detection of allele-specific gene expression was equally robust on autosomes, we used pyrosequencing as an independent method to validate our RNA-seq results. Previous studies used pyrosequencing to measure relative abundance of parental allele expression based on SNP differences using pyrosequencing (Landry et al., 2005; Wittkopp et al., 2004).

In this analysis, we treated hybrids from the reciprocal crosses as technical replicates of each day. We predict genes that exhibit strong *cis* effects to be biased equally in direction and magnitude. We selected candidates for pyrosequencing by calculating the fold change of normalized expression level (transcripts per million, TPM) between NN and HH (parents), and between alleles in NH or HN crosses (hybrids). The direction of allele-specific expression was determined and genes with allele-specific expression that correlated positively in direction were chosen to be tested.

We generated biological replicates independent from the samples used for RNA-seq and tested genes that exhibited expression bias towards the N2 allele (N2 *cis* effects) in Starved L1s (*rpt-2*, *rhgf-2*) and L4 male (*nspd-7*) (**Figure 3-3A**), and HW *cis* effects in Starved L1s (*nlp-20*, *K09H11.7*, *gst-39*) and at every stage and condition for *gst-39* (**Figure 3-3B-C**). We were able to detect differences by pyrosequencing for genes that were expressed as low

as ~19 TPMs (*rhgf-2*, **Figure 3-3A**) with $\log_2$ fold change as low as ~1 fold difference (*rpt-2*, **Figure 3-3A**).

From these experiments, we were able to independently confirm our analysis pipeline and experimental design, and the reproducibility and sensitivity for detecting allele-specific expression.

Figure 3-3

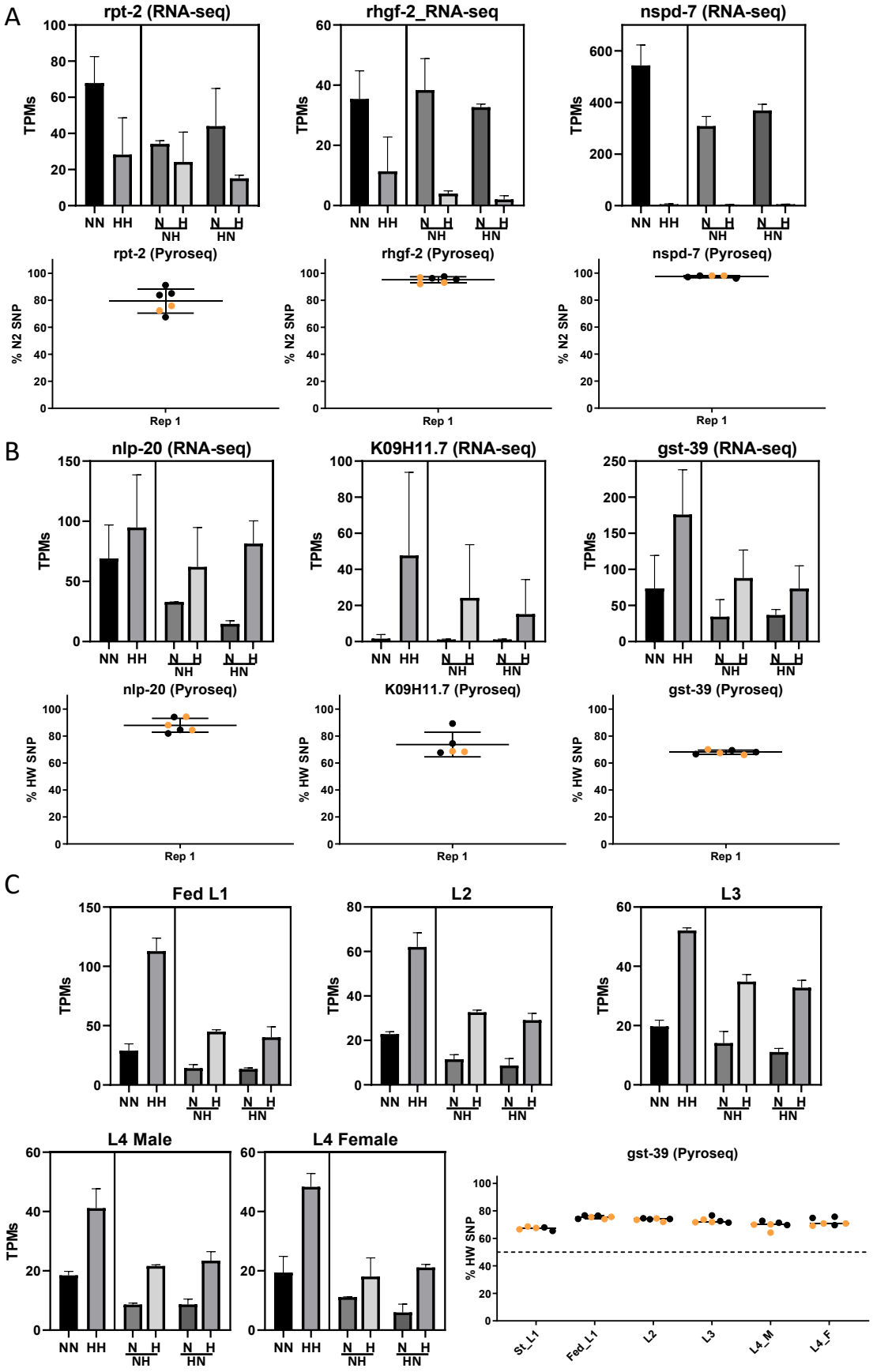

A
rpt-2 (RNA-seq)
rhgf-2_RNA-seq
nspd-7 (RNA-seq)
TPMs
NN HH N H NH N H HN
rpt-2 (Pyroseq)
rhgf-2 (Pyroseq)
nspd-7 (Pyroseq)
% N2 SNP
Rep 1
B
nlp-20 (RNA-seq)
K09H11.7 (RNA-seq)
gst-39 (RNA-seq)
nlp-20 (Pyroseq)
K09H11.7 (Pyroseq)
gst-39 (Pyroseq)
% HW SNP
C
Fed L1
L2
L3
L4 Male
L4 Female
gst-39 (Pyroseq)
St_L1
Fed_L1
L2
L3
L4_M
L4_F

Figure 3-3. Independent validation of allele-specific expression by pyrosequencing.

A) Genes that exhibit expression bias towards the N2 allele. *rpt-2* and *rhgf-2* were tested in Starved L1s, and *nspd-7* was tested in L4 male.

B) Genes that exhibit expression bias towards the HW allele. *nlp-20* and *K09H11.7* were tested in Starved L1s, and *gst-39* were tested in Starved L1 and all stages and conditions.

C) Bar charts are normalized RNA-seq results shown in transcripts per million (TPMs) comparing parental expression (NN and HH) and expression level from the N2 (N) or the HW (H) allele in the NH or the HN hybrid. Pyrosequencing results were shown in dot plots, with either the percent N2 expression (A) or % HW expression (B – C). Percent expression were taken directly from the pyrosequencing results by the percentage of SNP detected for the corresponding allele. Yellow dots represent hybrids generated from NH crosses, and black dots were hybrids from HN crosses.

Expression divergence between N2 and HW genomes

What is the extent of gene expression variation? To examine genome-wide expression divergence between the N2 and HW genomes, we first assigned an allelic imbalance (AI) score for every gene by dividing the normalized reads (TPMs) from the N2 allele by the sum of the TPMs from the N2 and the HW allele and comparing the AI score between the parents and hybrids in the six conditions (**Figure 3-4**). Comparison of the expression between the parents is technically not an allelic imbalance per se, but the final AI score reflects the difference in expression between N2 and HW genomes or alleles.

In this analysis, AI scores that are closer to 1 indicates that the expression is biased towards the N2 allele, and AI scores closer to 0 indicates a bias towards the HW allele. From visual inspection (**Figure 3-4A-G**) and quantification (**Table 4**, **Figure 3-5**) of genes with significant allelic imbalance, we identified three main classes of expression divergence based on AI scores and direction of allele-specific expression (Landry et al., 2005; Wittkopp et al., 2004):

cis: The direction of expression divergence is the same for parent and hybrid. If a gene falls into the top right quadrant, it indicates that gene expression is strongly biased towards the N2 allele thus exhibiting a N2-*cis* effect. If a gene falls into the bottom left quadrant, it is exhibiting a HW-*cis* effect.

trans: Regulatory differences in *trans* predicts that expression divergence is observed in parents, but the effects are equalized when both alleles reside in

the same *trans*-regulatory environment. Genes that fall under this category will fall into the top- and bottom-center of the vertical axis.

cis-trans: When expression divergence is observed in the hybrids but not in the parents, for example, parental score = 0.5, and hybrid AI scores near 0 or 1. This is likely the effect of compensatory *cis-trans* interactions that result in allele-specific over- or under-expression in the hybrids.

AI scores were used to quantify the number of significant AI genes that showed *cis*, *trans* and *cis-trans* effects. Using hybrid AI scores on the X-axis and parents AI scores on the Y-axis, the following table provides the scoring system for each expression divergence category:

Table 4. Expression divergence scoring system

	Hybrid	**Parent**
N2-cis	X greater than 0.66	Y greater than 0.66
HW-cis	X less than 0.33	Y less than 0.33
Trans	X is greater than 0.33 AND X is less than 0.66	Y is less than 0.33 OR Y is greater than 0.66
Cis-trans	X is less than 0.33 OR X is greater than 0.66	Y is greater than 0.33 AND Y is less than 0.66

In total, we identified 139 unique *cis* effect genes, 60 *trans* and 26 *cis-trans* effect genes across all conditions (**Figure 3-5B**, **Table 11**). The relative contribution of *cis* and *trans* effects for each condition were generally consistent (**Figure 3-5A**). From this analysis, we found that majority of gene expression variation are driven by *cis* regulatory divergence in five out of six conditions, with L4 Female being the exception. In the L4 Female, the

decreased number of *cis* effects and increased number of *trans* effects resulted in similar number of *cis* and *trans* effects. Compared to earlier stages (~11 genes), there are increased number of genes that showed *trans* regulatory divergence in the L4 Male (37 genes) and L4 Female (31 genes) stages. One possible explanation is the inability to separate sex-specific effects in earlier stages in this experiment. It has been reported previously in yeast that *trans* effects are more condition-dependent (Smith and Kruglyak, 2008; Tirosh et al., 2009), therefore *trans* effects are masked after averaging sample populations that contain both sexes (further discussed in Conclusions). Another possibility is that there is an overall increase in the number or identity of genes that are expressed in L4 stages compared to earlier stages. Upon inspection of the number of genes detected (>10 TPMs) between conditions, there are no obvious differences between conditions suggesting that the changes in relative contributions of *cis* and *trans* effects are related to the identity of genes expressed and not the total number of genes detected (**Figure 3-5C**). In the L4 Male stage, there were examples of expression divergence in genes related to major sperm proteins (*msp-64*, *msp-71*) and a male copulatory plug gene (*plg-1*). *plg-1* fall into the HW-cis category, consistent with a well-known plugging phenotype that is exhibited in HW males but lost in N2 males (Palopoli et al., 2008). There are also other L4 stage specific genes that showed divergence, such as *prom-1* (progression of meiosis) in both L4 Male and L4 Female, or *nspd-7* (nematode specific peptide family, group D) that is enriched in the male (curated on WormBase).

The full lists of divergent genes separated into N2-*cis*, HW-*cis*, *trans*, *cis-trans* for each condition are provided in **Table 5** - **Table 10**.

What are the functions of divergent genes? To determine any functional divergence, we performed Gene Ontology Enrichment Analysis (GEA) and Tissue Enrichment Analysis (TEA) on all divergent genes (**Figure 3-5B**, **Figure 3-6**, **Table 11**). Using GEA, there is significant enrichment in genes related to biotic response/defense response/immune process (*fipr-22*, *fipr-23*, *cyp-35A5*, etc.), tetrapyrrole binding or iron ion binding (*cyp-35A5, cyp-13A5, ctl-1, cyp-34A8, etc.*), and dephosphorylation (*pgph-1*, *pgph-2*, *mtm-6*, *K09F6.3*, *Y39A3A.4*, *Y54F10BM.3*) (**Figure 3-6A**, **Table 12**). From TEA, there is significant enrichment of divergent genes being expressed in the male and in the intestine (**Figure 3-6B**, **Table 13**). In addition to digesting and metabolizing food, the intestine of *C. elegans* serves as a key interface for host-pathogen interactions as defense against ingested pathogens (Cheesman et al., 2016; Pukkila-Worley and Ausubel, 2012). This is consistent with a previous report that performed comparative genomics analysis between N2 and HW pure breeds that found unique expression patterns of genes implicated in innate immunity during post-embryonic development (Capra et al., 2008).

As mentioned above, there are examples of stage-specific expression especially in the male (**Figure 3-6B**, **Table 9**). Most of the male-specific genes, however, are not curated with gene ontology terms therefore did not show particular enrichment in GEA. Given the unique feature of plugging in

HW males, we predict that there would be many genes that are functionally related to male reproduction that provide unique features between N2 and HW male mating and physiology. In all, these results agree with conventional knowledge that immune system and reproduction genes are among the most rapidly evolving genes among species.

Figure 3-4

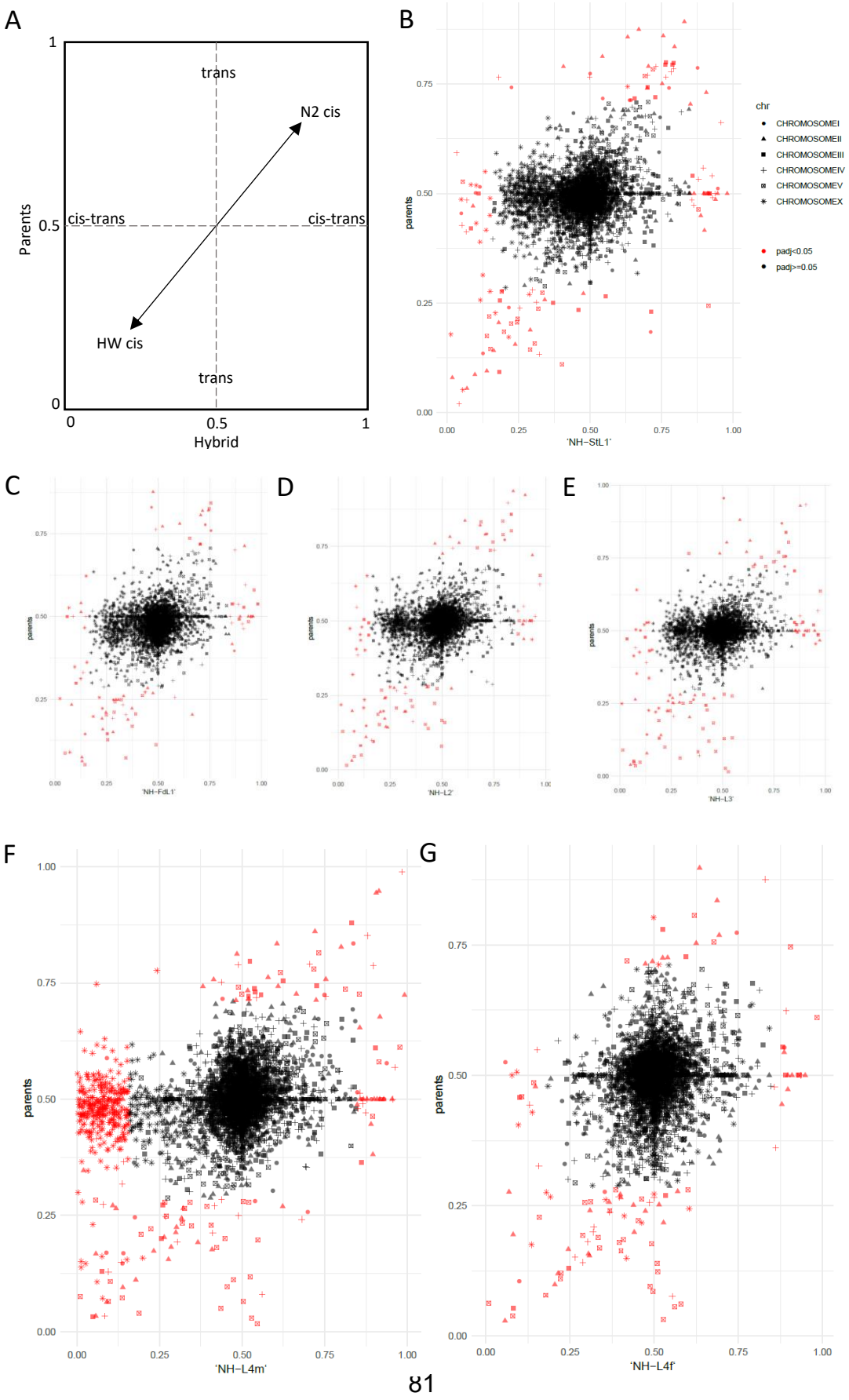

A
trans
N2 cis
cis-trans
cis-trans
HW cis
trans
Parents
Hybrid
B
parents
'NH-StL1'
chr
CHROMOSOMEI
CHROMOSOMEII
CHROMOSOMEIII
CHROMOSOMEIV
CHROMOSOMEV
CHROMOSOMEX
padj<0.05
padj>=0.05
C
'NH-FdL1'
D
'NH-L2'
E
'NH-L3'
F
'NH-L4m'
G
'NH-L4f'

Figure 3-4. Expression divergence in the parent and F1 hybrid systems between N2 and HW.

Allelic imbalance scores were calculated as follows: [NN / (NN + HH)] for the parents and [N / (N+H)] for the hybrids. Note that only the data for NH crosses are shown, but visual analysis did not indicate any qualitative differences between the NH and HN datasets.

A) With hybrid AI scores plotted on the X-axis and parental AI score on the Y-axis, classes of expression divergence are categorized into three main classes: *cis*, *trans*, and *cis-trans*.

Stages and conditions showed are:

B) Starved L1; C) Fed L1; D) L2; E) L3; F) L4 Male, note that the number of significant genes that fall to the far-left center quadrant are mostly X-chromosome genes; G) L4 Female, note that the X-chromosome genes no longer forms a cloud on the left center quadrant.

Genes highlighted in red are genes that showed significant allelic imbalance calculated based on a beta distribution and binomial test. Genes labelled black indicate there is no significant difference in AI score and suggest expression is conserved between genetic backgrounds.

Figure 3-5

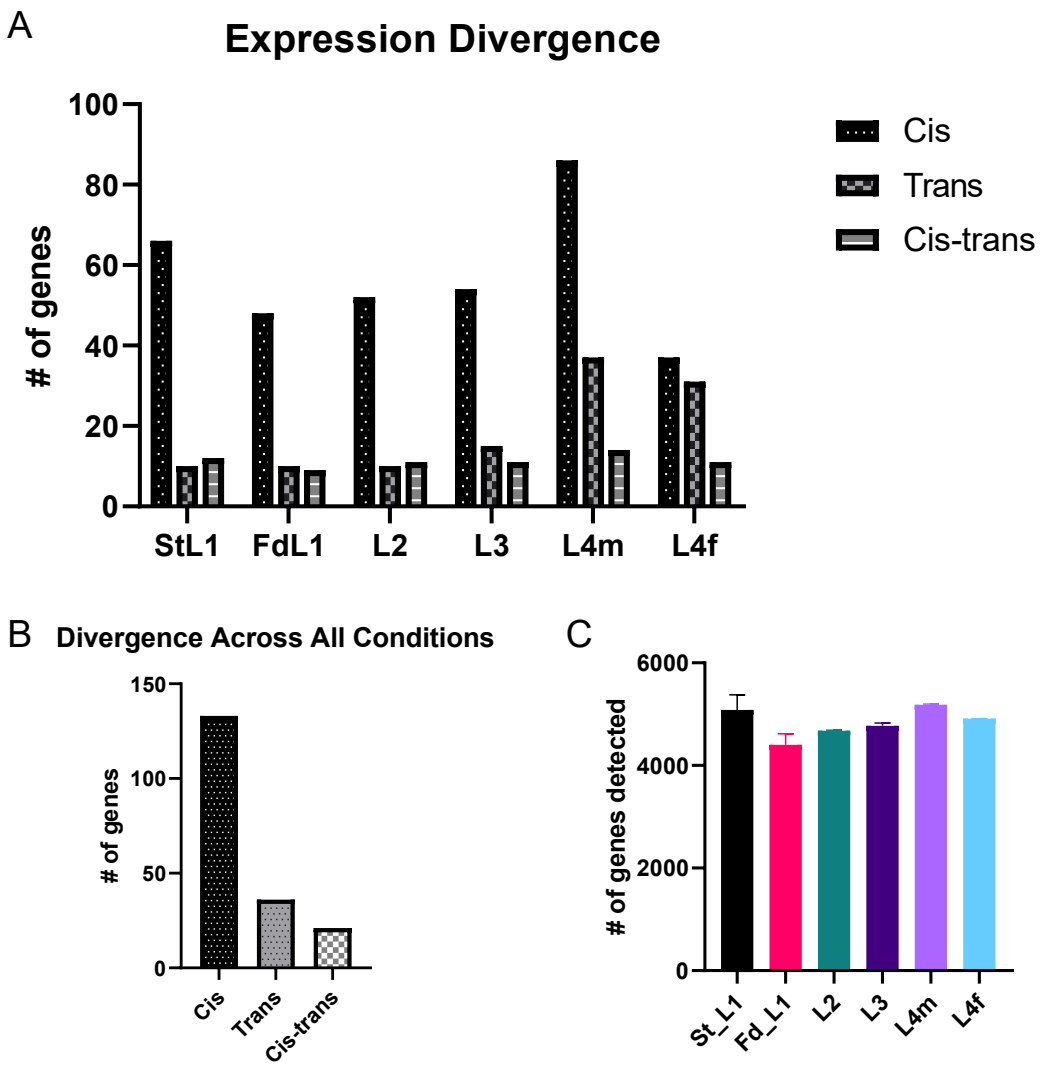


Figure 3-5. Overview of expression divergence across different conditions.

A) The number of genes that display *cis*, *trans*, and *cis-trans* regulatory divergence for each condition.

B) The cumulative number of genes that displayed *cis*, *trans*, or *cis-trans* effect at least once across all conditions.

C) The average number of genes detected with at least 10 TPMs at each condition between replicates.

Figure 3-6

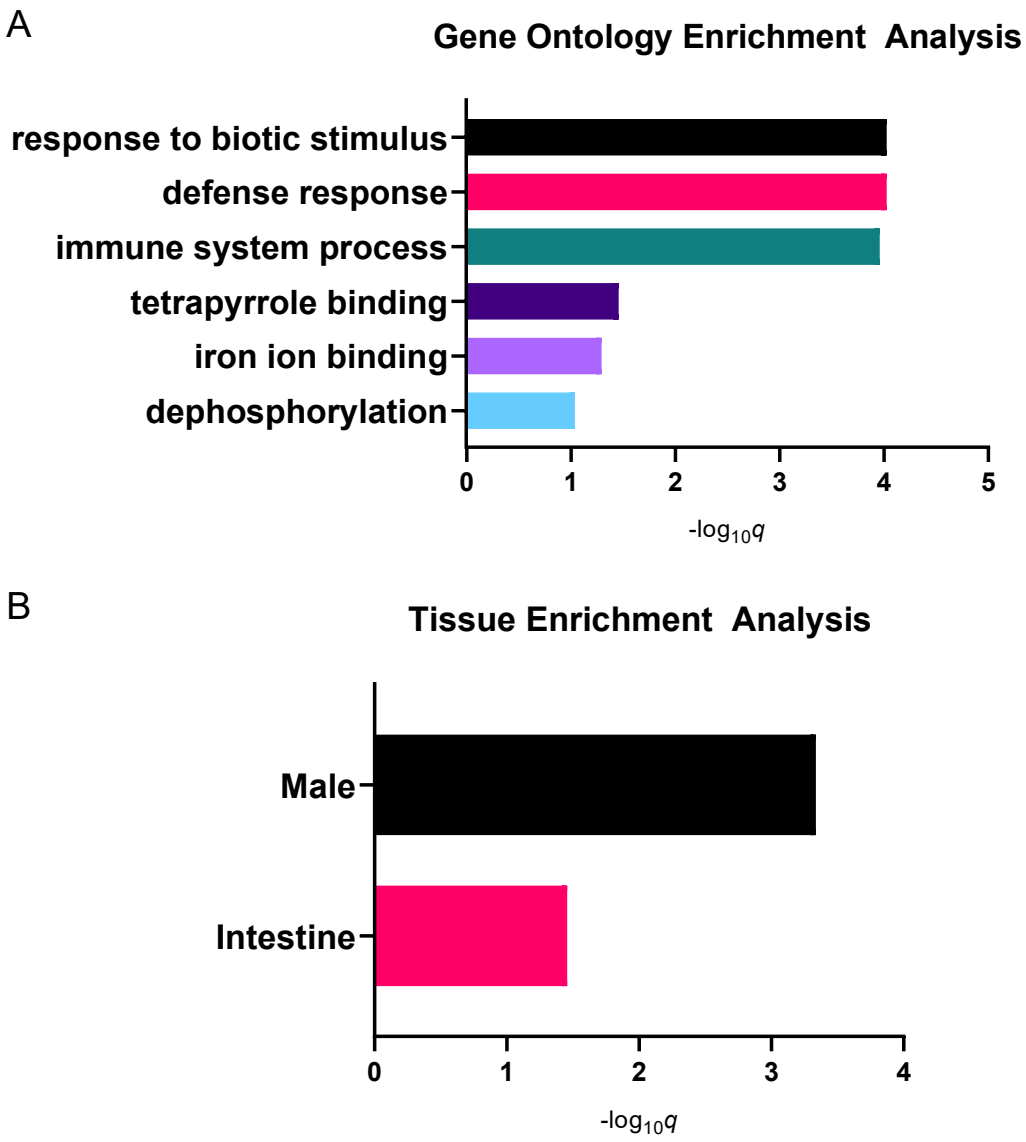


Figure 3-6. Functional annotations of divergent genes.

A) Gene Ontology Enrichment Analysis and B) Tissue Enrichment Analysis were performed using the Enrichment Analysis tool on WormBase (Angeles-Albores et al., 2018).

Conclusions

In this study, we analyzed the evolution of gene expression within the *C. elegans* species by implementing a novel approach using the parent/F1 hybrid system in gonochoristic *fog-2* mutants in the N2 and HW backgrounds. To our knowledge, this work provides the first genome-wide characterization of the evolution of gene expression in *C. elegans* using the parent/F1 hybrid system approach. Our analysis provides a comprehensive view of gene expression variation throughout development between two divergent isolates of *C. elegans*. Furthermore, this work provides a rich resource for future investigations into how gene expression variation can lead to phenotypic variation within species.

Taken together, our genome-wide analysis of expression divergence in *cis* and in *trans* provide a key insight into the evolution between the two most genetically divergent strains of *C. elegans* species. We observed that variation of gene expression across multiple conditions are predominately driven by *cis* regulatory divergence between N2 and HW. This result contrasts with previous observations in other models that showed *trans* effects to be the larger contributor to gene expression variation within species of *Drosophila* (Chen et al., 2015a; Coolon et al., 2014; Metzger et al., 2016; Wittkopp et al., 2008), yeast (Emerson et al., 2010), and plants (Rhoné et al., 2017). Our study provides new insights and a different view into within species evolution of gene expression. The predominance of *cis* regulatory differences is, so far,

a consistent feature observed in interspecific hybrids (Coolon et al., 2014; Fear et al., 2016; Goncalves et al., 2012; Mack et al., 2016; Metzger et al., 2017; Schaefke et al., 2013; Shi et al., 2012; Signor and Nuzhdin, 2018; Tirosh et al., 2009). Our findings indicate that the N2 and HW expression divergence are more similar to interspecific rather than intraspecific hybrids, suggesting the possibility where N2 and HW could be at the early stages of species divergence. However, more studies in different systems are needed to draw conclusions for the type of evolutionary mechanism that accounts for the relative contributions of *cis*, *trans*, and compensatory *cis-trans* interactions within and between species, such as, stabilizing selection, gene regulatory network feedback, or transvection as discussed extensively in a recent review (Signor and Nuzhdin, 2018).

Limitations in our current analysis

For the expression of any given gene in our dataset, current results reflect the population average of an equal mix of males and females in stages before L4. This may occlude any potential sex-specific effects that are present in earlier developmental stages. One approach to overcome this challenge would be to perform single worm RNA-seq on larvae from all developmental stages. However, it is challenging to differentiate between sexes from L1 to L3 stages, and the separation of sexes will likely be performed computationally. One straightforward prediction would be that individual samples will separate into two distinct groups by sex at each developmental stage. Based on our

single-worm RNA-seq results (**Figure 2-1, Figure 2-6**), it will require the sequencing of hundreds of animals and at high depth to obtain signal over two major sources of technical noise: allele-specific RNA-seq and low-input amount.

As mentioned in the results, we noticed the HW *fog-2* parental strain was only expressing N2 polymorphisms for large regions on Chr. I and Chr. V. Recently, a new genome assembly for CB4856 (HW) was published. Using Pacific Biosciences long-read sequencing platform, the authors identified numerous small rearrangements that ranged from 10 to 100 kb on the right arm of Chr V and a previously uncharacterized sub-telomeric region on the right arm of Chr V (Kim et al., 2019). These results suggest it is possible that these rearrangements could act to suppress recombination during backcrossing, however this does not explain the discrepancy for regions on Chr I. Various regions of the genomes are known to have different densities of sequence variations (Kim et al., 2019; Thompson et al., 2015). Therefore, it is possible that by removing these introgressed regions, the expression profile for the number of genes showing *cis*, *trans*, compensatory *cis-trans* regulatory effects could change significantly. This can be overcome by generating *fog-2* mutants by targeted genetic engineering (for example, using CRISPR/Cas9) in the CB4856 strain and re-perform the experiment for more robust analysis of evolution of gene expression in *C. elegans*.

Methods

C. elegans Strains and Maintenance

All strains were maintained and passaged under standard conditions at 20°C (Brenner, 1974). Strains were fed with *E. coli* OP50 on modified NGM plates that contained 1% agarose and 0.7% agar to prevent CB4856 from burrowing. Animals were passaged under non-starvation conditions for at least four generations before experimentation. To generate the HW *fog-2(q71)* parental strain, we introgressed *fog-2(q71)* from N2 Bristol background into the CB4856 Hawaiian background by performing ten rounds of backcrossing. We selected for the *fog-2* allele by checking whether males used for backcrossing can generate 50% hermaphrodites and 50% females by single animal test crosses into the N2 *fog-2(q71)* strain.

Preparing worms for RNA-seq and Pyrosequencing

Animals propagated under non-starvation conditions were bleached to collect embryos and plated at 200 embryos per 35 x 10mm petri dishes that contained modified NGM with OP50. This ensures that N2 and HW parental strains animals were never starved and nearly synchronized for crosses. After 2 days, 100 male and 100 female L4 animals were manually picked to fresh modified NGM OP50 plates for crosses to generate hybrids. There were two reciprocal crosses: N2 *fog-2* males to HW *fog-2* females (“NH”), and HW *fog-2* males to N2 *fog-2* females (“HN”). Animals used for parents were never manually picked, but instead left on their original plates for crossing. After one

day of crossing, female animals were visually checked to contain at least one row of embryos with some embryos laid. Next, animals were washed off plates with M9 supplemented with PEG 3350 (0.5%, w/v), and bleached to extract embryos. M9 buffer was always supplemented with PEG 3350 (0.5%, w/v). Embryos were placed on a rocking platform and hatched for 20 hours to obtain a synchronized population and this is the first collection timepoint for Starved L1s. Animals were then plated onto modified NGM plates with OP50 (0 hour, time post-feeding) and reared at 20°C until appropriate time points for subsequent collections as shown in (**Figure 3-1**).

To collect animals for RNA-seq, they were first washed off the plates with M9 into 15 mL conical tubes. Residual pieces of the NGM-agar/agarose would often contaminate the samples; therefore, all samples were next passed through a 30 µm cell strainer (MACS® SmartStrainers, #130-101-812). Next, samples were washed and centrifuged at 2,000 x g at 30 seconds for four cycles with 10 mL of M9 to remove as much bacteria as possible. After the fourth wash and centrifugation, M9 was removed leaving ~ 1 mL of buffer and worm pellet. The contents were then transferred to a Low Retention 1.7 mL microtube (Genesee Scientific #: 24-282LR). Finally, worms were centrifuged at 2,000 x g for 1 minute and M9 buffer was carefully removed leaving ~25 µL and pellet untouched. The pellet was immediately flash-frozen in liquid nitrogen and stored in -80ºC until used for RNA extraction.

Total RNA preparation

Samples were first processed by adding 30 µL of 2X worm lysis buffer (80mM Tris pH7.5, 20mM EDTA, 400mM NaCl, 1% SDS, and 0.8g/ul Proteinase K) to the frozen pellet, and incubated at 55° for 10 minutes. Next, total RNA was extracted using TRI reagent and 1–bromo–3–chloropropane (BCP) (Molecular Research Center) followed by isopropanol precipitation, ethanol wash, and resuspended in nuclease-free water. Next, samples were treated with Turbo DNase (Ambion) to eliminate genomic DNA and purified with RNA clean & concentrator (Zymo). Total RNA quality was verified with Agilent BioAnalyzer.

Library construction for RNA sequencing

Sequencing libraries were built starting with 100 ng of total RNA using the NEBNext Ultra II Directional RNA library Prep Kit for Illumina with polyA mRNA workflow. mRNAs were enriched using the NEBNext Poly(A) mRNA Magnetic Isolation Module. 48 libraries were built in parallel in 96 well format and NEBNext Multiplex Oligos (E6440) was used for adaptor ligation. Libraries were ran a polyacrylamide gel and a size range of 200-700 bp were extracted to remove un-ligated adaptors. Finally, quality of libraries were verified with Aligent BioAnalyzer and concentrations were verified with Qubit. All libraries were pooled to equimolar amounts and sequenced twice on a NextSeq 500 using a High Output kit producing 150 bp paired-end reads.

RNA-seq data analysis

RNA sequencing results and allele-specific analysis were analyzed using a custom pipeline. Briefly, a custom transcriptome was generated for CB4856 by using liftover to WS230 and genome published in Thompson et al., 2015. Reference files were generated and gene expression was estimated using RSEM (V1.3.0) with bowtie (v1.2.2). Genes were considered expressed if 15 TPMs were detected at any time point. Graphs in **Figure 3-4** were generated in R. Significance for allelic imbalance shown in graphs were calculated by fitting a beta distribution to the observed allelic ratio [NN / (NN + HH)] for the parents and [N / (N+H)] for the hybrids, and the probability of each allele being in the beta distribution were calculated. p-values were corrected for multiple hypothesis testing using FDR with a cut-off of p<=0.05.

The heatmap in **Figure 3-2** was generated by calculating Pearson correlation of the sum of normalized read counts for each gene that contains a polymorphism between samples using GraphPad Prism 8.1.2.

Pyrosequencing

Total RNA was converted to cDNA using SuperScript IV (Invitrogen) with a mix of oligo dT and random hexamer primers. cDNA was diluted 1:10 to 1:30 before used for PCR amplification of candidate genes.

The steps and parameters for choosing candidates from RNA-seq results for pyrosequencing were as follows:

1) Calculate the $\log_2$ fold change of normalized expression level (transcripts per million, TPM) between strains NN and HH [($\log_2$ (NN/HH)], and between alleles in hybrids [$\log_2$ (N/H)]. Ratios after $\log_2$ transformation will become zero-centered, with 0 meaning equal levels of expression between alleles. Positive values indicate higher level of expression from the N2 allele and negative values indicate higher level of expression from the HW allele.

2) Correlation between samples were tested by multiplying the fold change value between the parents and the hybrids and between biological replicates, the resulting value must be positive to indicate positive correlation and reproducibility. Only genes that showed positive correlation were considered for pyrosequencing analysis.

3) Candidates were next tested for compatibility with pyrosequencing. Pyrosequencing can only detect single-nucleotide differences. PCR amplification of targets is first performed and an amplicon size within the range of 70 – 200 bp is preferred. PCR products must show only one band before further analysis.

Pyrosequencing primers were designed with PyroMark Assay Design software 2.0, pyrosequencing was performed on PyroMark Q24 sequencer and raw data processed with PyroMark Q24 software (Qiagen).

Acknowledgements

We thank members of the Rando and Mello labs for helpful discussions. We thank Pranitha Vangala and Manuel Garber for their assistance with RNA-seq data analysis. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). This work was supported by NIH R01 HD080224. CCC is a Merck Fellow of the Helen Hay Whitney Foundation.

Table 5. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in Starved L1.

N2-cis	HW-cis	Trans	Cis-trans
B0281.3	B0238.18	C05D12.4	C33D9.9
C14C6.5	C07G3.10	F44F1.4	T10B5.7
F07E5.5	C49G7.13	K02E7.6	Y40C5A.4
F13B6.3	F53C3.8	Y71H2B.4	bra-2
F36H5.14	K02E7.4	clec-52	clec-91
F53C3.3	K09H11.7	fbxa-36	cyp-31A1
F53C3.4	R03H10.6	fipr-23	exc-5
K05F6.10	W08E12.3	grl-15	glh-4
R12C12.7	Y45G12C.3	math-24	hsp-12.3
T28A11.2	Y69A2AR.8	sod-5	npp-20
W04A8.4	ZC239.14		rpn-6.1
Y110A2AL.4	ZC239.6		vha-11
Y17G9B.8	ZC247.1		
Y82E9BR.22	clec-209		
Y92C3B.4	clec-77		
ZC204.12	cnc-3		
ZK1240.3	col-104		
ZK1248.13	decr-1.1		
atln-2	fbxa-182		
btb-16	fbxa-58		
clec-72	fipr-22		
ctl-1	gst-32		
ctsa-2	lipl-3		
drh-2	math-38		
fbxa-60	mct-1		
gcy-15	mut-16		
gln-3	nhr-226		
gpx-1	pqn-97		
gst-27			
linc-11			
math-3			
math-41			
pals-18			
pals-22			
rhgf-2			
sdz-24			
tag-234			
trp-2			

Table 6. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in Fed L1.

N2-cis	HW-cis	Trans	Cis-trans
C14C6.5	B0281.5	C41H7.6	F44E7.2
C29F9.2	C49G7.13	F58G6.9	bra-2
F07E5.5	F45D11.1	K08A2.1	cyp-31A1
R12C12.7	F53C3.8	Y46D2A.1	exc-5
T10B5.7	K02E7.4	Y73B6A.3	nhr-122
T28A11.2	K08D8.4	cyp-35A5	rhgf-2
W04A8.4	K09H11.7	dao-2	rpn-6.1
Y110A2AL.4	R03H10.6	fipr-22	vha-11
Y54G2A.45	R08E5.4	gst-32	vps-35
Y56A3A.18	Y19D10A.4	pud-2.2	
ZC204.12	Y40C5A.4		
ZK1248.13	ZC239.6		
clec-72	ZC247.1		
cpr-8	ceh-43		
gln-3	clec-170		
gpx-1	clec-209		
gst-27	clec-77		
linc-11	fbxa-182		
math-3	fbxa-58		
math-41	fbxc-36		
mtm-6	gst-39		
pals-22	math-38		
sdz-24	mct-1		
	mut-16		
	nhr-115		

Table 7. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in L2.

N2-cis	HW-cis	Trans	Cis-trans
B0281.3	B0281.5	C14C6.8	C01B10.6
C14C6.5	C36C5.14	C41H7.4	F44E7.2
C23H5.8	C36C5.15	C41H7.5	T10B5.7
C29F9.2	C49G7.13	C41H7.6	Y40C5A.4
F54E2.1	F45D11.1	C49G7.12	bra-2
K11D12.7	F45D11.4	K06H6.1	cyp-31A1
T20D4.10	F53C3.8	K06H6.2	exc-5
W04A8.4	K02E7.4	K08A2.1	rpn-6.1
Y54G2A.45	K08D8.4	decr-1.1	uso-1
ZC204.12	K09H11.7	nhr-155	vha-11
ZK6.11	R03H10.6		vps-28
arrd-22	R08E5.4		
clec-72	Y37E11AL.6		
ctl-1	Y52B11A.3		
ctsa-2	Y82E9BL.9		
fat-5	ZC239.6		
gln-3	clec-209		
gpx-1	clec-77		
gst-27	fbxa-182		
linc-11	fbxa-58		
math-3	fbxc-36		
math-41	glh-4		
rhgf-2	gst-39		
sdz-24	lips-6		
	math-38		
	mct-1		
	mut-16		
	pud-2.2		

Table 8. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in L3.

N2-cis	HW-cis	Trans	Cis-trans
C14C6.5	B0281.5	C14C6.6	F44E7.2
C23H5.8	C14C6.3	C41H7.5	Y17G9B.8
C29F9.2	C32H11.8	C41H7.6	Y40C5A.4
E03H4.8	C36C5.14	F36F12.1	Y47D3A.21
F54E2.1	C36C5.15	K02E7.6	Y52B11A.3
K11D12.7	C49G7.13	K06H6.1	bra-2
T10B5.7	C49G7.7	K06H6.2	cyp-31A1
T20D4.10	F45D11.1	K08A2.1	exc-5
Y54G2A.45	F45D11.14	T28A11.2	glh-4
ZC204.12	F53C3.8	Y82E9BR.5	rpn-6.1
ZK1248.13	F57G9.6	ZK1025.3	vha-11
ZK6.11	K02E7.4	ZK488.6	
aqp-1	K08D8.4	cyp-35A5	
clec-72	K09H11.7	nhr-155	
ctl-1	R03H10.6	nstp-7	
ctsa-2	Y19D10A.4		
fat-5	Y46D2A.1		
gln-3	Y82E9BL.9		
gpx-1	ZC239.6		
gst-27	clec-209		
linc-11	clec-77		
math-3	fbxa-182		
math-41	fbxa-58		
rhgf-2	fbxc-36		
sdz-24	gst-39		
tag-234	mut-16		
ugt-28	pud-2.2		

Table 9. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in L4 Male.

N2-cis	N2-cis	HW-cis	Trans	Cis-trans
C14C6.5	nep-8	B0213.18	B0513.90	C04F12.12
C23H5.8	nspd-7	B0281.5	BE0003N10.6	F44E7.2
C27D6.3	prom-1	C04F12.6	C06E8.5	T23B3.5
C29F9.2	rhgf-2	C31B8.12	C14C6.3	Y40C5A.4
E01G4.5	scl-15	C36C5.14	C14C6.6	Y49F6B.15
F26A3.5	scl-8	C36C5.15	C14C6.8	Y59E9AL.3
F34D6.8	siah-1	C49G7.13	C16C8.8	Y59E9AR.1
F43C11.12	tag-234	C49G7.7	C41H7.6	bra-2
F59H6.15	ugt-28	F13B6.1	F40G9.15	col-126
K09F6.3		F28A10.5	F40G9.7	cyp-31A1
R03H10.4		F41D3.13	K06H6.1	glh-4
T05F1.5		F53C3.8	K06H6.2	pinn-1
T10B5.7		F55B11.5	K12H6.8	rpn-6.1
T10D4.15		F55F10.3	K12H6.9	vha-11
T26E3.6		F56D6.11	Y47D7A.15	
T28A11.2		F57G9.6	Y47H10A.5	
W08E3.4		K02F6.4	Y58A7A.5	
Y48G9A.6		K09H11.7	ZK488.6	
Y49F6B.8		R03H10.6	acp-6	
Y54F10BM.3		R08E5.4	clec-17	
Y54G2A.45		Y37E11AL.4	clec-170	
Y69A2AR.47		Y37E11AL.6	col-139	
Y75B7B.1		Y39A3A.4	col-88	
ZC204.12		Y46D2A.1	ctl-1	
ZK1248.13		Y82E9BL.9	cyp-13A5	
ZK783.6		ZC239.6	cyp-35A5	
clec-103		clec-110	cysl-2	
clec-104		clec-132	cysl-3	
clec-119		clec-209	grd-3	
clec-124		clec-77	grl-27	
clec-94		cyp-34A8	linc-41	
clp-6		fbxa-182	lips-6	
fat-5		fbxa-58	nhr-155	
gln-3		fbxc-50	nstp-7	
gpx-1		msp-71	tba-7	
gst-27		nhr-115	ugt-43	
linc-11		nspe-4	zmp-3	
msp-64		plg-1		
nep-13				

Table 10. List of genes that showed expression divergence between N2 and HW alleles in F1 hybrids in L4 Female.

N2-cis	HW-cis	Trans	Cis-trans
C14C6.5	B0281.5	B0238.18	C01B10.6
C29F9.2	C36C5.14	BE0003N10.6	F44E7.2
K11D12.7	C36C5.15	C14C6.3	Y17G7B.12
R12C12.7	C49G7.13	C14C6.6	Y17G9B.8
T10B5.7	C49G7.7	C14C6.8	bra-2
Y54G2A.45	F45D11.14	C34F11.8	chk-1
ZC204.12	F57G9.6	C41H7.4	cyp-31A1
ZK1248.13	K02E7.4	F14F9.4	glh-4
gln-3	K08D8.4	F39E9.1	gmps-1
gpx-1	K09H11.7	F58G6.9	rpn-6.1
gst-27	R03H10.6	K06H6.1	vha-11
prom-1	Y19D10A.4	K06H6.2	
tag-234	Y40C5A.4	R05A10.8	
	Y46D2A.1	Y47D7A.15	
	Y82E9BL.9	Y58A7A.5	
	ZC239.6	clec-170	
	clec-209	col-88	
	clec-77	ctl-1	
	cyp-34A8	cyp-35A5	
	fbxa-182	cysl-2	
	fbxa-58	cysl-3	
	mct-1	fbxc-36	
	nhr-115	grd-3	
	pud-2.2	grl-27	
		linc-8	
		math-38	
		nhr-155	
		nstp-7	
		pqn-32	
		pqn-97	
		srap-1	

Table 11. Unique divergent genes across all conditions.

N2-cis	N2-cis	HW-cis	HW-cis	trans	trans	cis-trans
math-3	F53C3.4	F41D3.13	fbxc-50	cyp-13A5	F40G9.15	exc-5
clec-72	prom-1	Y45G12C.3	fbxa-182	clec-52	BE0003N10.6	hsp-12.3
nspd-7	gst-27	R08E5.4	plg-1	ugt-43	K06H6.1	Y49F6B.15
C27D6.3	F34D6.8	clec-77	col-104	K12H6.9	C14C6.6	T23B3.5
F53C3.3	sdz-24	nhr-226	C07G3.10	C41H7.5	Y73B6A.3	col-126
clec-103	math-41	clec-110	F55F10.3	Y47H10A.5	C16C8.8	npp-20
btb-16	linc-11	F45D11.1	lipl-3	ZK488.6	nstp-7	pinn-1
Y82E9BR.22	Y49F6B.8	C31B8.12	W08E12.3	dao-2	clec-17	bra-2
drh-2	scl-8	Y37E11AL.6	cnc-3	grl-15	zmp-3	rpn-6.1
F59H6.15	tag-234	K09H11.7	Y19D10A.4	pqn-32	R05A10.8	Y47D3A.21
siah-1	ZK783.6	F28A10.5	K02E7.4	linc-41	Y71H2B.4	Y59E9AR.1
clec-119	F26A3.5	clec-209	F56D6.11	ZK1025.3	Y58A7A.5	vps-28
arrd-22	Y48G9A.6	nhr-115		F39E9.1	C41H7.6	F44E7.2
Y75B7B.1	nep-8	F53C3.8		Y47D7A.15	acp-6	Y17G7B.12
gln-3	ugt-28	nspe-4		K08A2.1		nhr-122
E03H4.8	clec-104	K08D8.4		K06H6.2		Y59E9AL.3
K11D12.7	F13B6.3	Y37E11AL.4		grl-27		C04F12.12
F43C11.12	mtm-6	clec-132		F36F12.1		cyp-31A1
T10D4.15	Y54F10BM.3	B0213.18		col-139		vps-35
aqp-1	ZK1248.13	C36C5.15		cyp-35A5		chk-1
T26E3.6	clp-6	C32H11.8		grd-3		clec-91
scl-15	E01G4.5	ZC239.6		fbxa-36		uso-1
T05F1.5	nep-13	cyp-34A8		nhr-155		vha-11
K05F6.10	cpr-8	fbxa-58		cysl-3		C01B10.6
ZK6.11	F36H5.14	F57G9.6		C41H7.4		C33D9.9
F07E5.5	fat-5	Y69A2AR.8		B0513.90		gmps-1
C29F9.2	ZC204.12	K02F6.4		col-88		
T20D4.10	clec-94	Y82E9BL.9		Y82E9BR.5		
R03H10.4	F54E2.1	F45D11.14		C05D12.4		
Y92C3B.4	gcy-15	R03H10.6		linc-8		
K09F6.3	atln-2	B0281.5		F58G6.9		
gpx-1	clec-124	ZC239.14		sod-5		
C14C6.5	trp-2	C49G7.13		srap-1		
fbxa-60	Y56A3A.18	F55B11.5		F44F1.4		
ctsa-2	msp-64	C36C5.14		C06E8.5		
B0281.3		msp-71		K12H6.8		
Y110A2AL.4		mct-1		C34F11.8		
pals-22		gst-39		math-24		
ZK1240.3		F13B6.1		K02E7.6		
R12C12.7		ZC247.1		F40G9.7		
Y54G2A.45		F45D11.4		C49G7.12		
W08E3.4		mut-16		tba-7		
W04A8.4		ceh-43		fipr-23		
pals-18		Y39A3A.4		C14C6.8		
C23H5.8		C49G7.7		cysl-2		
Y69A2AR.47		C04F12.6		F14F9.4		

Table 12. Divergent genes enriched in Gene Ontology Expression Analysis (GEA).

gene	term
cyp-35A5	iron ion binding GO:0005506
cyp-13A5	iron ion binding GO:0005506
cyp-34A8	iron ion binding GO:0005506
fat-5	iron ion binding GO:0005506
cyp-35A5	defense response GO:0006952
aqp-1	defense response GO:0006952
C14C6.5	defense response GO:0006952
nhr-115	defense response GO:0006952
gpx-1	defense response GO:0006952
clec-52	defense response GO:0006952
K08D8.4	defense response GO:0006952
fipr-23	defense response GO:0006952
ZK6.11	defense response GO:0006952
math-38	defense response GO:0006952
fbxa-60	defense response GO:0006952
fipr-22	defense response GO:0006952
grd-3	defense response GO:0006952
cyp-35A5	response to biotic stimulus GO:0009607
aqp-1	response to biotic stimulus GO:0009607
C14C6.5	response to biotic stimulus GO:0009607
nhr-115	response to biotic stimulus GO:0009607
gpx-1	response to biotic stimulus GO:0009607
clec-52	response to biotic stimulus GO:0009607
K08D8.4	response to biotic stimulus GO:0009607
fipr-23	response to biotic stimulus GO:0009607
ZK6.11	response to biotic stimulus GO:0009607
math-38	response to biotic stimulus GO:0009607
fbxa-60	response to biotic stimulus GO:0009607
fipr-22	response to biotic stimulus GO:0009607
grd-3	response to biotic stimulus GO:0009607
cyp-35A5	immune system process GO:0002376
aqp-1	immune system process GO:0002376
C14C6.5	immune system process GO:0002376
nhr-115	immune system process GO:0002376
gpx-1	immune system process GO:0002376
K08D8.4	immune system process GO:0002376
fipr-23	immune system process GO:0002376
ZK6.11	immune system process GO:0002376
fbxa-60	immune system process GO:0002376
fipr-22	immune system process GO:0002376
cyp-35A5	tetrapyrrole binding GO:0046906
cyp-13A5	tetrapyrrole binding GO:0046906
ctl-1	tetrapyrrole binding GO:0046906

cyp-34A8	tetrapyrrole binding GO:0046906
Y52B11A.3	tetrapyrrole binding GO:0046906
pgph-1	dephosphorylation GO:0016311
mtm-6	dephosphorylation GO:0016311
K09F6.3	dephosphorylation GO:0016311
Y54F10BM.3	dephosphorylation GO:0016311
Y39A3A.4	dephosphorylation GO:0016311
pgph-2	dephosphorylation GO:0016311

Table 13. Genes enriched in Tissue Enrichment Analysis (TEA).

gene	term
C14C6.8	intestine WBbt:0005772
Y47H10A.5	intestine WBbt:0005772
fbxa-60	intestine WBbt:0005772
Y110A2AL.4	intestine WBbt:0005772
F45D11.1	intestine WBbt:0005772
pals-22	intestine WBbt:0005772
ZK6.11	intestine WBbt:0005772
hsp-12.3	intestine WBbt:0005772
K05F6.10	intestine WBbt:0005772
Y82E9BR.5	intestine WBbt:0005772
Y54G2A.45	intestine WBbt:0005772
lido-10	intestine WBbt:0005772
clec-52	intestine WBbt:0005772
ZK1248.13	intestine WBbt:0005772
lido-9	intestine WBbt:0005772
col-104	intestine WBbt:0005772
math-38	intestine WBbt:0005772
gpx-1	intestine WBbt:0005772
lips-6	intestine WBbt:0005772
clec-209	intestine WBbt:0005772
arrd-22	intestine WBbt:0005772
lido-8	intestine WBbt:0005772
uso-1	intestine WBbt:0005772
E03H4.8	intestine WBbt:0005772
E01G4.5	intestine WBbt:0005772
cysl-2	intestine WBbt:0005772
ugt-43	intestine WBbt:0005772
dao-2	intestine WBbt:0005772
ctl-1	intestine WBbt:0005772
fbxa-58	intestine WBbt:0005772
T10B5.7	intestine WBbt:0005772
F36H5.14	intestine WBbt:0005772
Y47D3A.21	intestine WBbt:0005772
K02E7.6	intestine WBbt:0005772
mtm-6	intestine WBbt:0005772
cyp-35A5	intestine WBbt:0005772
math-3	intestine WBbt:0005772
K11D12.7	intestine WBbt:0005772
siah-1	intestine WBbt:0005772
C29F9.2	intestine WBbt:0005772
F54E2.1	intestine WBbt:0005772
sdz-24	intestine WBbt:0005772
linc-8	intestine WBbt:0005772
lido-11	intestine WBbt:0005772

ZC239.14	intestine WBbt:0005772
K08D8.4	intestine WBbt:0005772
btb-16	intestine WBbt:0005772
nhr-226	intestine WBbt:0005772
gln-3	intestine WBbt:0005772
F26A3.5	intestine WBbt:0005772
F58G6.9	intestine WBbt:0005772
C23H5.8	intestine WBbt:0005772
C05D12.4	intestine WBbt:0005772
nhr-115	intestine WBbt:0005772
aqp-1	intestine WBbt:0005772
clec-170	intestine WBbt:0005772
C14C6.5	intestine WBbt:0005772
F45D11.14	intestine WBbt:0005772
pals-18	intestine WBbt:0005772
pud-2.2	intestine WBbt:0005772
pqn-32	intestine WBbt:0005772
rpn-6.1	intestine WBbt:0005772
math-24	intestine WBbt:0005772
R03H10.6	intestine WBbt:0005772
B0281.3	intestine WBbt:0005772
clec-17	intestine WBbt:0005772
Y45G12C.3	intestine WBbt:0005772
nep-8	intestine WBbt:0005772
C49G7.12	intestine WBbt:0005772
drh-2	intestine WBbt:0005772
fbxa-182	intestine WBbt:0005772
tba-7	intestine WBbt:0005772
nspd-7	intestine WBbt:0005772
ZK1240.3	intestine WBbt:0005772
tag-234	intestine WBbt:0005772
C07G3.10	intestine WBbt:0005772
cyp-13A5	intestine WBbt:0005772
R12C12.7	intestine WBbt:0005772
fat-5	intestine WBbt:0005772
nhr-122	intestine WBbt:0005772
cnc-3	intestine WBbt:0005772
K10C2.1	intestine WBbt:0005772
cpr-8	intestine WBbt:0005772
ZC204.12	intestine WBbt:0005772
Y56A3A.18	intestine WBbt:0005772
C01B10.6	intestine WBbt:0005772
F39E9.1	intestine WBbt:0005772
F14F9.4	intestine WBbt:0005772
pgph-2	intestine WBbt:0005772
pgph-1	intestine WBbt:0005772
pinn-1	intestine WBbt:0005772

math-41	intestine WBbt:0005772
gst-39	intestine WBbt:0005772
gst-27	intestine WBbt:0005772
T23B3.5	intestine WBbt:0005772
vps-35	intestine WBbt:0005772
vha-11	intestine WBbt:0005772
clec-72	intestine WBbt:0005772
Y40C5A.4	intestine WBbt:0005772
C16C8.8	male WBbt:0007850
ZK783.6	male WBbt:0007850
fmil-1	male WBbt:0007850
F45D11.1	male WBbt:0007850
msp-64	male WBbt:0007850
scl-8	male WBbt:0007850
clec-52	male WBbt:0007850
K12H6.9	male WBbt:0007850
C04F12.12	male WBbt:0007850
scl-15	male WBbt:0007850
T10D4.15	male WBbt:0007850
Y48G9A.6	male WBbt:0007850
T26E3.6	male WBbt:0007850
E01G4.5	male WBbt:0007850
msp-71	male WBbt:0007850
Y75B7B.1	male WBbt:0007850
ugt-43	male WBbt:0007850
clec-119	male WBbt:0007850
K12H6.8	male WBbt:0007850
C27D6.3	male WBbt:0007850
F43C11.12	male WBbt:0007850
C29F9.2	male WBbt:0007850
clp-6	male WBbt:0007850
W08E3.4	male WBbt:0007850
Y54F10BM.3	male WBbt:0007850
ZC239.6	male WBbt:0007850
F26A3.5	male WBbt:0007850
plg-1	male WBbt:0007850
C04F12.6	male WBbt:0007850
C14C6.5	male WBbt:0007850
K09F6.3	male WBbt:0007850
clec-17	male WBbt:0007850
nep-8	male WBbt:0007850
C49G7.12	male WBbt:0007850
T05F1.5	male WBbt:0007850
nspd-7	male WBbt:0007850
C07G3.10	male WBbt:0007850
F40G9.15	male WBbt:0007850
F40G9.7	male WBbt:0007850

nspf-2	male WBbt:0007850
R03H10.4	male WBbt:0007850
ceh-43	male WBbt:0007850
cnc-3	male WBbt:0007850
nep-13	male WBbt:0007850
Y69A2AR.8	male WBbt:0007850
pgph-2	male WBbt:0007850
pgph-1	male WBbt:0007850
Y49F6B.8	male WBbt:0007850
clec-124	male WBbt:0007850
T23B3.5	male WBbt:0007850
F13B6.3	male WBbt:0007850
Y59E9AL.3	male WBbt:0007850

Chapter 4: Discussion

There are two fundamental sources that contribute to phenotypic variation: environmental variation and genetic variation. Here, I characterized a novel environmental interaction that affects gene expression during L1 starvation and explored the contributions of natural genetic variation towards gene expression divergence within a species using a variety of sequencing methods.

Density-dependence and *daf*-22 independent signaling

In Chapter 2, I identified that population density can alter gene expression during L1 developmental arrest, I found that this is regulated through a *daf-22* independent pathway. Density-dependent chemical communication pathways are crucial for community behaviors in many organisms, such as quorum sensing and biofilm formation in bacteria (Whiteley et al., 2018). Insects and animals can also regulate their behavior using chemical signals (Wyatt, 2009), but the genetic basis for chemical communication in animals is generally poorly understood. Our study provides a framework for investigation into a novel chemical signaling pathway in *C. elegans*. This study also raises two curious questions: are there other chemical communication pathways beyond *daf-22* and the presumably disparate L1 density signaling pathway? Under what conditions can we reveal

them? Density-dependent behaviors and physiology responses during other developmental arrest in other life stages, except dauer, are poorly understood. Future work needs to address whether density affects gene expression and physiological changes in stages beyond L1 and dauer.

The elusive L1 density signaling pathway

To date, the evidence for *daf-22* independent inter-nematode communication pathways have been scarce and the mechanism remain elusive (Artyukhin et al., 2013; Ludewig et al., 2017, 2019). The *Plips-15::GFP* density reporter and other potential reporter candidate genes present an exciting avenue for expanding the known collection of metabolites used during inter-nematode signaling that ultimately alters worm physiology, like starvation survival (Artyukhin et al., 2013). Our reporter-based system present opportunities for experimental work addressing two key questions: 1) identifying the genetic pathways required for ascaroside-independent signaling, 2) identifying and validating functional compounds and determining the chemical structures of these compounds.

Using an L1 density reporter, two approaches can be taken to investigate the genetic pathways required for L1 density signaling. First, a targeted genetic perturbation approach either by RNAi or CRISPR/Cas9 against the most differentially expressed genes, for example *lips-15, lips-16, nspe-1, nspe-2* and other potential targets mentioned in Chapter 2 (**Figure 2-1**). Second, an unbiased genome-wide screen in the *Plips-15::GFP* background, by whole

genome RNAi or mutagenesis screening (Fraser et al., 2000; Rual et al., 2004). These experiments will help elucidate the density signaling mechanisms that seem to contain two components: sensing density and responding to density signals. From our unpublished findings, reporter GFP is not expressed at birth and gradually becomes expressed with similar a time frame of 3 to 5 hours after releasing reporters from high density conditions (**Figure 2-7D**; data not shown). This result supports the idea that density is first sensed then followed by a gene regulation response as reflected in reporter GFP expression. Third, the density reporter can be used to screen for active compounds using activity-guided fractionation approaches and the chemical structures can be determined using analytical chemistry techniques such as HPLC-MS. These experiments will help elucidate the genetics and biochemistry of the elusive L1 density signaling pathway.

One challenge encountered while developing a screening assay in a multi-well format (96 well or 384 well) was that L1s are small and the GFP signal was difficult to detect in wells as opposed to on slides with high magnification (63x), even with the GFP signal coming from the largest cell in the body (excretory cell) (**Figure 2-4, Figure 2-5, Figure 2-7**). This approach was unsuccessful due to two reasons: 1) our inability to differentiate signal from background in images; 2) when attempting to enhance the signal for quantitative analysis, we were unable to take Z-stack images because small movements occur due to either incomplete immobilization and from subtle movements of the microscope stage. To overcome these challenges, perhaps a strategy similar to fluorescence activated cell sorting such as the large-

particle flow cytometer-based approach can be used to avoid the data collection issues with microscopy.

One caveat in using the *Plips-15::GFP* reporter system for screening is that one should not assume that changes in *Plips-15::GFP* signal is sufficient to reflect the entirety of the effects exerted by density signaling. Based on activity guided fractionation experiments, density signaling appears to be regulated by a multifactorial signal as combined fractions were necessary while single fractions were insufficient to drive the starvation survival phenotype (Artyukhin et al., 2013). Even though the response by *Plips-15::GFP* in the excretory cell could reflect the final integrated response from all signals, secondary screens with additional reporter strains are likely necessary for the complete understanding of the synergies within the density signaling pathway that produce the starvation survival phenotype.

One question that was frequently raised while performing this work was: can L1 density signaling act through a non-chemical signal, like oxygen? Oxygen is known to affect aggregation behaviors on plates (Gray et al., 2004; Rogers et al., 2006). Even though oxygen may play a role during L1 density signaling and survival, but it is unlikely based on the results from fractionation and add-back experiments (Artyukhin et al., 2013).

Paternal effects on offspring phenotype in *C. elegans*

As mentioned in the Introduction section, the original motivation for the work presented in Chapter 2 was to investigate paternal effects on phenotype. Our goal was to test whether paternal starvation would alter offspring phenotypes. Our approach was to compare offspring sired by males that were exposed to 6 days of starvation during the L1 stage vs. continuously well-fed (control) males. The discovery of density effects was made possible due to the difficulty in generating large number of cross-progeny. In a typical experiment, I must manually pick starved fathers to well-fed mothers to generate cross progeny. To generate the control groups, the strategy was to use animals that were directly plated on OP50 NGM plates. Due to the ease of generating control populations (no picking involved) compared to starved populations, we always had excess of control progeny relative to starved father progeny thus control progenies were always generated at much higher numbers and densities. This led to the initial paternal effects studies to be confounded by population density.

After identifying the effects of density on gene expression, I repeated the paternal effects studies with densities controlled at low density (1 worm/μL). I tested progeny from starved fathers vs. control fathers for gene expression by single L1 RNA-seq, L1 starvation survival, L1 heat shock survival, brood size, and aging (data not shown). However, I did not identify any significant differences between the two groups. This suggests previous studies that showed transgenerational effects in ancestral L1 starvation

(Jobson et al., 2015) or dauer (Webster et al., 2018) may be either a signal that can only be passed on maternally, an effect of maternal provisioning, or both. Another possibility may be differences in culture or experimental conditions, and closer examination between our studies and Jobson et al. showed that their starvation regime involved starving arrested L1s in glass tubes for 8 days (Jobson et al., 2015) while we performed starvation on NGM plates without food for 6 days. It will be worthwhile to perform paternal studies using culture conditions exactly as Jobson et al. and also paternal passage through dauer to provide a more significant change in both environmental and developmental conditions.

Despite my preliminary experiments that showed negative results to paternal exposure to L1 starvation, the potential for paternal transmission of environmental conditions in *C. elegans* has not yet been tested comprehensively. Future efforts should be directed toward testing paternal exposure to drugs or toxins. Previous studies in mice have shown paternal effects of nicotine exposure (Vallaster et al., 2017). *C. elegans* have been shown to respond to drugs of abuse, such as alcohol, nicotine, cocaine and methamphetamine (Engleman et al., 2016). It would be curious to test whether the effects of drugs can be transmitted for multiple generations and whether this can be passed on through the male germline.

The genetic basis of evolution of gene expression in *C. elegans*

In Chapter 3, I explored the genetic basis of how gene expression can vary within a species during post-embryonic development. I found that the overall gene expression pattern seems to resemble an interspecies cross. However, this is an extrapolation based on studies in a handful of systems. A more expanded strategy that includes comparison between other divergent strains beyond N2 and HW, and analysis of hybrids between these strains will provide further insight into within species expression divergence and evolution.

With current sequencing technologies, it is straightforward to detect changes in gene expression but resolving the underlying mechanisms require extensive molecular characterization and experimentation. Future studies need to address the following questions in order to draw conclusions for the molecular basis of gene regulatory variation, the relationship between expression variation and developmental dynamics, and evolutionary implications.

What are the molecular mechanisms driving gene expression divergence?

For *cis* regulatory divergence, mutations in *cis* regulatory regions, such as, enhancer or promoter sequences that alter transcription factor binding affinities, histone modifications or nucleosome positioning/occupancy are likely candidates that directly affect transcriptional outcomes. Further studies measuring chromatin accessibility, histone modifications and nucleosome positioning could uncover how allele-specific chromatin architecture affects gene expression in *cis* regulatory regions (Daugherty et al., 2017; Meers et al., 2019). Given the repertoire of post-transcriptional regulation mechanisms in *C. elegans*, another potential source for *cis*-regulatory divergence are variations in 3' UTR regions, a well-known region for post-transcriptional gene regulation (Jan et al., 2011; Kaymak et al., 2016; Merritt et al., 2008). Further computational analysis needs to systematically characterize whether genes that display *cis* regulatory divergence contain similar sequence differences in 3'UTR regions that could be used to infer differential binding for gene regulatory agents, such as miRNAs and RNA binding proteins.

For compensatory *cis-trans* interactions, computational analysis to identify mutations in any consensus motif sequences in *cis* regulatory regions could be used to reveal evolution of specific gene regulatory mechanisms or networks. There are two predictions: 1) no general consensus sequence motifs can be found, suggesting each particular *cis*-regulatory element has evolved to avoid mis-expression targeted by a particular *trans*-acting factor; 2)

general consensus sequence motifs can be found, suggesting co-evolution of *cis*-regulatory elements for multiple genes in the same gene regulatory network. Similar analysis can also be performed on genes that exhibit *trans* regulatory divergence, except the expectation is that there should not be any differences in the *cis*-regulatory regions and the *trans*-acting factor contributing to expression variation will have to be extrapolated from sequence motif analysis to identify any potential *trans*-acting factor variants.

In our current analysis, we are invisible to a type of compensating *cis*-*trans* interaction where both alleles in the hybrid are either over- or under-expressed compared to parents. This type of interaction would be represented as no allelic imbalance in our graphical analysis, but in fact a type of novel interaction in the hybrid. To identify this will require expanding the analysis to include expression level in addition to direction of allelic imbalance.

Regardless of computational analysis outcomes, analysis of sequence conservation in non-coding regions is a poor predictor of functional outcomes and results must be validated *in vivo* to draw any conclusions about interactions in gene regulatory networks. Future studies need to genetically convert sequences in *cis*-regulatory regions from one allele form to another (i.e. converting N2 sequence to HW sequence, or vice versa), and measure gene expression outcomes to identify true functional mutations.

How does gene expression variation affect a dynamic yet tightly controlled biological process such as development?

To date, there have been two genome-wide studies that investigated gene expression variation during development in *C. elegans*. One study used comparative genomics approach comparing gene expression between wild-type N2 and CB4856 strains from egg to young adult stages and characterized the types of genes and the amount of expression variation between the two strains across development (Capra et al., 2008). A more recent study mapped polymorphic regions of the genome that contributed to gene expression variation during development using an eQTL approach with a panel of 206 recombinant inbred advanced intercross lines. They identified thousands of *cis*- and *trans*-eQTLs during development with high temporal resolution (every 3 hours) from the L3 to young adult stages. This study uncovered >900 loci that showed a local effect on variation in gene expression (*cis*-eQTLs) and 773 genes were affected by variation in 10 major distal loci (*trans*-eQTLs hotspots) (Francesconi and Lehner, 2014). These works provided insight into developmental gene regulation with either stage specificity between species or variation with high variation over a short period of time (Capra et al., 2008; Francesconi and Lehner, 2014). Our analysis using parent and F1 hybrids made the discovery for compensatory *cis-trans* interactions possible, which were undetectable using previous approaches, and increased the temporal resolution to every developmental stage.

Discussed in the foundational work about the course of organismal development as a stabilized and optimal process, Waddington described that: “developmental reactions, as they occur in organisms submitted to natural section, are in general canalized. That is to say, they are adjusted so as to bring about one definite end-result regardless of minor variations in conditions during the course of the reaction.” (Waddington, 1942). Considering our results, there were several genes that showed compensatory *cis-trans* interactions, and this can be viewed as a form of mis-expression in the hybrid. Furthermore, extensive gene expression divergence was observed in hybridized offspring (**Figure 3-4**). During experimentation, hybrids were viable, and no obvious differences were noticed during development between parents and hybrids, suggesting canalization was acting to protect the developmental process. One hypothesis is that expression variation for developmentally important genes are minimal. Further analysis integrating temporal expression patterns of individual genes from our dataset promise new insights into pathways that can tolerate variation and potentially developmentally controlled pathways.

The study by Capra et al. (2013) comparing gene expression between N2 and HW strains during development showed that some genes are expressed one stage earlier or later in one background over the other. What will happen to the expression of genes with this type of pattern in the hybrid? Consider gene expression as a function of time in development, a simple model where a given gene is expressed during L2 stage in N2 parental strain and expressed during L3 in HW parental strain (**Figure 4-1A-B**). There are

many possible outcomes, but let us consider two models that illustrate the detection of variable genes (can tolerate variation) and controlled genes (cannot tolerate variation):

1) For variable genes, the hybrids recapitulate the pattern of expression as the parents in an allele-specific manner (**Figure 4-1C**). This is a case of mis-expression (expressed twice during development) of the given gene, but it is tolerated during normal development. This suggests co-evolution of both *cis* and *trans* regulatory elements for the given gene achieved a specialized sequence recognition and temporal expression pattern in a background-specific manner. In this case, the divergence class of this gene remains *cis*, but the direction switches from N2-*cis* in L2 to HW-*cis* in L3 (**Figure 4-1D**).
2) For controlled genes, the given gene switches divergence class in the hybrid (**Figure 4-1E**). In this case, the pattern switches from N2-*cis* in L2 to *trans* affecting both alleles. This result indicates a mechanism that contains two components: a *trans*-activator that activated expression from the N2 allele in L2, and a *trans*-silencer that silences both alleles in L3 and beyond (**Figure 4-1F**). This suggests a feedback mechanism of the gene regulatory pathway to inhibit mis-expression during development. This result resembles a mechanism similar to controlling for heterochrony, meaning changes in the timing or rate of events during development (Keyte and Smith, 2014; Klingenberg, 1998). Heterochronic mutants in *C.*

elegans exhibit precocious, retarded, or repetitive events during development (Abbott et al., 2005; Ambros, 1997; Ambros and Horvitz, 1984, 1987). In our model, this would be inhibiting the expression of the given gene after its been expressed once to prevent repetitive events.

Further characterization of variable and controlled genes will help better understand the genetic control of development, and pathways and networks that allow for innovation or protected and selected for during evolution. Directed approaches targeting chromatin regulators or transcription factors can be used to dissect the molecular basis of gene activating and silencing activities proposed (**Figure 4-1C-F**) .

One of the outstanding puzzles in the field is understanding the logic behind evolution of gene expression in the context of gene regulatory networks and development (Signor and Nuzhdin, 2018). Our experimental model provides a genetically tractable system that allows the molecular characterization of this relationship. Our current analysis is restricted to gene expression at the mRNA level, it is possible that a buffering mechanism can act to limit variability on the protein level. For example, if a gene exhibits mis-expression through a *cis* regulated mechanism by transcription, a *trans* regulatory mechanism on a protein level can act to reduce the translational output of the gene to stabilize overall variability (Signor and Nuzhdin, 2018).

To gain a more comprehensive understanding of evolution of gene expression, an interesting avenue to pursue would be to compare

transcriptome variability to proteome variability. Is gene expression more conserved or variable on the transcriptional or post-transcriptional level? Future efforts should aim to expand the analysis to connect the dots between chromatin structure, transcription, post-transcriptional regulation, translation and post-translational regulatory mechanisms to understand the complete landscape of the evolution of *cis* and *trans* gene regulation and ultimately variation in phenotypes.

Figure 4-1

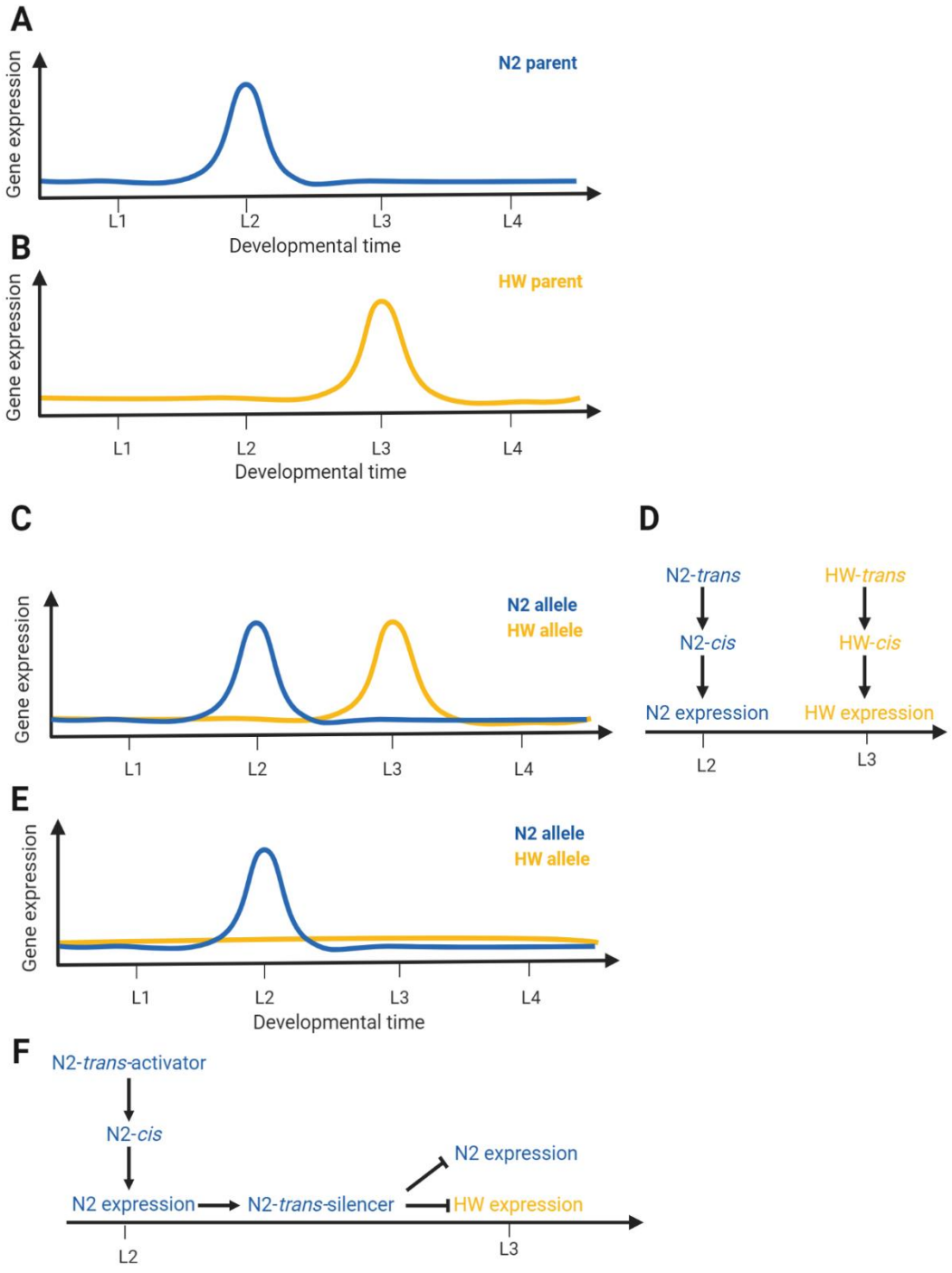


Figure 4-1. Models of *cis-trans* regulation of variable and controlled genes.

A) Expression of a given gene from the N2 parent.

B) Expression of a given from the HW parent.

C-D) Model and proposed mechanism for expression of a variable gene from the N2 and HW alleles in the hybrid.

E-F) Model and proposed mechanism for expression of a controlled gene from the N2 and HW alleles in the hybrid.

Are there any hybrid-specific effects?

Gene expression changes often lead to changes in phenotype. Our results revealed remarkable expression divergence and novel *cis-trans* interactions in the hybrid (**Figure 3-4**). It is therefore conceivable that hybrids may display unique phenotypes compared to their parents under certain environmental conditions. Interspecific hybrids generally produce offspring that produce lower fitness than their parents (i.e. reduced survival or reproductive capabilities), but sometimes the hybrid offspring show improved qualities compared to their parents and this is described as heterosis or hybrid vigor (Chen, 2013). Our system using intraspecific hybrids produced seemingly healthy offspring, so how well do hybrids fare compared to parents during stress? This has not been challenged rigorously.

Given that there are many distinct phenotypes between N2 and HW, however the gene or genes responsible for distinct or variable phenotypes cannot always be solved through classical genetic mapping strategies. For example, there is a significant difference between N2 and HW in cold tolerance and temperature acclimation (Okahata et al., 2016). In this study, Okahata and colleagues posed several temperature challenges to a panel of

thirteen *C. elegans* strains including N2 and HW. In their cold tolerance assay, worms were first reared at 15°C, 20°C or 25°C and then cold shocked at 2°C to measure survival. From the 15°C to 2°C shift N2 showed near 100% survival rates while HW showed ~50%. When shifted from 20°C or 25°C to 2°C, however, N2 survival rates were reduced to near 0% while HW survival rates were above 10%. Next, they performed a cold acclimation assay, where worms were first reared at 25°C, and the temperatures were switched from 25°C to 15°C for several hours (0, 3, 5, 8 hrs) and then followed by cold shock. They found that it takes N2 three hours to adapt and survive cold shock, raising from ~5% survival to ~25%; however there was no significant improvement in HW at 3 hours, from near 0% to 5%, but eventually acclimated after 8 hours and improving survival to ~50% (Okahata et al., 2016). This result indicate that N2 can better adapt to temperature changes if given enough time, but HW are more resistant to sudden temperature shifts but adapts much slower.

How would the hybrids perform? Is cold tolerance and acclimation dominant or recessive or do heterozygotes gain the benefits or fall out of the range of both parents? Performing cold tolerance and acclimation experiments in hybrids will help the understanding of adaptation and penetrance on a molecular level, and assaying for global gene expression alongside these changing conditions will provide mechanistic insight into the regulatory networks and interactions required for a temperature-dependent adaptive trait. Temperature fluctuations and climate change are arguably the most variable and steadily changing environmental conditions. How will

species adjust and adapt to this change? The genetic and phenotypic diversity in *C. elegans* present a versatile toolbox that can provide understanding into how multicellular organisms adapt to not only temperature changes but environmental changes in general. Future studies should combine eQTL and parent/hybrid approaches to link changes in gene expression as a first line of evidence to understand the genotype-phenotype relationship, and the evolution of gene regulatory mechanisms that contribute to phenotypic variation and potentially adaptation.

Another curious direction to take would be to assay for any hybrid-specific cell-type specific effects. Previous studies from Ruvinsky and colleagues examined how *cis*-regulatory elements between *Caenorhabditis* species contribute to variation in gene expression. They generated non-*elegans* promoter GFP reporter constructs and looked at cell-type specific expression and found ectopic expression outside the expected cell types when expressed in *C. elegans* (Barrière and Ruvinsky, 2014; Barrière et al., 2012). This result demonstrated that evolution of *cis* regulatory elements can alter not only expression level but also cell-type specificity. It would be interesting to uncover any unique cell-type specific expression patterns that can potentially be informative about any physiologically relevant phenotypes in hybrids within a species.

C. elegans as a genetic model for studying cytosine methylation independent parent-of-origin effects

Although the existence of genomic imprinting in *C. elegans* has been traditionally rejected in the field due to the absence of DNA cytosine methylation (Simpson et al., 1986), a well-characterized epigenetic modification involved in genomic imprinting in mammals and plants (Ferguson-Smith, 2011; Tucci et al., 2019), and genetic tests that did not validate any imprinting effects (Haack and Hodgkin, 1991). However, previous work from Fire lab showed the potential of imprinting in *C. elegans* by introducing GFP transgenes paternally or maternally and noticed differential GFP expression levels in progeny (Sha and Fire, 2005). More recent work from Strome lab discovered that specific histone modifications are inherited and propagated in a parent-of-origin specific manner, and some offspring phenotypes associated with sperm-inherited modified histones in *C. elegans* (Gaydos et al., 2014; Kaneshiro et al., 2019; Tabuchi et al., 2018). These results suggest that genes could be expressed in a parent-of-origin specific manner through a histone-based mechanism. Cases of histone-based genomic imprinting have been recently found in mice and human (Inoue et al., 2017, 2018; Xia et al., 2019; Xu et al., 2019; Zhang et al., 2019).

The experimental system presented in Chapter 3 also allows the analysis of parent-of-origin effects (**Figure 3-1**). From separate experimental work that stemmed from Chapter 3, I tested several candidates for parent-of-origin effects based on both bulk animal RNA-seq and single-worm RNA-seq

results; however, I was unable to validate those results using pyrosequencing. Generally, genes that exhibited parent-of-origin effects were lowly expressed by RNA-seq and results were inconsistent in pyrosequencing making it difficult to draw robust conclusions. One way to overcome this challenge would be to increase the number of biological and technical replicates, and substantially increase the sequencing depth to obtain more robust statistical power. However, our approach is invisible to cell-type specific events, as for the case for imprinting in mammals (Tucci et al., 2019), since we perform sequencing with whole worms. We also used a mixed population of males and females, therefore unable to separate any potential sex-specific effects. Genomic imprinting is generally considered to be a phenomenon in organisms of placental habit, with some cases in insects (Ferguson-Smith, 2011; Herrick and Seger, 1999). It would be a worthy endeavor to study this question at the single-cell, allele-specific level to address the potential for parent-of-origin effects in *C. elegans*, and potentially shed light on the evolution of cytosine methylation independent imprinting mechanisms.

Chapter 5: Bibliography

Abbott, A.L., Alvarez-Saavedra, E., Miska, E.A., Lau, N.C., Bartel, D.P., Horvitz, H.R., and Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. Dev. Cell.

Ambros, V. (1997). Heterochronic Genes.

Ambros, V., and Horvitz, H.R. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. Science (80-.).

Ambros, V., and Horvitz, H.R. (1987). The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. Genes Dev.

Ancel, L.W., and Fontana, W. (2000). Plasticity, evolvability, and modularity in RNA. J. Exp. Zool.

Andersen, E.C., Gerke, J.P., Shapiro, J.A., Crissman, J.R., Ghosh, R., Bloom, J.S., Félix, M.A., and Kruglyak, L. (2012). Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nat. Genet.

Anderson, L.M., Riffle, L., Wilson, R., Travlos, G.S., Lubomirski, M.S., and Alvord, W.G. (2006). Preconceptional fasting of fathers alters serum glucose in offspring of mice. Nutrition.

Angeles-Albores, D., Lee, R., Chan, J., and Sternberg, P. (2018). Two new functions in the WormBase Enrichment Suite. MicroPublication Biol.

Anway, M.D., Cupp, A.S., Uzumcu, N., and Skinner, M.K. (2005). Toxicology: Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science (80-.).

Aprison, E.Z., and Ruvinsky, I. (2016). Sexually Antagonistic Male Signals Manipulate Germline and Soma of C. elegans Hermaphrodites. Curr. Biol.

Arda, H.E., Taubert, S., MacNeil, L.T., Conine, C.C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K.R., and Walhout, A.J.M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol.

Artyukhin, A.B., Schroeder, F.C., and Avery, L. (2013). Density dependence in Caenorhabditis larval starvation. Sci. Rep. *3*, 2777.

Artyukhin, A.B., Zhang, Y.K., Akagi, A.E., Panda, O., Sternberg, P.W., and Schroeder, F.C. (2018). Metabolomic “dark Matter” Dependent on Peroxisomal β-Oxidation in Caenorhabditis elegans. J. Am. Chem. Soc.

Bale, T.L. (2015). Epigenetic and transgenerational reprogramming of brain development. Nat. Rev. Neurosci.

Bargmann, C.I. (2006). Chemosensation in C. elegans. WormBook.

Barrière, A., and Ruvinsky, I. (2014). Pervasive Divergence of Transcriptional Gene Regulation in Caenorhabditis Nematodes. PLoS Genet.

Barrière, A., Gordon, K.L., and Ruvinsky, I. (2012). Coevolution within and between Regulatory Loci Can Preserve Promoter Function Despite Evolutionary Rate Acceleration. PLoS Genet.

Baugh, L.R. (2013). To grow or not to grow: Nutritional control of development during Caenorhabditis elegans L1 Arrest. Genetics *194*, 539–555.

De Bono, M., and Bargmann, C.I. (1998). Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. Cell.

Bošković, A., and Rando, O.J. (2018). Transgenerational Epigenetic Inheritance. Annu. Rev. Genet. *52*, 21–41.

Brady, S.C., Zdraljevic, S., Bisaga, K.W., Tanny, R.E., Cook, D.E., Lee, D., Wang, Y., and Andersen, E.C. (2019). A Novel Gene Underlies Bleomycin-Response Variation in Caenorhabditis elegans . Genetics.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., et al. (2011). The evolution of gene expression levels in mammalian organs. Nature.

Brem, R.B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. Science (80-.).

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics.

Broman, K.W. (2005). The genomes of recombinant inbred lines. Genetics.

Buechner, M., Hall, D.H., Bhatt, H., and Hedgecock, E.M. (1999). Cystic canal mutants in Caenorhabditis elegans are defective in the apical membrane domain of the renal (excretory) cell. Dev. Biol.

Burga, A., Ben-David, E., Lemus Vergara, T., Boocock, J., and Kruglyak, L.

(2019). Fast genetic mapping of complex traits in C. elegans using millions of individuals in bulk. Nat. Commun.

Butcher, R.A., Ragains, J.R., and Clardy, J. (2009a). An indole-containing dauer pheromone component with unusual dauer inhibitory activity at higher concentrations. Org. Lett.

Butcher, R.A., Ragains, J.R., Li, W., Ruvkun, G., Clardy, J., and Mak, H.Y. (2009b). Biosynthesis of the Caenorhabditis elegans dauer pheromone . Proc. Natl. Acad. Sci.

Byerly, L., Cassada, R.C., and Russell, R.L. (1976). The life cycle of the nematode Caenorhabditis elegans. I. Wild-type growth and reproduction. Dev. Biol.

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science (80-.).

Capra, E.J., Skrovanek, S.M., and Kruglyak, L. (2008). Comparative developmental expression profiling of two C. elegans isolates. PLoS One.

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascenção, K., Rummel, C., Ovchinnikova, S., et al. (2019). Gene expression across mammalian organ development. Nature.

Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Hart, C.E., Li, R., Bock, C., Li, C., Gu, H., Zamore, P.D., et al. (2010). Paternally induced transgenerational environmental reprogramming of metabolic gene

expression in mammals. Cell *143*, 1084–1096.

Cassada, R.C., and Russell, R.L. (1975). The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. Dev. Biol.

Cheesman, H.K., Feinbaum, R.L., Thekkiniath, J., Dowen, R.H., Conery, A.L., and Pukkila-Worley, R. (2016). Aberrant activation of p38 MAP kinase-dependent innate immune responses is toxic to Caenorhabditis elegans. G3 Genes, Genomes, Genet.

Chen, Z.J. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet.

Chen, J., Nolte, V., and Schlötterer, C. (2015a). Temperature Stress Mediates Decanalization and Dominance of Gene Expression in Drosophila melanogaster. PLoS Genet.

Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., Feng, G.-H., Peng, H., Zhang, X., Zhang, Y., et al. (2015b). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. Science science.aad7977-.

Coolon, J.D., McManus, C.J., Stevenson, K.R., Graveley, B.R., and Wittkopp, P.J. (2014). Tempo and mode of regulatory evolution in Drosophila. Genome Res.

Corsi, A.K., Wightman, B., and Chalfie, M. (2015). A transparent window into biology: A primer on Caenorhabditis elegans. Genetics.

D'Andrea, L., Pérez-Rodríguez, F.-J., de Castellarnau, M., Guix, S., Ribes, E.,

Quer, J., Gregori, J., Bosch, A., and Pintó, R.M. (2019). The Critical Role of Codon Composition on the Translation Efficiency Robustness of the Hepatitis A Virus Capsid. Genome Biol. Evol.

Darwin, C. (1859). On the Origin of the Species.

Daugherty, A.C., Yeo, R.W., Buenrostro, J.D., Greenleaf, W.J., Kundaje, A., and Brunet, A. (2017). Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. Genome Res.

Davis, M.W., and Hammarlund, M. (2006). Single-nucleotide polymorphism mapping. Methods Mol. Biol.

Denver, D.R., Morris, K., Streelman, J.T., Kim, S.K., Lynch, M., and Thomas, W.K. (2005). The transcriptional consequences of mutation and natural selection in Caenorhabditis elegans. Nat. Genet.

Doroszuk, A., Snoek, L.B., Fradin, E., Riksen, J., and Kammenga, J. (2009). A genome-wide library of CB4856/N2 introgression lines of Caenorhabditis elegans. Nucleic Acids Res.

Ellis, R., and Schedl, T. (2007). Sex determination in the germ line. WormBook.

Emerson, J.J., Hsieh, L.C., Sung, H.M., Wang, T.Y., Huang, C.J., Lu, H.H.S., Lu, M.Y.J., Wu, S.H., and Li, W.H. (2010). Natural selection on cis and trans regulation in yeasts. Genome Res.

Emmons, S.W., Klass, M.R., and Hirsh, D. (1979). Analysis of the constancy of DNA sequences during development and evolution of the nematode

Caenorhabditis elegans. Proc. Natl. Acad. Sci.

Engleman, E.A., Katner, S.N., and Neal-Beliveau, B.S. (2016). Caenorhabditis elegans as a Model to Study the Molecular and Genetic Mechanisms of Drug Addiction. In Progress in Molecular Biology and Translational Science, p.

Fear, J.M., León-Novelo, L.G., Morse, A.M., Gerken, A.R., Van Lehmann, K., Tower, J., Nuzhdin, S. V., and McIntyre, L.M. (2016). Buffering of genetic regulatory networks in Drosophila melanogaster. Genetics.

Ferguson-Smith, A.C. (2011). Genomic imprinting: The emergence of an epigenetic paradigm. Nat. Rev. Genet.

Francesconi, M., and Lehner, B. (2014). The effects of genetic variation on gene expression dynamics during development. Nature.

Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature.

Frazier, H.N., and Roth, M.B. (2009). Adaptive Sugar Provisioning Controls Survival of C. elegans Embryos in Adverse Environments. Curr. Biol.

Frøkjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.-P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. *40*, 1375–1383.

Fullston, T., Teague, E.M.C.O., Palmer, N.O., Deblasio, M.J., Mitchell, M., Corbett, M., Print, C.G., Owens, J.A., and Lane, M. (2013). Paternal obesity

initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. FASEB J. *27*, 4226–4243.

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E., and Mansuy, I.M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nat. Neurosci. *17*, 667–669.

Gaydos, L.J., Wang, W., and Strome, S. (2014). H3K27me and PRC2 transmit a memory of repression across generations and during development. Science (80-.).

Gilad, Y., Rifkin, S.A., and Pritchard, J.K. (2008). Revealing the architecture of gene regulation: the promise of eQTL studies. Trends Genet.

Golden, J.W., and Riddle, D.L. (1982). A pheromone influences larval development in the nematode Caenorhabditis elegans. Science (80-.).

Golden, J.W., and Riddle, D.L. (1984). The Caenorhabditis elegans dauer larva: Developmental effects of pheromone, food, and temperature. Dev. Biol. *102*, 368–378.

Golden, J.W., and Riddle, D.L. (1985). A gene affecting production of the Caenorhabditis elegans dauer-inducing pheromone. MGG Mol. Gen. Genet.

Goncalves, A., Leigh-Brown, S., Thybert, D., Stefflova, K., Turro, E., Flicek, P., Brazma, A., Odom, D.T., and Marioni, J.C. (2012). Extensive compensatory cis-trans regulation in the evolution of mouse gene expression.

Genome Res.

Goodman, M.B. (2006). Mechanosensation. WormBook.

Gray, J.M., Karow, D.S., Lu, H., Chang, A.J., Chang, J.S., Ellis, R.E., Marietta, M.A., and Bargmann, C.I. (2004). Oxygen sensation and social feeding mediated by a C. elegans guanylate cyclase homologue. Nature.

Grishkevich, V., and Yanai, I. (2013). The genomic determinants of genotype × environment interactions in gene expression. Trends Genet.

Gumienny, T.L., and Savage-Dunn, C. (2013). TGF-β signaling in C. elegans. WormBook.

Haack, H., and Hodgkin, J. (1991). Tests for parental imprinting in the nematode Caenorhabditis elegans. MGG Mol. Gen. Genet.

Hansen, B.G., Halkier, B.A., and Kliebenstein, D.J. (2008). Identifying the molecular basis of QTLs: eQTLs add a new dimension. Trends Plant Sci.

Heard, E., and Martienssen, R.A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. Cell.

Herrick, G., and Seger, J. (1999). Imprinting and paternal genome elimination in insects. Results Probl. Cell Differ.

Hodgkin, J., and Doniach, T. (1997). Natural variation and copulatory plug formation in Caenorhabditis elegans. Genetics.

Hollick, J.B. (2016). Paramutation and related phenomena in diverse species. Nat. Rev. Genet.

Hsueh, Y.P., Mahanti, P., Schroeder, F.C., and Sternberg, P.W. (2013). Nematode-trapping fungi eavesdrop on nematode pheromones. Curr. Biol.

Hu, P.J. (2007). Dauer. WormBook 1–19.

Inoue, A., Jiang, L., Lu, F., Suzuki, T., and Zhang, Y. (2017). Maternal H3K27me3 controls DNA methylation-independent imprinting. Nature.

Inoue, A., Chen, Z., Yin, Q., and Zhang, Y. (2018). Maternal Eed knockout causes loss of H3K27me3 imprinting and random X inactivation in the extraembryonic cells. Genes Dev.

Jakobson, C.M., and Jarosz, D.F. (2019). Molecular Origins of Complex Heritability in Natural Genotype-to-Phenotype Relationships. Cell Syst.

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. Nature.

Jansen, R.C., and Nap, J.P. (2001). Genetical genomics: The added value from segregation. Trends Genet.

Jimenez-Chillaron, J.C., Isganaitis, E., Charalambous, M., Gesta, S., Pentinat-Pelegrin, T., Faucette, R.R., Otis, J.P., Chow, A., Diaz, R., Ferguson-Smith, A., et al. (2009). Intergenerational transmission of glucose intolerance and obesity by in utero undernutrition in mice. Diabetes.

Jobson, M.A., Jordan, J.M., Sandrof, M.A., Hibshman, J.D., Lennox, A.L., and Baugh, L.R. (2015). Transgenerational effects of early life starvation on growth, reproduction, and stress resistance in Caenorhabditis elegans. Genetics *201*, 201–212.

Kammenga, J.E., Doroszuk, A., Riksen, J.A.G., Hazendonk, E., Spiridon, L., Petrescu, A.J., Tijsterman, M., Plasterk, R.H.A., and Bakker, J. (2007). A Caenorhabditis elegans wild type defies the temperature-size rule owing to a single nucleotide polymorphism in tra-3. PLoS Genet.

Kaneshiro, K.R., Rechtsteiner, A., and Strome, S. (2019). Sperm-inherited H3K27me3 impacts offspring transcription and development in C. elegans. Nat. Commun.

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

Kelly, S.A., Panhuis, T.M., and Stoehr, A.M. (2012). Phenotypic plasticity: Molecular mechanisms and adaptive significance. Compr. Physiol.

Keyte, A.L., and Smith, K.K. (2014). Heterochrony and developmental timing mechanisms: Changing ontogenies in evolution. Semin. Cell Dev. Biol.

Kim, C., Kim, J., Kim, S., Cook, D.E., Evans, K.S., Andersen, E.C., and Lee, J. (2019). Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in C. elegans. Genome Res.

Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. Dev. Biol.

Kishimoto, S., Uno, M., Okabe, E., Nono, M., and Nishida, E. (2017).

Environmental stresses induce transgenerationally inheritable survival advantages via germline-to-soma communication in Caenorhabditis elegans. Nat. Commun.

Klass, M., and Hirsh, D. (1976). Non-ageing developmental variant of Caenorhabditis elegans. Nature.

Klingenberg, C.P. (1998). Heterochrony and allometry: The analysis of evolutionary change in ontogeny. Biol. Rev.

Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T., and Lehner, B. (2017). Transgenerational transmission of environmental information in C. elegans. Science (80-.).

L'Hernault, S.W. (2006). Spermatogenesis. WormBook.

Landry, C.R., Wittkopp, P.J., Taubes, C.H., Ranz, J.M., Clark, A.G., and Hartl, D.L. (2005). Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of drosophila. Genetics.

Lee, D., Yang, H., Kim, J., Brady, S., Zdraljevic, S., Zamanian, M., Kim, H., Paik, Y.K., Kruglyak, L., Andersen, E.C., et al. (2017). The genetic basis of natural variation in a phoretic behavior. Nat. Commun.

Lewis, J.A., and Fleming, J.T. (1995). Chapter 1: Basic Culture Methods. Methods Cell Biol.

Li, X.C., and Fay, J.C. (2017). Cis-regulatory divergence in gene expression between two thermally divergent yeast species. Genome Biol. Evol.

Ludewig, A.H., Gimond, C., Judkins, J.C., Thornton, S., Pulido, D.C., Micikas, R.J., Döring, F., Antebi, A., Braendle, C., and Schroeder, F.C. (2017). Larval crowding accelerates C. elegans development and reduces lifespan. PLoS Genet.

Ludewig, A.H., Artyukhin, A.B., Aprison, E.Z., Rodrigues, P.R., Pulido, D.C., Burkhardt, R.N., Panda, O., Zhang, Y.K., Gudibanda, P., Ruvinsky, I., et al. (2019). An excreted small molecule promotes C. elegans reproductive development and aging. Nat. Chem. Biol. *15*, 838–845.

Ludewig , A, H., and Schroeder, F, C. (2013). Ascaroside signalling in C.elegans (January 18, 2013).

Mack, K.L., Campbell, P., and Nachman, M.W. (2016). Gene regulation and speciation in house mice. Genome Res.

MacOsko, E.Z., Pokala, N., Feinberg, E.H., Chalasani, S.H., Butcher, R.A., Clardy, J., and Bargmann, C.I. (2009). A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans. Nature.

Manosalva, P., Manohar, M., Von Reuss, S.H., Chen, S., Koch, A., Kaplan, F., Choe, A., Micikas, R.J., Wang, X., Kogel, K.H., et al. (2015). Conserved nematode signalling molecules elicit plant defenses and pathogen resistance. Nat. Commun.

Martin, A., and Orgogozo, V. (2013). The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. Evolution (N. Y).

Maures, T.J., Booth, L.N., Benayoun, B.A., Izrayelit, Y., Schroeder, F.C., and

Brunet, A. (2014). Males shorten the life span of C. elegans hermaphrodites via secreted compounds. Science (80-.).

Maydan, J.S., Flibotte, S., Edgley, M.L., Lau, J., Selzer, R.R., Richmond, T.A., Pofahl, N.J., Thomas, J.H., and Moerman, D.G. (2007). Efficient high-resolution deletion discovery in Caenorhabditis elegans by array comparative genomic hybridization. Genome Res.

Maydan, J.S., Lorch, A., Edgley, M.L., Flibotte, S., and Moerman, D.G. (2010). Copy number variation in the genomes of twelve natural isolates of Caenorhabditis elegans. BMC Genomics.

McManus, C.J., Coolon, J.D., Duff, M.O., Eipper-Mains, J., Graveley, B.R., and Wittkopp, P.J. (2010). Regulatory divergence in Drosophila revealed by mRNA-seq. Genome Res.

Meers, M.P., Bryson, T.D., Henikoff, J.G., and Henikoff, S. (2019). Improved CUT&RUN chromatin profiling tools. Elife.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3′ UTRs Are the Primary Regulators of Gene Expression in the C. elegans Germline. Curr. Biol.

Metzger, B.P.H., Duveau, F., Yuan, D.C., Tryban, S., Yang, B., and Wittkopp, P.J. (2016). Contrasting Frequencies and Effects of cis- and trans-Regulatory Mutations Affecting Gene Expression. Mol. Biol. Evol.

Metzger, B.P.H., Wittkopp, P.J., and Coolon, J.D. (2017). Evolutionary Dynamics of Regulatory Changes Underlying Gene Expression Divergence

among Saccharomyces Species. Genome Biol. Evol.

Miersch, C., and Döring, F. (2012). Paternal dietary restriction affects progeny fat content in Caenorhabditis elegans. IUBMB Life.

Moore, R.S., Kaletsky, R., and Murphy, C.T. (2019). Piwi/PRG-1 Argonaute and TGF-β Mediate Transgenerational Learned Pathogenic Avoidance. Cell.

Murphy, C.T., and Hu, P.J. (2013). Insulin/insulin-like growth factor signaling in C. elegans. WormBook.

Nelson, F.K., and Riddle, D.L. (1984). Functional study of the Caenorhabditis elegans secretory-excretory system using laser microsurgery. J. Exp. Zool.

Ng, S.-F., Lin, R.C.Y., Laybutt, D.R., Barres, R., Owens, J. a, and Morris, M.J. (2010). Chronic high-fat diet in fathers programs β-cell dysfunction in female rat offspring. Nature *467*, 963–966.

Nica, A.C., and Dermitzakis, E.T. (2013). Expression quantitative trait loci: Present and future. Philos. Trans. R. Soc. B Biol. Sci.

Nicholas, W.L., Dougherty, E.C., and Hansen, E.L. (1959). AXENIC CULTIVATION OF CAENORHARDITIS BRIGGSAE (NEMATODA: RHABDITIDAE) WITH CHEMICALLY UNDEFINED SUPPLEMENTS; COMPARATIVE STUDIES WITH RELATED NEMATODES. Ann. N. Y. Acad. Sci.

Nuzhdin, S. V., Friesen, M.L., and McIntyre, L.M. (2012). Genotype-phenotype mapping in a post-GWAS world. Trends Genet.

Ohta, A., Ujisawa, T., Sonoda, S., and Kuhara, A. (2014). Light and pheromone-sensing neurons regulates cold habituation through insulin signalling in Caenorhabditis elegans. Nat. Commun.

Okahata, M., Ohta, A., Mizutani, H., Minakuchi, Y., Toyoda, A., and Kuhara, A. (2016). Natural variations of cold tolerance and temperature acclimation in Caenorhabditis elegans. J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.

Öst, A., Lempradl, A., Casas, E., Weigert, M., Tiko, T., Deniz, M., Pantano, L., Boenisch, U., Itskov, P.M., Stoeckius, M., et al. (2014). Paternal diet defines offspring chromatin state and intergenerational obesity. Cell.

Palopoli, M.F., Rockman, M. V., TinMaung, A., Ramsay, C., Curwen, S., Aduna, A., Laurita, J., and Kruglyak, L. (2008). Molecular basis of the copulatory plug polymorphism in Caenorhabditis elegans. Nature.

Perez, M.F., and Lehner, B. (2019). Intergenerational and transgenerational epigenetic inheritance in animals. Nat. Cell Biol.

Perez, M.F., Francesconi, M., Hidalgo-Carcedo, C., and Lehner, B. (2017). Maternal age generates phenotypic variation in Caenorhabditis elegans. Nature.

Pigliucci, M. (2001). Phenotypic plasticity: Beyond nature and nurture: Syntheses in ecology and evolution.

Pukkila-Worley, R., and Ausubel, F.M. (2012). Immune defense mechanisms in the Caenorhabditis elegans intestinal epithelium. Curr. Opin. Immunol.

Radford, E.J., Ito, M., Shi, H., Corish, J.A., Yamazawa, K., Isganaitis, E.,

Seisenberger, S., Hore, T.A., Reik, W., Erkek, S., et al. (2014). In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. Science *345*, 1255903.

Ramsköld, D., Luo, S., Wang, Y.C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat. Biotechnol.

Rando, O.J. (2012). Daddy issues: Paternal effects on phenotype. Cell.

Rando, O.J., and Simmons, R.A. (2015). I'm eating for two: Parental dietary effects on offspring metabolism. Cell.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of Genetic and Epigenetic Inheritance. Cell.

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res.

Remy, J.J. (2010). Stable inheritance of an acquired behavior in caenorhabditis elegans. Curr. Biol.

Von Reuss, S.H., Bose, N., Srinivasan, J., Yim, J.J., Judkins, J.C., Sternberg, P.W., and Schroeder, F.C. (2012). Comparative metabolomics reveals biogenesis of ascarosides, a modular library of small-molecule signals in C. elegans. J. Am. Chem. Soc.

Rhoné, B., Mariac, C., Couderc, M., Berthouly-Salazar, C., Ousseini, I.S., and

Vigouroux, Y. (2017). No excess of CIS-regulatory variation associated with intraspecific selection in wild Pearl Millet (Cenchrus Americanus). Genome Biol. Evol.

Rockman, M. V., and Kruglyak, L. (2006). Genetics of global gene expression. Nat. Rev. Genet.

Rockman, M. V., and Kruglyak, L. (2009). Recombinational landscape and population genomics of caenorhabditis elegans. PLoS Genet.

Rodgers, A.B., Morgan, C.P., Leu, N.A., and Bale, T.L. (2015). Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. Proc. Natl. Acad. Sci. U. S. A.

Rogers, C., Persson, A., Cheung, B., and de Bono, M. (2006). Behavioral Motifs and Neural Pathways Coordinating O2 Responses and Aggregation in C. elegans. Curr. Biol.

Rual, J.F., Ceron, J., Koreth, J., Hao, T., Nicot, A.S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S.H., Hill, D.E., van den Heuvel, S., et al. (2004). Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. Genome Res.

Sarropoulos, I., Marin, R., Cardoso-Moreira, M., and Kaessmann, H. (2019). Developmental dynamics of lncRNAs across mammalian organs and species. Nature.

Schaefke, B., Emerson, J.J., Wang, T.Y., Lu, M.Y.J., Hsieh, L.C., and Li, W.H. (2013). Inheritance of gene expression level and selective constraints on

trans-and cis-regulatory changes in yeast. Mol. Biol. Evol.

Schedl, T., and Kimble, J. (1988). fog-2, a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in Caenorhabditis elegans. Genetics *119*, 43–61.

Schindler, A.J., Baugh, L.R., and Sherwood, D.R. (2014). Identification of Late Larval Stage Developmental Checkpoints in Caenorhabditis elegans Regulated by Insulin/IGF and Steroid Hormone Signaling Pathways. PLoS Genet. *10*.

Schulenburg, H., and Müller, S. (2004). Natural variation in the response of Caenorhabditis elegans towards Bacillus thuringiensis. Parasitology.

Scott, E., Hudson, A., Feist, E., Calahorro, F., Dillon, J., De Freitas, R., Wand, M., Schoofs, L., O'Connor, V., and Holden-Dye, L. (2017). An oxytocin-dependent social interaction between larvae and adult C. Elegans. Sci. Rep.

Seidel, H.S., and Kimble, J. (2011). The oogenic germline starvation response in c. elegans. PLoS One *6*.

Seidel, H.S., Rockman, M. V., and Kruglyak, L. (2008). Widespread genetic incompatibility in C. elegans maintained by balancing selection. Science (80-.).

Seidel, H.S., Ailion, M., Li, J., van Oudenaarden, A., Rockman, M. V., and Kruglyak, L. (2011). A novel sperm-delivered toxin causes late-stage embryo lethality and transmission ratio distortion in C. elegans. PLoS Biol.

Sha, K., and Fire, A. (2005). Imprinting capacity of gamete lineages in

Caenorhabditis elegans. Genetics.

Sharma, U., Conine, C.C., Shea, J.M., Boskovic, A., Derr, A.G., Bing, X.Y., Belleannee, C., Kucukural, A., Serra, R.W., Sun, F., et al. (2015). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science (80-.). science.aad6780-.

Shelton, C.A. (2006). Quantitative PCR approach to SNP detection and linkage mapping in Caenorhabditis elegans. Biotechniques.

Shen, S.Q., Turro, E., and Corbo, J.C. (2014). Hybrid mice reveal parent-of-origin and Cis- and Trans-regulatory effects in the retina. PLoS One.

Shi, C., Runnels, A.M., and Murphy, C.T. (2017). Mating and male pheromone kill Caenorhabditis males through distinct mechanisms. Elife.

Shi, X., Ng, D.W.K., Zhang, C., Comai, L., Ye, W., and Jeffrey Chen, Z. (2012). Cis- and trans-regulatory divergence between progenitor species determines gene-expression novelty in Arabidopsis allopolyploids. Nat. Commun.

Signor, S.A., and Nuzhdin, S. V. (2018). The Evolution of Gene Expression in cis and trans. Trends Genet.

Simpson, V.J., Johnson, T.E., and Hammen, R.F. (1986). Caenorhabditis elegans DNA does not contain 5-methylcytosine at any time during development or aging. Nucleic Acids Res.

Smith, E.N., and Kruglyak, L. (2008). Gene-environment interaction in yeast gene expression. PLoS Biol.

Snoek, B.L., Sterken, M.G., Bevers, R.P.J., Volkers, R.J.M., van't Hof, A., Brenchley, R., Riksen, J.A.G., Cossins, A., and Kammenga, J.E. (2017). Contribution of trans regulatory eQTL to cryptic genetic variation in C. elegans. BMC Genomics.

Sonoda, S., Ohta, A., Maruo, A., Ujisawa, T., and Kuhara, A. (2016). Sperm Affects Head Sensory Neuron in Temperature Tolerance of Caenorhabditis elegans. Cell Rep.

Srinivasan, J., Kaplan, F., Ajredini, R., Zachariah, C., Alborn, H.T., Teal, P.E.A., Malik, R.U., Edison, A.S., Sternberg, P.W., and Schroeder, F.C. (2008). A blend of small molecules regulates both mating and development in Caenorhabditis elegans. Nature.

Srinivasan, J., von Reuss, S.H., Bose, N., Zaslaver, A., Mahanti, P., Ho, M.C., O'Doherty, O.G., Edison, A.S., Sternberg, P.W., and Schroeder, F.C. (2012). A modular library of small molecule signals regulates social behaviors in Caenorhabditis elegans. PLoS Biol.

Stern, D.B., and Crandall, K.A. (2018). The evolution of gene expression underlying vision loss in cave animals. Mol. Biol. Evol.

Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Dev Biol.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol.

Tabuchi, T.M., Rechtsteiner, A., Jeffers, T.E., Egelhofer, T.A., Murphy, C.T.,

and Strome, S. (2018). Caenorhabditis elegans sperm carry a histone-based epigenetic memory of both spermatogenesis and oogenesis. Nat. Commun.

Thompson, O.A., Snoek, L.B., Nijveen, H., Sterken, M.G., Volkers, R.J.M., Brenchley, R., van't Hof, A., Bevers, R.P.J., Cossins, A.R., Yanai, I., et al. (2015). Remarkably divergent regions punctuate the genome assembly of the Caenorhabditis elegans hawaiian strain CB4856. Genetics.

Tian, J., Keller, M.P., Broman, A.T., Kendziorski, C., Yandell, B.S., Attie, A.D., and Broman, K.W. (2016). The dissection of expression quantitative trait locus hotspots. Genetics.

Tijsterman, M., Okihara, K.L., Thijssen, K., and Plasterk, R.H.A. (2002). PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of C. elegans. Curr. Biol.

Tirosh, I., Reikhav, S., Levy, A.A., and Barkai, N. (2009). A yeast hybrid provides insight into the evolution of gene expression regulation. Science (80-.).

Trombetta, J.J., Gennert, D., Lu, D., Satija, R., Shalek, A.K., and Regev, A. (2014). Preparation of single-cell RNA-Seq libraries for next generation sequencing. Curr. Protoc. Mol. Biol.

Tucci, V., Isles, A.R., Kelsey, G., Ferguson-Smith, A.C., Bartolomei, M.S., Benvenisty, N., Bourc'his, D., Charalambous, M., Dulac, C., Feil, R., et al. (2019). Genomic Imprinting and Physiological Processes in Mammals. Cell.

Vallaster, M.P., Kukreja, S., Bing, X.Y., Ngolab, J., Zhao-Shea, R., Gardner,

P.D., Tapper, A.R., and Rando, O.J. (2017). Paternal nicotine exposure alters hepatic xenobiotic metabolism in offspring. Elife.

Vassoler, F.M., White, S.L., Schmidt, H.D., Sadri-Vakili, G., and Christopher Pierce, R. (2013). Epigenetic inheritance of a cocaine-resistance phenotype. Nat. Neurosci.

Vergara, I.A., Tarailo-Graovac, M., Frech, C., Wang, J., Qin, Z., Zhang, T., She, R., Chu, J.S.C., Wang, K., and Chen, N. (2014). Genome-wide variations in a natural isolate of the nematode Caenorhabditis elegans. BMC Genomics.

Verta, J.P., and Jones, F.C. (2019). Predominance of cis-regulatory changes in parallel expression divergence of sticklebacks. Elife.

Van Voorhies, W.A. (1996). Bergmann size clines: A simple explanation for their occurrence in ectotherms. Evolution (N. Y).

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. Nature.

Walker, A.K., Jacobs, R.L., Watts, J.L., Rottiers, V., Jiang, K., Finnegan, D.M., Shioda, T., Hansen, M., Yang, F., Niebergall, L.J., et al. (2011). A conserved SREBP-1/phosphatidylcholine feedback circuit regulates lipogenesis in metazoans. Cell.

Wang, M., Uebbing, S., and Ellegren, H. (2017). Bayesian inference of allele-specific gene expression indicates abundant Cis-regulatory variation in natural flycatcher populations. Genome Biol. Evol.

Wang, X., Werren, J.H., and Clark, A.G. (2016). Allele-Specific Transcriptome and Methylome Analysis Reveals Stable Inheritance and Cis-Regulation of DNA Methylation in Nasonia. PLoS Biol.

Webster, A.K., Jordan, J.M., Hibshman, J.D., Chitrakar, R., and Ryan Baugh, L. (2018). Transgenerational effects of extended dauer diapause on starvation survival and gene expression plasticity in Caenorhabditis elegans. Genetics.

Wei, Y.-P.P.Y., Yang, C.-R.R., Zhao, Z.-A. a, Hou, Y., Schatten, H., and Sun, Q.-Y.Y. (2014). Paternally induced transgenerational inheritance of susceptibility to diabetes in mammals. Proc Natl Acad Sci U S A *111*, 1873–1878.

Whiteley, M., Diggle, S.P., and Greenberg, E.P. (2018). Bacterial quorum sensing: the progress and promise of an emerging research area. Nature.

Wicks, S.R., Yeh, R.T., Gish, W.R., Waterston, R.H., and Plasterk, R.H.A. (2001). Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat. Genet.

Wittkopp, P.J. (2013). V.7. Evolution of Gene Expression. In The Princeton Guide to Evolution, p.

Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2004). Evolutionary changes in cis and trans gene regulation. Nature.

Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2008). Regulatory changes underlying expression differences within and between Drosophila species. Nat. Genet.

Wyatt, T.D. (2009). Fifty years of pheromones. Nature.

Xia, W., Xu, J., Yu, G., Yao, G., Xu, K., Ma, X., Zhang, N., Liu, B., Li, T., Lin, Z., et al. (2019). Resetting histone modifications during human parental-to-zygotic transition. Science (80-.).

Xu, Q., Xiang, Y., Wang, Q., Wang, L., Brind'Amour, J., Bogutz, A.B., Zhang, Y., Zhang, B., Yu, G., Xia, W., et al. (2019). SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. Nat. Genet.

Yamada, K., Hirotsu, T., Matsuki, M., Butcher, R.A., Tomioka, M., Ishihara, T., Clardy, J., Kunitomo, H., and Iino, Y. (2010). Olfactory plasticity is regulated by pheromonal signaling in Caenorhabditis elegans. Science (80-.).

Yang, B., and Wittkopp, P.J. (2017). Structure of the transcriptional regulatory network correlates with regulatory divergence in Drosophila. Mol. Biol. Evol.

Zeybel, M., Hardy, T., Wong, Y.K., Mathers, J.C., Fox, C.R., Gackowska, A., Oakley, F., Burt, A.D., Wilson, C.L., Anstee, Q.M., et al. (2012). Multigenerational epigenetic adaptation of the hepatic wound-healing response. Nat. Med.

Zhang, W., Chen, Z., Yin, Q., Zhang, D., Racowsky, C., and Zhang, Y. (2019). Maternal-biased H3K27me3 correlates with paternal-specific gene expression in the human morula. Genes Dev.

Zhang, Y., Lu, H., and Bargmann, C.I. (2005). Pathogenic bacteria induce aversive olfactory learning in Caenorhabditis elegans. Nature.

Zhao, L., Zhang, X., Wei, Y., Zhou, J., Zhang, W., Qin, P., Chinta, S., Kong,

X., Liu, Y., Yu, H., et al. (2016). Ascarosides coordinate the dispersal of a plant-parasitic nematode with the metamorphosis of its vector beetle. Nat. Commun.