A MULTIPARTITE APPROACH TO MAPPING THE GENE NETWORK DIRECTING *CAENORHABDITIS ELEGANS* VULVAL ORGANOGENESIS

Thesis by

Ted Olin Ririe

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 22, 2009)

Acknowledgments

There are many people who have helped shape me as a person, provided me with opportunities, and acted as role models. I know that any of my accomplishments are due in large part to the selfless efforts and encouragement of others. My advisor, Paul Sternberg, personifies all these characteristics. Paul cares about science, he cares about people, and he shows it. This can be summed up by some advice he gave me recently in regards to emailing another lab about some overlapping research interests, "try to write it in a way that enhances science and makes everyone happier." I do not think I have ever met anyone else with such unbridled excitement and passion about their career. When I met obstacles during my graduate work he was unfailingly supportive and encouraging (I never would have come up with the "Ririegram" without you, Paul). He has been a fabulous mentor and I am grateful he allowed me to be a member of his lab.

I am thankful as well for the members of my thesis committee: Scott Fraser, Angela Stathopoulos, and Kai Zinn. They all made me think harder about my project, and they were always generous with their time and assistance. My science improved because of their questions and insights.

One of the great things about Caltech is the community of students and scientists who call it home, in particular all of the members of the Sternberg Lab with whom I have crossed paths. I cannot count all the times when my lab-mates took the time to teach me a protocol, gave me reagents, or just offered me a listening ear. I hope I returned the favor, and I have tried to follow their examples in working with others.

I would be remiss if I did not mention at least a few of them by name. When I joined the lab I was assigned to a room that had not been remodeled in several decades. This room's occupants (Jolene Fernandes, John "DeMo" DeModena, and Shahla Gharib) more than compensated, however, for what the room lacked in physical qualities. Who would have guessed that Jolene and I would become nearly inseparable friends? We didn't exchange much more than a hello or two during our first few weeks in the room, but we certainly made up for that in the years that followed. I will always remember our Capitol Steps addiction, the epic Mallomar quest, lunches/snacks/dinners at any one of the CBs, and those late nights of lab karaoke. Jolene was always ready to talk science, her advice was consistently thoughtful and intelligent, and she's my living, breathing *C. elegans* encyclopedia. DeMo's sleeping "schedule" is a thing of lab legend, but who else could I call when I saw some unidentified liquid oozing out from underneath a door at 3 a.m.? He was always the one I knew could (or at least would try to) fix a broken piece of equipment, find some missing reagent, or figure out a complicated protocol. I greatly appreciate him and his genuine kindness.

Takao helped me transition into the lab, and get my project up and running. Adler Dillman and Paul Minor were always good for a laugh, and they kept it from ever getting too lonely around the lab. Many thanks to Barb, Mary, and Gladys for making the lab run so smoothly and for adding some much needed continuity. Steven, Cheryl, Jagan, Mihoko, Adeline, Gary, and Bhagwati, to name a few past and present members of the lab, made this a truly unique and memorable place.

Three scientists were the most instrumental in preparing me for graduate work, Larry St. Clair, Adam Karpf, and David Jones. Dr. St. Clair provided me with a science

project while I was still in high school, and his generosity continued as I worked in his lab throughout my undergraduate years. His kind influence helped propel me into a PhD program. Dr. Jones gave me the opportunity to work as a lab technician right out of college, and he made the inspired decision of assigning Dr. Karpf as my supervisor. The skills and knowledge I learned from them were invaluable during my time at Caltech.

I will be eternally grateful for my parents, Delbert and Rebecca Ririe. Growing up we had a microscope for examining water samples from a local pond. Since before I could walk, Dad carried me around his garden while he taught me all about his favorite plants. Family vacations were filled with excursions to museums and natural wonders. I love them for providing me with a home where the pursuit of knowledge was encouraged and an interest in the world was nurtured. Their influence fostered within me a desire to pursue a career in science. Thank you as well to my sister Anneli and her family, and my brothers Dan and Peter and their families, for setting examples of goodness and achievement.

The most wonderful consequence of my graduate work is that while living in Pasadena I met, wooed, was dumped by, reconnected with, re-wooed, got engaged to, married, and procreated with the incomparable Allison Syphus. My wife is an amazingly talented woman (she teaches, sings, bakes, dances…I could go on), and she has an unparalleled ability to balance an ever-increasing roster of competing responsibilities. She makes me want to be a better person, and her love and patience while I have completed my thesis has been unmatched. Lastly, I thank Tessa. She only came into our lives a few short months ago, but she lights up our home and the love we feel for her is impossible to describe.

Abstract

The complex interactions of signaling molecules, transcription factors, and target genes that direct the development of diverse organisms and their component parts are known as gene regulatory networks (GRNs). The *Caenorhabditis elegans* vulva, with its invariant lineage, 22 cells of seven types (vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF), and availability of 30 reporter constructs exhibiting a variety of differentiated vulval cell expression patterns, provides us with an elegant system for studying a GRN directing organogenesis. Several regulators of vulval differentiation are known (COG-1, EGL-38, LIN-11, and LIN-29). Utilizing the available reporter constructs of genes expressed in the maturing vulva allowed me to further define the roles of LIN-11 (LIM homeodomain) and LIN-29 (zinc finger). The heterochronic gene pathway member LIN-29 was originally thought to play a direct role in the temporal expression of genes in the vulva. The analysis of multiple pairwise interactions between *lin-29* (loss of function) and members of the vulval GRN, however, has shown that LIN-29's role is more nuanced than previously assumed. I determined that *lin-29* positively regulates *lin-11* expression and directs a significant aspect of gene expression patterning in the maturing vulva.

The known regulators of the vulval network fail to account for all of the regulatory interactions required for proper vulval differentiation. As such, there are likely other transcriptional regulators that have not yet been identified. I took several approaches for identifying additional factors. First, I conducted a screen in a *zmp-1::GFP* (zinc metalloproteinase) background by disrupting activity of 836 transcription factors by RNA interference (RNAi). *zmp-1* is a readout for differentiation of the vulA, vulD, and vulE cells. This screen identified *nhr-67* (ortholog of *Drosophila tailless*) and

nhr-113 (orphan nuclear hormone receptor) as positive regulators of *zmp-1* expression in the vulA cells. *nhr-113* appears to have a narrow role in vulval organogenesis; perhaps only as a partial regulator of vulA cell differentiation. Furthering the analysis of *zmp-1*, I utilized a phylogenetic footprinting approach for analyzing the upstream region of *C. elegans zmp-1*, and its homologs in the related species *C. briggsae* and *C. remanei*. This analysis identified four conserved motifs upstream of the ZMP-1 translational start site. Characterization of these motifs, by deleting them from a *zmp-1::GFP* transgene, revealed that three of these four motifs exhibit *cis*-regulatory activity.

In addition to investigation of *zmp-1* regulation, we undertook a parallel study where we performed sequence alignment analysis on 17 genes with reported vulval expression in combination with *in vivo* testing. This identified nine ~ 200-bp vulva-specific enhancer elements associated with six of these genes. Yeast one-hybrid analysis of six of these enhancer elements isolated six protein-DNA interactions. Further characterization of these interactions uncovered the role of the zinc finger transcription factor, ZTF-16, in vulval induction and differentiation. These results show that by taking multiple approaches, including the use of post-genome technologies, we can expand our understanding of a gene regulatory network.

TABLE OF CONTENTS

CHAPTER I:

Gene Regulatory Networks and Development

Introduction

All metazoans arise initially from single cells that subsequently pass through multiple rounds of cell division. When a cell divides, its progeny each inherit the same genetic code, however, these cellular offspring can, and often do, express subsets of genes that are divergent from their parent and from each other. Each cell can assume incredibly diverse characteristics of size, shape, and function in spite of the fact that they each inherit the same genetic information. Gene regulatory networks (GRNs) describe the complex interactions that account for these differences (1).

There are, at a minimum, three components of a gene regulatory network. First, it consists of a system for processing information from either intra- or intercellular signaling molecules. Second, these signaling molecules modulate the expression levels of transcription factors (TFs). Third, target genes are acted upon by these transcription factors via *cis*-regulatory modules, and their expression products, either RNA or protein, ultimately account for the functional and structural differences between cells (1–4).

Each protein-encoding gene involved in a gene regulatory network is generally associated with a core promoter element, enhancers (activating sequences), and silencers (repressing sequences). Other elements may include insulators and barriers. 5' of the transcriptional start site is the core promoter element, an approximately 100 base pair (bp) region where the transcriptional machinery binds and transcription is initiated. Enhancers are bound by activators which act to increase transcription; while silencers are bound by repressors which decrease transcription (5). The enhancers, silencers, and insulators are collectively referred to as *cis*-regulatory modules, and are often located

upstream of the gene-coding sequence, but they have also been identified within intronic sequences and even downstream of their concomitant gene. These modules serve as the targets of the transcription factors that operate in a given gene regulatory network, and it is through them that the signals that direct the complex patterns of cellular gene expression are mediated (3).

The details of exactly when, where, and at what levels each gene is expressed are represented by the gene regulatory network. The linkages between genes and their targets are the building blocks of the network and reveal how the inherited genomic sequence can result in a given function and structure. These linkages cannot be fully appreciated at the level of individual gene-gene interactions, but by stepping back, and observing them in the context of multiple interacting subcircuits controlling a complex system (cell fate specification, embryogenesis, organogenesis, etc.), their significance can be more fully recognized (2, 3).

The process of development is extremely complex. To aid in understanding their underlying logic, developmental gene regulatory networks should be analyzed at the level of their component modules or subcircuits. Subcircuits consist of a group of genes and *cis*-regulatory elements that perform a specific developmental "job." There are a host of common subcircuits including auto and cross regulatory feedback, double-negative gates, intra- and intergenic feedback loops, regulatory autorepression, and transcriptional exclusion of alternative states (1–3, 6, 7).

1. How and why do we investigate GRNs?

BioTapestry (www.biotapestry.org) is an open source software package that aids in the computational modeling of gene regulatory networks (8). This software is useful in that it provides a uniform method for representing the architecture of developmental GRNs, and it is now utilized by many of the groups working in the field. Modeling in this manner also aids in the cross-species comparison of networks—comparisons that can illuminate evolutionary diversification that has occurred among related processes of differentiation and patterning (2, 9, 10). This research also reveals the blueprints for how genetic pathways generate different cell types, a central topic of study in the burgeoning field of stem cell biology (11, 12).

In studying developmental regulatory networks, we uncover the processes that account for the extraordinary diversity of cells, tissues, organs, and species that are present on Earth. By thoroughly dissecting these processes we can more fully understand how they operate. A potential benefit of this knowledge is that many of the diseases that afflict humans, animals, and plant species involve developmental processes at their core. These conditions can potentially be treated or reversed by repairing, disrupting, or circumventing defective genetic architecture.

1.1 New Approaches

Our understanding of gene regulatory networks is based on the results of many different techniques including genetic screens; genetic interaction analysis using mutant alleles, overexpression constructs and other tools; reporter assays; and *in vitro* binding

assays. In addition to these methods, analysis of *cis*-regulatory modules has typically involved labor-intensive molecular biology techniques including sequential deletion and dissection of regulatory elements, DNase footprinting assays, and gel-shift analysis. Advances in whole-genome techniques in the post-genomic sequencing era are opening up new avenues of exploration into gene regulatory networks.

1.1.1 A Genome-Wide Analysis of Kidney Organogenesis

An international consortium, the GenitoUrinary Development Molecular Anatomy Project (GUDMAP), is using the mouse kidney model for the first genomics-level analysis of organogenesis (13). They utilize either fluorescence-activated cell sorting (FACS) or laser capture microdissection (LCM) in combination with component-specific GFP transgenic mice to isolate and purify individual elements of the developing kidney. These techniques allow the consortium to isolate discrete components of the kidney at multiple developmental time points. RNA is then isolated from these samples and the transcription profiles are analyzed using microarray technology.

The data from this microarray analysis can then be used to identify linkages in the gene regulatory network of kidney development. When genes are expressed in the same region and at the same time point it is more likely that they are regulated by the same transcription factors. By determining the distribution of conserved transcription factor-binding sites in coexpressed genes, a regulatory circuit can be revealed. In the consortium's analysis of the promoters of genes showing enriched expression in the proximal tubule of the nephron, a statistically significant overabundance of HNF1β-

binding sites were identified. HNF1β is a homeodomain transcription factor that binds DNA either as a homo- or heterodimer. 44 genes were identified as candidate target genes of HNF1β due to the enriched presence of HNF1β-binding sites. One interesting putative target gene is the transcription factor *HNF4α*. Subsequent analysis revealed a further 19 genes in the proximal tubule that are putative targets of HNF4α. These findings suggest a possible regulatory circuit within the proximal tubule, with HNF1β regulating HNF4α, and HNF4α regulating at least a subset of these 19 implicated genes. These findings show the efficacy of utilizing this type to data to identify additional components of a gene regulatory network (13).

1.1.2 Genomic Techniques and Patterning of the *Drosophila* Embryo

The sequence-specific transcription factor Dorsal, a relative of mammalian NF-κB, controls the asymmetric dorsal-ventral (DV) patterning of the *Drosophila* embryo. Extensive analysis of this process has identified numerous Dorsal target genes and allowed for the elucidation of the regulatory framework that directs DV patterning. The recent implementation of several post-genomic technologies has resulted in the identification of many more potential Dorsal target genes and members of the dorsal-ventral patterning gene regulatory network (14).

A whole-genome tiling array that contains the entire *Drosophila* genome is now available. Because tiling arrays include the entire genome, unlike traditional microarrays, they are not biased by gene prediction models, and they include nonprotein-coding genes such as miRNAs. Using this technique, Biemer et al. were able to identify 30 additional

protein encoding genes and at least two miRNA genes that are regulated by the DV gradient (15).

Two of the earliest targets of Dorsal in the DV patterning network are the genes that encode Twist (basic helix-loop-helix activator) and Snail (zinc-finger repressor). To identify the enhancer elements through which these transcription factors regulate gene expression, Zeitlinger et al. used chromatin immunoprecipitation coupled with microarray analysis (ChIP-chip). This analysis identified several hundred additional potential target enhancers, many of which are associated with previously unidentified target genes. Confirming the efficacy of this approach, their ChIP-chip analysis identified 20 of 22 known target enhancers of Dorsal (16). The results of the tiling array and the ChIP-chip analysis reveal that the entire complement of genes that comprise even the most thoroughly studied gene regulatory networks have not yet been identified. These techniques, and others at the forefront of genetic research, have now proven their potential for greatly enhancing our understanding of genetic regulatory circuitry.

2. The *C. elegans* Vulva

In addition to the mouse kidney and *Drosophila* embryo, regulatory networks are being investigated in diverse model systems and at various developmental time points (see Table 1). These range from the network controlling environmentally induced physiological responses in the free-living, single-celled organism *Halobacterium salinarum* (17) to development of the hindbrain in mouse and chicken (18). Whereas the majority of gene regulatory networks being analyzed in other invertebrates focus

primarily on pathways involved in embryogenesis (i.e., germ layer specification in sea urchin (19), and specification and patterning of the *Drosophila* syncytial blastoderm (14, 20)), the *C. elegans* vulva provides us with a model that is well-suited for mapping the architecture of a gene regulatory network controlling organogenesis.

2.1 Vulval Development

The life cycle of *C. elegans* consists of four larval stages (L1–4) and an adult stage, with each stage separated by a molt. The vulva, which allows for passage of sperm and eggs by connecting the uterus to the external environment, is derived postembryonically from six vulval precursor cells (VPCs) termed P3.p–P8.p (see Figure 1). All six VPCs are competent to receive an inductive signal from a specialized somatic gonadal cell, the anchor cell (AC), during the L2 stage. P6.p, which is closest to the AC, is induced to generate the 1° vulval lineage, producing the inner cells of the vulva. The P5.p and P7.p cells generate the 2° vulval lineages, producing the outer cells of the vulva. The P3.p, P4.p, and P8.p are not induced and adopt the 3° fate, and fuse to the hypodermal syncytium hyp7. The L4 stage vulva comprises 22 differentiated cells that are descendents of P5.p, P6.p, and P7.p. There are seven different vulval cell types: vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF (21).

Each vulval cell type has specialized roles that contribute to vulval function and morphology. For example, vulF cells, the innermost vulval cells, contact the AC and the uterus. As such, they are the targets for AC invasion, thus creating the vulval-uterine connection required for egg-laying (22). Vulval muscles that regulate egg-laying connect

to the vulva between vulC and vulD, and the outermost portion of the vulva comprises the vulA cells, which attach the vulva to the hypodermis (21). The unique patterns of gene expression in each of the vulval cell types are likely responsible for their individual properties.

2.2 Vulval Organogenesis as a Model GRN

While the organs of many higher species (i.e., the mouse kidney with its roughly one million nephrons) may contain hundreds of millions of cells, the *C. elegans* vulva, as described above, contains only 22 cells. This relative simplicity allows for easier identification and mapping of the intra- and intercellular communication that occurs during organogenesis. In spite of this relative simplicity though, the vulva possesses many hallmarks of an organ. It is composed of rudimentary tissue types, specifically its seven cell types, and it performs a specialized task. It also uses several mechanisms that are common to organogenesis such as reciprocal inductive interactions and epithelial cell polarity. For example, during vulval organogenesis an inductive interplay between the vulva and the gonad results in cell fate specification. The anchor cell (a specialized cell of the somatic gonad) ensures the 1˚ vulval fate (23), and subsequently the vulF (1˚ vulval fate cells) induce uv1 cell fate in the gonad (24). Additionally, epithelial cell polarization, as evidenced by the reversed polarity of the 2˚ vulval lineages, has been demonstrated by both conserved cell division patterns and gene expression (25).

Other key properties of *C. elegans* include its amenability to genetic manipulation, its invariant cell lineage, and its well-characterized genome. Additional

advantages of the vulva as a model for systematic study of organogenesis include that it is not essential for viability, thus allowing for the *in vivo* observation of varied genetic perturbations, and that we have access to a wide variety of reporter genes with spatially and temporally unique expression patterns in the seven vulva cell types (21, 26).

2.3 Dissection of the Vulval GRN

EGF, Notch, and Wnt signaling pathways specify which VPCs generate the 1° and 2° lineages (21), but we are just now identifying the network of transcription factors that control cell-fate differentiation of the seven vulval cell types. When I commenced my thesis research, four transcription factors had been identified as major regulators of vulval cell differentiation and morphogenesis: *lin-11* (LIM homeodomain protein), *lin-29* (C2H2-type zinc finger, heterochronic gene pathway member), *cog-1* (Nkx6.1/6.2 homeoprotein), and *egl-38* (Pax2/5/8 protein). Mutants of these transcription factors were originally isolated from ethyl methanesulfonate (EMS) mutagenesis screens and subsequently characterized. The original mutant alleles of *lin-11*, *lin-29*, and *egl-38* were isolated from a large EMS mutagenesis screen (145 fertile mutants) for egg-laying defective (Egl) mutants (27). *cog-1(sy275)* was isolated from an EMS screen for Egl mutants with abnormal late-L3 vulval morphology (28). These screens, and others like them, are biased toward the isolation of mutants with strong morphological defects (i.e., protruding vulva, no vulva, etc.) and for those that act nonredundantly. The *C. elegans* genome encodes approximately 934 known and putative transcription factors (29). Since functions have not been observed for many of these, we find it highly plausible that

transcription factors that act redundantly, or that play more subtle roles in vulval development exist. It will likely require novel approaches to old methods and implementation of newly developed post-genome techniques to identify these factors.

Thesis Prelude

The *C. elegans* vulva provides us with an excellent opportunity for studying a post-embryonic model of organogenesis. Its level of complexity strikes an elegant balance between the almost overwhelming complexity of mammalian organogenesis and the relative simplicity of single cell-type differentiation and specification. My project initially focused on the specification of the vulA cells. I took two different approaches for analyzing this process: a reverse genetic screen utilizing an RNA interference (RNAi) library, and phylogenetic footprinting coupled with enhancer element dissection.

Libraries of RNAi clones consisting of the majority of genes in *C. elegans* genome have become available in the last several years. We compiled a list of known and predicted *C. elegans* transcription factors and generated a transcription factor RNAi library consisting of 836 clones. I conducted a screen in a *zmp-1::GFP* (zinc metalloproteinase) background using our transcription factor subset of the whole-genome RNAi library. *zmp-1* is expressed in the vulA, vulD, and vulE cells, but its vulA expression has been shown to be differentially regulated from that in vulD and vulE (26). This screen identified *nhr-67* (ortholog of *Drosophila tailless*) and *nhr-113* (orphan nuclear hormone receptor) as positive regulators of *zmp-1* expression in vulA cells. *nhr-113* appears to have a narrow role in vulval organogenesis; perhaps only as a partial

regulator of vulA cell differentiation. It however, serves as an example of a mutation with a subtle phenotype that was less likely to be identified by a traditional mutagenesis screen.

Furthering our analysis of *zmp-1*, we used Cistematic to analyze the upstream region of *C. elegans zmp-1*, and its homologs in the related species *C. briggsae* and *C. remanei*. Cistematic is a software package that incorporates the phylogenetic-footprinting and motif-finding capabilities of MEME, AlignACE, Co-Bind, and Footprinter for identifying *cis*-regulatory elements (30). This scan identified four conserved motifs upstream of the ZMP-1 translational start site. Characterization of these motifs, by deleting them from a *zmp-1::GFP* transgene, revealed that three of these four motifs exhibit *cis*-regulatory activity.

The availability of thirty transgenes with a variety of expression patterns in the cells of the vulva allowed me to further define the roles of two transcription factors acting in the vulval gene regulatory network: *lin-11* and *lin-29*. The heterochronic gene pathway member *lin-29* was originally thought to play a direct role in the temporal expression of genes in the vulva. The analysis of multiple pairwise interactions between *lin-29* and members of the vulval GRN however, has shown that *lin-29*'s role is more nuanced than previously perceived. We now know that *lin-29*, at least partially, regulates *lin-11* expression to direct a major component of vulval cell fate specification.

The final aspect of my project further dealt with the concept of *cis*-regulation of transcription, and was done in collaboration with Takao Inoue and Marian Walhout. We were able to use bioinformatics techniques to identify nine ~ 200-bp vulva-specific

enhancer elements associated with six genes. By submitting these sequences to yeast one-hybrid analysis we were able to isolate five transcription factors that bind these regions in yeast nuclei. Further investigation of these transcription factors identifies one of them as a new member of the vulval gene regulatory network. These results show that by taking multiple approaches, and embracing post-genomic sequencing technologies, previously unidentified members of gene regulatory networks can be unveiled.

References

1. Davidson EH (2006) The regulatory genome: gene regulatory networks in development and evolution (Academic, Burlington, MA; San Diego) pp. xi, 289.
2. Davidson EH & Levine MS (2008) Properties of developmental gene regulatory networks. Proc Natl Acad Sci U S A 105(51):20063–20066.
3. Levine M & Davidson EH (2005) Gene regulatory networks for development. Proc Natl Acad Sci U S A 102(14):4936–4942.
4. Davidson EH (2001) Genomic regulatory systems: development and evolution (Academic Press, San Diego) pp. xii, 261.
5. Lee TI & Young RA (2000) Transcription of eukaryotic protein-coding genes. Annu Rev Genet 34:77–137.
6. Stathopoulos A & Levine M (2005) Genomic regulatory networks and animal development. Dev Cell 9(4):449–462.
7. Oliveri P & Davidson EH (2007) Development. Built to run, not fail. Science 315(5818):1510–1511.
8. Longabaugh WJ, Davidson EH & Bolouri H (2005) Computational representation of developmental genetic regulatory networks. Dev Biol 283(1):1–16.
9. Usui K, Goldstone C, Gibert JM & Simpson P (2008) Redundant mechanisms mediate bristle patterning on the Drosophila thorax. Proc Natl Acad Sci U S A 105(51):20112–20117.
10. Gross JB et al. (2008) Synteny and candidate gene prediction using an anchored linkage map of Astyanax mexicanus. Proc Natl Acad Sci U S A 105(51):20106–20111.

11. Matthias P & Rolink AG (2005) Transcriptional networks in developing and mature B cells. Nat Rev Immunol 5(6):497–508.

12. Shaywitz DA & Melton DA (2005) The molecular biography of the cell. Cell 120(6):729–731.

13. Brunskill EW et al. (2008) Atlas of gene expression in the developing kidney at microanatomic resolution. Dev Cell 15(5):781–791.

14. Hong JW, Hendrix DA, Papatsenko D & Levine MS (2008) How the Dorsal gradient works: insights from postgenome technologies. Proc Natl Acad Sci U S A 105(51):20072–20076.

15. Biemar F et al. (2006) Comprehensive identification of Drosophila dorsal-ventral patterning genes using a whole-genome tiling array. Proc Natl Acad Sci U S A 103(34):12763–12768.

16. Zeitlinger J et al. (2007) Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes Dev 21(4):385–390.

17. Bonneau R et al. (2007) A predictive model for transcriptional control of physiology in a free living cell. Cell 131(7):1354–1365.

18. Tumpel S, Cambronero F, Sims C, Krumlauf R & Wiedemann LM (2008) A regulatory module embedded in the coding region of Hoxa2 controls expression in rhombomere 2. Proc Natl Acad Sci U S A 105(51):20077–20082.

19. Ben-Tabou de-Leon S & Davidson EH (2007) Gene regulation: gene control network in development. Annu Rev Biophys Biomol Struct 36:191.

20. Ochoa-Espinosa A et al. (2005) The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. Proc Natl Acad Sci U S A 102(14):4960–4965.

21. Sternberg PW. Vulval development. (WormBook, Community TCeR, ed.).

22. Sherwood DR & Sternberg PW (2003) Anchor cell invasion into the vulval epithelium in *C. elegans*. Dev Cell 5(1):21–31.

23. Wang M & Sternberg PW (2000) Patterning of the *C. elegans* 1° vulval lineage by RAS and Wnt pathways. Development 127(23):5047–5058.

24. Chang C, Newman AP & Sternberg PW (1999) Reciprocal EGF signaling back to the uterus from the induced *C. elegans* vulva coordinates morphogenesis of epithelia. Curr Biol 9(5):237–246.

25. Green JL, Inoue T & Sternberg PW (2008) Opposing Wnt pathways orient cell polarity during organogenesis. Cell 134(4):646–656.

26. Inoue T, Wang M, Ririe TO, Fernandes JS & Sternberg PW (2005) Transcriptional network underlying *Caenorhabditis elegans* vulval development. Proc Natl Acad Sci U S A 102(14):4972–4977.

27. Trent C, Tsuing N & Horvitz HR (1983) Egg-laying defective mutants of the nematode Caenorhabditis elegans. Genetics 104(4):619–647.

28. Palmer RE, Inoue T, Sherwood DR, Jiang LI & Sternberg PW (2002) *Caenorhabditis elegans cog-1* locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. Dev Biol 252(2):202–213.

29. Reece-Hoyes JS et al. (2005) A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol 6(13):R110.

30. Mortazavi A, Leeper Thompson EC, Garcia ST, Myers RM & Wold B (2006) Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. Genome Res 16(10):1208–1221.

31. Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL & Meyerowitz EM (2006) Genome-wide analysis of gene expression during early Arabidopsis flower development. PLoS Genet 2(7):e117.

32. Baugh LR et al. (2005) The homeodomain protein PAL-1 specifies a lineage-specific regulatory network in the C. elegans embryo. Development 132(8):1843–1854.

33. Gaudet J & Mango SE (2002) Regulation of organogenesis by the Caenorhabditis elegans FoxA protein PHA-4. Science 295(5556):821–825.

34. Gaudet J, Muttumu S, Horner M & Mango SE (2004) Whole-genome analysis of temporal gene expression during foregut development. PLoS Biol 2(11):e352.

35. Hobert O (2008) Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. Proc Natl Acad Sci U S A 105(51):20067–20071.

36. Imai KS, Levine M, Satoh N & Satou Y (2006) Regulatory blueprint for a chordate embryo. Science 312(5777):1183–1187.

37. Stathopoulos A & Levine M (2002) Dorsal gradient networks in the Drosophila embryo. Dev Biol 246(1):57–67.

38. Nikitina N, Sauka-Spengler T & Bronner-Fraser M (2008) Dissecting early regulatory relationships in the lamprey neural crest gene network. Proc Natl Acad Sci U S A 105(51):20083–20088.

39. Mori AD et al. (2006) Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. Dev Biol 297(2):566–586.

40. Mu X & Klein WH (2004) A gene regulatory hierarchy for retinal ganglion cell specification and differentiation. Semin Cell Dev Biol 15(1):115–123.

41. Hoffman BG et al. (2008) Identification of transcripts with enriched expression in the developing and adult pancreas. Genome Biol 9(6):R99.

42. Tamplin OJ et al. (2008) Microarray analysis of Foxa2 mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. BMC Genomics 9:511.

43. Georgescu C et al. (2008) A gene regulatory network armature for T lymphocyte specification. Proc Natl Acad Sci U S A 105(51):20100–20105.

44. Koide T, Hayata T & Cho KW (2005) *Xenopus* as a model system to study transcriptional regulatory networks. Proc Natl Acad Sci U S A 102(14):4943–4948.

45. Morley RH et al. (2009) A gene regulatory network directed by zebrafish No tail accounts for its roles in mesoderm formation. Proc Natl Acad Sci U S A 106(10):3829–3834.

Figure Legend

Figure 1. Vulval lineages. The fates of the six vulval precursor cells (VPCs) are specified during the third larval stage (L3). The 1˚ and 2˚ VPCs divide and generate the 22 cells of the adult vulva. The 1˚ lineage generates four vulE and four vulF cells. The 2˚ lineage generates four vulA cells, two vulB1 cells, two vulB2 cells, four vulC cells, and two vulD cells. The 3˚ lineage fuses with the hypodermal syncytium hyp7. L indicates a longitudinal axis of cell division; T, transverse axis of cell division; N, no cell division. The underlined Ls indicate where the vulva adheres to the ventral cuticle. (Figure taken from Sternberg, WormBook).

Figure 1

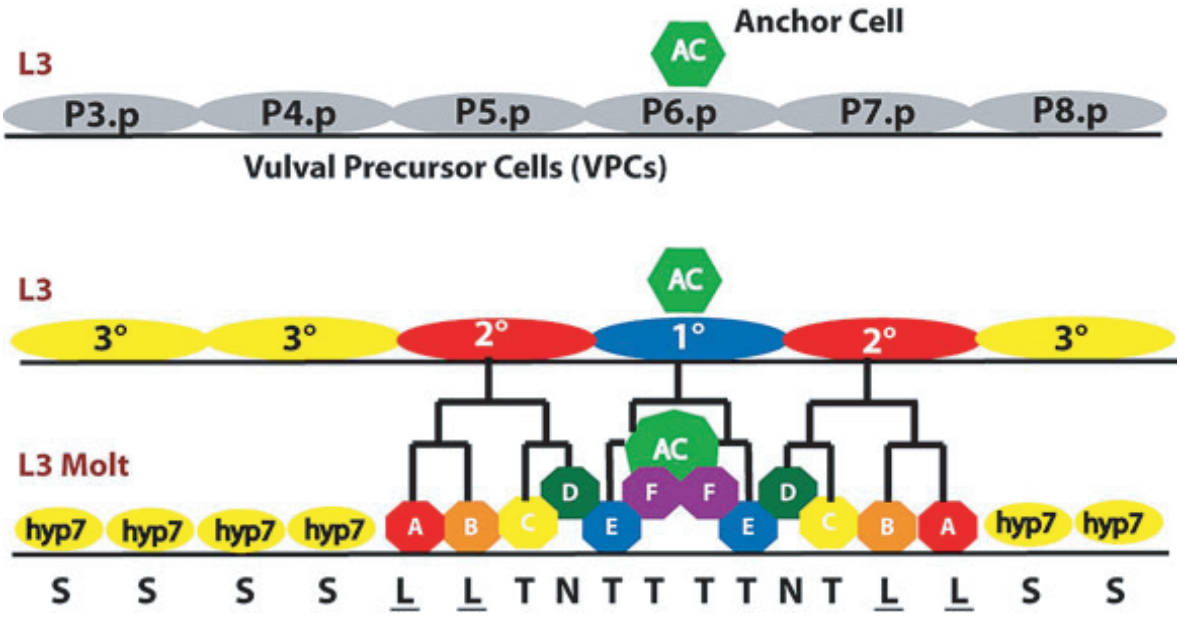


Figure from Sternberg, WormBook

Table 1: A sampling of gene regulatory networks in diverse species

Species	Process	Reference
Arabidopsis	early flower development	(31)
C. elegans	C-lineage specification in the embryo	(32)
C. elegans	foregut differentiation	(33, 34)
C. elegans	neuron specification	(35)
chicken	hindbrain development	(18)
Ciona	embryogenesis	(36)
Drosophila	Bicoid (Bcd): anterior-posterior patterning of the embryo	(20)
Drosophila	bristle patterning on the thorax	(9)
Drosophila	Dorsal (Dl): dorsal-ventral patterning of the embryo	(3, 37)
Halobacterium salinarum	physiology of free living cell	(17)
lamprey	neural crest formation	(38)
mouse	cardiac development	(39)
mouse	eye development	(40)
mouse	hindbrain development	(18)
mouse	pancreatic development	(41)
mouse	regulation of gastrula organizer	(42)
mouse	renal development	(13)
mouse	T lymphocyte fate specification	(43)
sea urchin	endomesoderm specification and patterning	(19)
Xenopus	mesoderm specification and patterning	(44)
zebrafish	mesoderm specification and patterning	(45)

CHAPTER II:

The *Caenorhabditis elegans* Vulva: A Post-Embryonic Gene Regulatory Network Controlling Organogenesis

Ted O. Ririe, Jolene S. Fernandes, and Paul W. Sternberg

A shorter version of this chapter was published in *PNAS*, **105**(51):20095-20099 (2008).

Abstract

The *Caenorhabditis elegans* vulva is an elegant model for dissecting a gene regulatory network (GRN) that directs post-embryonic organogenesis. The mature vulva comprises seven cell types (vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF), each with its own unique pattern of spatial and temporal gene expression. The mechanisms that generate the spatially defined pattern of these vulval fates are not well understood. Here we describe a provisional transcription factor network that defines some of these discrete cell types. Reverse genetic screens identified novel components of the vulval GRN, including *nhr-113* in vulA. Several transcription factors (*lin-11*, *lin-29*, *cog-1*, *egl-38*, and *nhr-67*) interact with each other and act in concert to regulate target gene expression in the diverse vulval cell types. For example, a pivotal role of *egl-38* (Pax2/5/8) is to stabilize the vulF fate by positively regulating vulF characteristics and by inhibiting characteristics associated with the neighboring vulE cells. Computational approaches have been successfully used to identify functional *cis*-regulatory motifs in the *zmp-1* (zinc metalloproteinase) promoter. Collectively, these findings provide an overview of the regulatory network architecture for each vulval cell type.

Introduction

Complex interactions of signaling molecules, transcription factors and effector genes direct spatial and temporal patterning during organogenesis (1). The differentiation and morphogenesis of the *Caenorhabditis elegans* vulva is useful for studying the gene regulatory network (GRN) of larval stage organogenesis due to its invariant cell lineage, its amenability to genetic manipulation, and the availability of reporter genes with a wide variety of spatial and temporal expression patterns in the seven vulval cell types (2). Previously we described a regulatory network of interactions between a set of evolutionarily conserved transcription factors and an array of genes expressed in the differentiated cells of the *C. elegans* vulva (3). Here, we briefly review vulval development and aspects of the provisional GRN directing its organogenesis. We then refine the network model by synthesizing data gathered by investigating additional pairwise interactions, conducting RNAi screens and computationally analyzing gene regulatory sequences. This analysis has uncovered common network themes such as boundary formation, combinatorial control, and stable feedback loops. Also, these additional data reflect precise spatial and temporal gene expression in an intact organism and demonstrate that the overall network architecture is unique for each of the vulval cell types.

The life cycle of *C. elegans* consists of four larval stages (L1–4) and an adult stage, each stage separated by a molt (2). The *C. elegans* vulva is derived post-embryonically from six vulval precursor cells (VPCs) termed P3.p–P8.p. All VPCs are competent to receive an inductive signal from a specialized somatic gonadal cell, the anchor cell (AC), during the L2 stage. P6.p, which is closest to the AC, is induced to

adopt the 1° vulval lineage fate that constitutes the inner portion of the vulva. The P5.p and P7.p cells adopt the 2° fate and give rise to the outer cells of the vulva. P3.p, P4.p, and P8.p are uninduced and adopt the 3° fate by fusing to the hypodermal syncytium (hyp7). The L4 stage vulva comprises 22 differentiated cells that are descendents of P5.p, P6.p, and P7.p, and that are of seven different types: vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF (Figure S1). EGF, Notch and Wnt signaling pathways specify the 1° and 2° cell fates, but we are just now identifying the network of transcription factors that control cell-fate differentiation in the seven vulval cell types (2).

The *C. elegans* vulva allows passage of sperm and eggs by connecting the uterus to the outside environment (2). Each vulval cell type has specialized physiological roles that contribute to vulval function and morphology. For example, vulF cells, the innermost vulval cells, contact the AC and the uterus; they are the target for AC invasion, thus creating the vulval-uterine connection required for egg-laying (4). Vulval muscles that regulate egg-laying connect to the vulva between vulC and vulD, and the outermost portion of the vulva comprises the vulA cells, which attach the vulva to the hypodermis (2). The unique patterns of gene expression in each of the vulval cell types are likely responsible for their individual properties. Comparison of the vulval GRN to those in other organisms is necessary for expanding our knowledge of organ development.

Results and Discussion

Functional Roles of Gene Expression During Vulval Differentiation. Much is known about the signaling network that initiates vulval development and sets up the general

pattern of cell differentiation, but our understanding of the GRN that specifies the terminal seven vulval cell types is limited (2). Five transcription factors (*lin-11*, *lin-29*, *cog-1*, *egl-38*, and *nhr-67*) are major regulators of cell-fate determination and morphogenesis in the vulva. *lin-11* encodes a LIM homeodomain protein, a class of proteins consisting of a homeodomain and two specialized zinc-fingers called LIM domains (5). These proteins have been identified in organisms as diverse as arthropods and vertebrates, and functional analyses of LIM homeodomain family members highlight their common role in differentiation and pattern formation (6, 7). *lin-29*, a C2H2 type zinc finger, plays a role in many events occurring at the larva-to-adult transition, including terminal differentiation of the seam cells (8), spicule development in the male tail (9), vulval morphogenesis (10) and formation of the vulval-uterine-seam cell connection (11). *cog-1* encodes a Nkx6.1/6.2 homeoprotein transcription factor (12); Nkx6.1 proteins in vertebrates have been implicated in neuronal and pancreatic endocrine cell formation (13, 14). *egl-38* encodes a Pax2/5/8 protein, which are known to be involved in organogenesis; e.g., mouse *Pax2* mediates nephrogenesis during kidney development (15, 16). *nhr-67* is an ortholog of *Drosophila melanogaster tailless* (*tll*) (17). The *tailless* nuclear hormone receptor plays major roles in determining the terminal-specific structures of the *Drosophila* embryo and in neuronal development (18). The pattern of vulval cell types is specified by the differential interactions of the transcription factors that operate within each cell (2).

We have identified 30 genes dynamically expressed in specific subsets of the cells of the mature *C. elegans* vulva (Table S1 has a complete list). These genes encode transcription factors, guidance cues, proteases, structural proteins, signaling molecules

and novel proteins with unknown function. The physiological relevance is known for several genes. For example, *egl-17*, which encodes a fibroblast growth factor (FGF)-related ligand, is necessary for migration of the sex myoblasts to the vulva (19). Induction of the uterine uv1 cells depends on the epidermal growth factor (EGF) family member *lin-3* (20). *bam-2* (neurexin-related transmembrane protein) and *syg-2* (transmembrane immunoglobulin superfamily protein) are required for vulval innervation (21, 22). *sqv-4*, which encodes an UDP-glucose dehydrogenase-related protein, is involved in the structural integrity and morphology of the vulva (23). The significance of some of the genes expressed in the vulva is not yet known, including: the *Drosophila empty spiracles* (*ems*) homolog *ceh-2* (24), a cadherin-related protein encoded by *cdh-3* (25), and *zmp-1* which encodes a zinc metalloproteinase (MT4-MMP-related) (24, 26). Genetic perturbations that result in altered expression patterns of these effector genes are helpful in elucidating the regulatory network controlling vulval organogenesis.

Several additional genes with detectable expression in the mature vulva, including two putative transcription factors, have been identified since we last described the vulval GRN (3). The Pax2/5/8 gene *pax-2* is expressed exclusively in the vulD cells (17). *pax-2* is the result of a recent duplication of the *egl-38* gene (26). *egl-38*, described above, has been previously identified as a regulator of cell-fate specification in the *C. elegans* vulva (3, 16, 20). *lin-39* encodes a Hox protein that is an ortholog of *Drosophila Sex combs reduced* (*Scr*) (27, 28). *lin-39* is required for proper acquisition of 1°, 2° and 3° vulval fates (29, 30). During the late L3 stage, *lin-39* expression increases in the vulA precursor cells; this expression persists in vulA until late L4 (31). The Patched-related protein DAF-6 is expressed in vulE and vulF (32). Hao et al. reported the expression of seven

hedgehog-related genes in the vulva: *grd-5* in vulB and vulD, *grd-12* in vulC, *grl-4* in vulA and vulB, *grl-10* in vulA and vulB, *grl-15* in vulB, vulC, vulD, and vulE, *grl-25* in vulA, and *grl-31* in vulF (33). Lastly, *nas-37* (34), which encodes a metalloprotease, is expressed in vulB (data not shown).

Conserved Regulatory Strategies in the Vulval GRN. Previously we described how the expression patterns of a subset of reporter constructs are affected in transcription factor mutant backgrounds (3). We have increased the number of known interactions more than twofold, from 15 to 36, and also identified many of the regulatory relationships among the transcription factors (Table 1, Figure S2 and Table S2). This network includes strategies that are shared by other GRNs. For example, the cell-type specific expression of *cog-1* appears to be restricted to the vulC and vulD cells by a variety of mutual and autoregulatory controls (17). In a second example, *lin-11* is necessary for vulA-specific expression of *nhr-67* (17). In turn, *nhr-67* is necessary for the expression of the vulA effector gene *zmp-1*. The differentiated state of vulA may be further stabilized, since *nhr-67* is positively autoregulated in vulA. This is an instance where multiple positive inputs, including feedback loops, ensure the maintenance of a terminal cell fate.

Other network strategies that are present in the vulva include combinatorial control circuits. *egl*-17 expression in vulF is not perturbed by either *cog-1* or *egl-38* mutants (17). Simultaneous deletion of both genes, however, leads to derepression of *egl-17* expression in vulF. This redundancy possibly acts to ensure proper execution of cell fate. Finally, negative autoregulation, as in *cog-1* in vulA, vulB, vulE, and vulF and

nhr-67 in vulC, vulD, vulE, and vulF, appears to be a fundamental strategy employed in the vulval GRN (17). Negative autoregulation has been found to accelerate gene circuit response time and assists in making quick cell fate decisions (35). This is analogous to an automobile: the better the brakes, the faster it can travel between stop signs.

LIN-29 and LIN-11 Interact to Determine Vulval Cell Fate. By examining the effects of specific transcription factor mutations on a wider array of genes expressed in the mature vulva, we have found that the roles of some transcription factors in the vulval GRN are more complex than previously suspected. We originally classified *lin-29* as a potential temporal regulator of gene expression because it is required for vulval gene expression at the mid to late L4 stage (*egl-17* in vulC and vulD, *ceh-2* in vulC, and *zmp-1* in vulD and vulE) (3, 24), and it is known to function in a variety of developmental processes that occur at the transition from the L4 to adult stages (8–11). The above scenario is not as straightforward as originally perceived, because we now know that in a *lin-29* mutant background the mid-L4 expression of *lin-3* in vulF cells is not abolished, while expression of *dhs-31*, which is initiated in gravid adults, is abolished (Table 1, Figure S2).

We also show that LIN-29 is necessary for wild type levels of *lin-11* transgene expression, and is thus a key regulator of *lin-11* (Table 1). In addition to genes described previously (3), *lin-11* is also required for *dhs-31*, *egl-26*, *lin-3*, and *pepm-1* expression (Table 1). These data reinforce that, of the known transcription factors, *lin-11* exhibits the broadest effect on cell-type specific expression in the vulva (3). To date, *egl-17*

during L3 and *cdh-3* in vulF are the only vulva expression patterns not abolished in a *lin-11* background. Reporter constructs are powerful tools for identifying genes' spatial and temporal expression patterns. It should be noted, however, that they exhibit limitations, in particular, relevant regulatory motifs may be lacking. Consequently, it is conceivable that the reporter constructs tested in *lin-11* and *lin-29* backgrounds lack their full complement of repressor elements, and as such the pairwise interactions tested may not reveal all regulatory activity.

The interplay between *lin-11*, *lin-29*, and their downstream targets is complex. *lin-11::GFP* expression is not completely abolished in a *lin-29* mutant background, and several of the gene expression patterns perturbed by loss of *lin-11* are not affected in a *lin-29* deficient background (Figure S2). It is possible that LIN-29 acts in concert with another, as yet unidentified, factor to ensure the proper temporal and spatial expression of the general cell-fate regulator LIN-11. Considering LIN-29's role as a regulator of developmental timing, it seems logical to postulate that it could serve as a temporal input, while another factor served as the spatial input.

Identification of Additional Components of the Vulval GRN. Transcription factors identified by forward genetic screens are biased by ascertainment toward those with strong, nonredundant effects, as well as those that affect multiple aspects of vulval development. Two categories of transcription factors are relatively more difficult to identify in genetic screens. Mutations that result in lethality or other severe defects will not be identified for roles that a partial loss of function mutation might reveal in post-

embryonic development. Conversely, genes with more subtle developmental phenotypes are also more likely to be overlooked. We identified genes in both these categories by RNA interference (RNAi) screens of known and predicted transcription factors. One of these screens (475 transcription factors, Table S3) was conducted in a *ceh-2::YFP* background, a readout for vulB fate during the L4 stage, and identified *nhr-67* (*tailless*) (17). Since *nhr-67* deletion mutants die as young larvae, its role in vulval development was not previously identified. RNAi often causes a partial-loss-of-function phenotype instead of a null phenotype. It can also serve as a temporal or conditional downregulator of gene function. In the case of *nhr-67*, the larval lethality phenotype was bypassed since RNAi was administered to L1 larvae, thus making its vulval phenotypes visible.

The second screen (836 transcription factors, Table S4) was conducted in a *zmp-1::GFP* background, focusing on perturbations of vulA expression. This screen also identified *nhr-67,* and the orphan nuclear hormone receptor *nhr-113* as positive regulators of *zmp-1* expression in vulA cells. *nhr-113* might have a narrow role in vulval organogenesis since *nhr-113* RNAi has no effect on the regulation of several other genes: *cdh-3, ceh-2, dhs-31, lin-3*, or *pepm-1* (data not shown). These results, however, show that RNAi screens can identify new components of GRNs.

Differentiation of Discrete 1° **Vulval Cell Fates**. The 1° lineage of the vulva generates four vulE and four vulF cells. A signal from the AC and a Wnt signal are required for proper specification of these cell fates (36). The GRN, however, that acts downstream of these intercellular signals to guide differentiation of vulE and vulF fates is not known.

egl-38::GFP is detectable solely in the vulF cells (37); but analysis of the mutant *egl-38* with *nhr-67* RNAi and mutant *cog-1* shows that *egl-38* not only functions in vulF cells, but also inhibits both *ceh-2* and *egl-17* expression in vulE (17) as well as vulF (3). Consequently, *egl-38*, *cog-1* and *nhr-67* enforce spatial boundaries by preventing the 2° cell-fate associated genes *ceh-2* and *egl-17* from being expressed in 1° cells (Figure S2, vulE and vulF).

It was previously speculated that the vulF cells, which are nearer the AC, have higher levels of *nhr-67* activity versus *cog-1*, while vulE cells have higher levels of *cog-1* than *nhr-67* (17). The recent availability of an *egl-38::GFP* reporter has allowed us to dissect further the GRN controlling the vulE versus vulF fates. *egl-38* expression in the vulF cells is positively regulated by *nhr-67* (Figure 1). *egl-38* expression in the vulF cells is necessary for specification of the uv1 fate via regulation of the vulF specific gene *lin-3*, thus allowing for the proper development of a vulval-uterine connection (20). Conversely, *zmp-1* is expressed in vulE but not vulF, and *egl-38* is required for inhibiting *zmp-1* expression in the vulF cells (3). In vulE cells, where *egl-38* expression is absent, *lin-3* is not expressed and *cog-1* activated expression of *zmp-1* is observed (20). *egl-38* expression is unaffected in a *cog-1(sy275)* background, but in an *egl-38(n578)* background *egl-38::GFP* expression is decreased, suggesting that *egl-38* positively autoregulates in vulF (Figure 1). Therefore, *egl-38* appears to stabilize vulF fate by repressing vulE characteristics and by reinforcing its own expression (Figure 2).

In the vulE and vulF cells, *cog-1* and *nhr-67* negatively regulate both each other and themselves (17). We previously speculated that this might allow 1° cells to rapidly switch their fates upon altered intra- and intercellular inputs. Presumptive vulE cells can

induce uterine uv1 fate specification in the absence of vulF cells, thus ensuring the establishment of a proper vulval-uterine connection (20). These observations fit a model in which presumptive vulE and vulF cells are bipotential and postitional cues help specify their fates. In this model, vulE, would then be biased for increased *nhr-67* activity in the absence of vulF. *egl-38* levels would thus increase, and vulF characteristics would be activated, while vulE characteristics would be inhibited.

Dissection of *zmp-1* Regulatory Elements. The vulA cells occupy a unique position in the vulva in that they form the outermost ring of cells and fuse to the surrounding hypodermal syncytium. *zmp-1* (zinc metalloproteinase) is first expressed in vulD and vulE cells beginning at the late L4 stage, and in the vulA cells at the L4 to adult transition (24). The fact that vulA-specific expression of *zmp-1* is initiated in a different temporal window than its expression in vulD and vulE is particularly interesting, because in a *lin-29* background vulD and vulE expression is abolished yet vulA expression is unaffected. Since *lin-29* affects early *zmp-1* expression (vulD and vulE) but not late expression (vulA), and since *lin-29* temporally regulates gene expression, it seems likely that modular *cis*-regulatory elements drive *zmp-1* expression.

A 386-bp fragment of upstream intergenic sequence (mk50-51) is sufficient to drive vulA and vulE expression, but not the vulD-specific expression of *zmp-1* (38). To computationally analyze *zmp-1* regulatory sequences defined by this transgenic reporter gene analysis we used Cistematic (39). This software incorporates the motif-finding and phylogenetic footprinting capabilities of MEME, AlignACE, Co-Bind and FootPrinter to

identify *cis*-regulatory elements. Specifically, we analyzed the intergenic region upstream of *zmp-1* and its orthologs in *C. briggsae* (CBG09053) and *C. remanei* (CRE04503). Addition of the *C. remanei* genome refined the identification of *cis*-regulatory motifs in the *zmp-1* upstream region (Table 2, Figure S3). This analysis identified three motifs within the mk50-51 *zmp-1* enhancer element, and a fourth motif five base pairs upstream. We deleted the instances of the motifs via site-directed mutagenesis and analyzed their effects on *zmp-1* reporter expression. Deletion of element 103/4 decreased vulA expression and abolished vulE expression. Previous *cis*-regulatory studies did not identify an element competent to drive vulA expression in the absence of vulE. Deletion of element 107/8 decreased vulA and vulE expression of *zmp-1::GFP*. Deletion of element 105/6, however, resulted in ectopic expression of mk50-51 *zmp-1::GFP* in vulC, vulD, and vulF. Thus, these elements likely act as positive (103/4, 107/8) and negative (105/6) regulatory sites for controlling of *zmp-1* expression. While we are only beginning to define *cis*-regulatory architecture of vulval specific gene expression, the availability of additional genomes for computational analysis should accelerate this process.

Conclusion

The regulatory architecture of the vulval GRN differs in all seven cell types. We postulate that this accounts for the differences in vulval cell fate, function, and morphology. We find that the differentiation and organogenesis of the *C. elegans* vulva utilizes several types of gene regulatory strategies that have been identified in other

networks. For example, COG-1 participates in a network of mutual inhibitions with NHR-67 in the 1° vulval cells to differentiate between the vulE and vulF fates. This is reminiscent of the cross-inhibitory interaction of COG-1 and DIE-1 in the *C. elegans* ASE neurons (40). Vertebrate COG-1 homologs, the homeodomain proteins Nkx6.1 and Nkx6.2, also appear to act in a similar manner (3). These proteins interact with the transcription factors Dbx1 and Dbx2 in a network of mutual inhibitions to specify motor neuron and interneuron fates during neural tube development (15). In another example, *lin-11* function is necessary for EGL-17 (FGF) expression in the vulva. An analogous network interaction is present during heart development in mice. Isl1, which, like *lin-11*, is a LIM homeodomain transcription factor, is required for the expression of FGFs (41).

We also describe new interactions within the vulva that may aid in the understanding of analogous regulatory interactions in other transcriptional networks, since the majority of the transcription factors and effectors present in this GRN have relatives in a diverse array of organisms. These interactions include positive regulation of *lin-11* by the heterochronic transcription factor LIN-29, *lin-3* (EGF) by LIN-11, and *egl*-38 (Pax2/5/8) by NHR-67 (*tailless*). Our increased knowledge of the roles of transcription factors such as *lin-11* (LIM homeodomain) and *egl-38* (Pax2/5/8) may help to further characterize other GRNs. For example, a LIM homeodomain protein (*Lim1*) and a Pax2/5/8 protein (*Pax2)* are involved in murine kidney development (42). Thus the GRN that directs nephrogenesis in mice appears to share at least two components with the vulval network in *C. elegans*. Our understanding of the kidney morphogensis GRN in mammals may be enhanced by investigating the mouse orthologs of other components of the vulval GRN.

New approaches are needed to elucidate all the constituents of the complex regulatory architecture that directs organogenesis. Until now, most of the key players in vulval organogenesis have been isolated using traditional mutagenesis screens. These types of screens, however, are often limited to identifying only those factors with severe phenotypes. To account for the regulation of all genes that are expressed and function during vulva development it is apparent that other members of the vulval GRN are yet to be identified. There are transcription factors with major and minor effects. For example, LIN-11 is required in two tissues for egg-laying and regulates gene expression in multiple vulval cell types. By contrast, using an RNAi screen we were able to identify NHR-113, a factor that is possibly only required for fine-tuning gene expression in the vulA cells.

Analysis of the effects of various genetic perturbations on vulval organogenesis has revealed detailed spatial and temporal distinctions in the regulation of diverse yet related cell types. Our approach provides for precise and accurate study of gene expression in an intact organism, and unveils the distinct network architecture in the different cell types. Further dissection of the genomic network within the differentiated cell types would extend our knowledge of vulval organogenesis and could also provide further insights into organogenesis in other systems.

Materials and Methods

Genetics and RNAi. *C. elegans* strains were grown and constructed using standard protocols (43, 44). Reporter transgenes used were *cdh-3::CFP* [*syIs51*] (24), *dhs-31::YFP* formerly known as T04B2.6 [*sysIs101*] (24), *egl-26::GFP* [*kuIs36*] (45), *egl-38::GFP* [*guEx877*] (gift from Chamberlin lab) (37), *lin-3::GFP* [*syIs107*] (46), *lin-11::GFP* [*syIs80*] (47), *pepm-1::GFP* formerly known as F47B8.6 [syIs61] (24), *unc-53::GFP* [*bgEx11*] (48), *unc-53::GFP* [*syEx756*], mk50-51 *zmp-1::GFP* [*syEx1012*], mk50-51 motif 103/4Δ *zmp-1::GFP* [*syEx1009*], mk50-51 motif 105/6Δ *zmp-1::GFP* [*syEx1091*], and mk50-51 motif 107/8Δ *zmp-1::GFP* [*syEx1018*]. Alleles used were LGI, *lin-11(n389)*; LGII, *cog-1(sy275)*, *egl-38(n578), lin-29(n333)*, *lin-29(sy292)*; LGIV, *egl-38(n578)*.

nhr-67 RNAi by feeding was used in this study because the mutant allele *ok631* is lethal at the L1 stage (International *C. elegans* Knockout Consortium). The protocol used was described previously (17).

Generation of reporter transgenes. *syEx1009, syEx1091* and *syEx1018* were generated by deleting one motif each, 103/4 (5'-CGAGTACGTTTACAC-3'), 105/6 (5'-GTACGTATTGCTT-3'), or 107/8 (5'-AGAAAAAGTAGAAGG-3'), respectively, from the mk50-51 construct and replacing the motif with a *Sac*II restriction site. Forward primer TOR109 (5'-CAAATGGCGATTACCGTTGATGTTGAAGTGGC-3') lies upstream of a *Spe*I site in the mk50-51 construct and approximately 4 kb upstream of the motifs, and reverse primer TOR110 (5'-

CATATCCTGATCTTCCAGATAACTGCCGTCAC-3') lies downstream of a *Pst*I site in the mk50-51 construct and approximately 2.5 kb downstream of the motifs. To generate a construct with deleted motif 103/4, two primers were designed: TOR104 (5'-GTGTCGCCGCGGAATACTTTGATGATCAATACAG-3') is a reverse primer directly upstream of the motif and engineered with a *Sac*II restriction site and six bases of spacer sequence; TOR103 (5'-GAATACCCGCGGTGGTTTCTGTTCTTTCCGTTT-3') is a forward primer directly downstream of the motif that is also engineered with a *Sac*II site and six bases of spacer sequence. Using mk50-51 as the template, two PCRs were conducted using the following primer pairs: TOR109/104 and 103/110. Each PCR product was then digested with *Sac*II. The TOR109/104 product was then ligated to 103/110. This produced a ligation construct with motif 103/4 replaced by a *Sac*II site. The ligation product was then PCR amplified using the TOR109/110 primer pair. The resulting product, and mk50-51 were subsequently digested with both *Spe*I and *Pst*I. The digested PCR product was then ligated with the digested mk50-51 to generate the deletion construct. The 105/6 and 107/8 deletion constructs were generated by the same method using the following primers: TOR106 (5'-GTATCACCGCGGACATGAATCAGTTTGCATCTG-3'), TOR105 (5'-GACTCGCCGCGGGAAAAAAAGAGTAACAAGAAAAAG-3'), TOR108 (5'-GTATCCCCGCGGTGTTACTCTTTTTTTCAAGC-3') and TOR107 (5'-GAATCCCCGCGGGTATTAGTCGTAGTAGTAGTAT-3'). These constructs were then individually injected into the gonads of *unc-119(ed4); him-5(e1490)* animals (49) using *unc-119*(+) (50) and pBSK+ (Stratagene). *syEx756* and *syEx1012* were generated by

injecting the constructs pNP10 (48) and mk50-51 (38), respectively, into a *unc-119(ed4); him-5(e1490)* background.

Microscopy. 3 mM levamisole was used to anesthetize transgenic animals for observation using Nomarski DIC optics (http://www.nomarski.com). Images were captured with a monochrome Hamamatsu Orca II digital camera (http://www.hamamatsu.com) and Improvision Openlab 4.0.4 software (http://www.improvision.com). The fluorescent images were overlaid with their respective DIC images using Adobe Photoshop 7.0.1 (http://www.adobe.com).

Acknowledgments

We thank Ali Mortazavi for help with comparative genomics, Helen Chamberlin for *guEx877*, Andy Fire for GFP vectors, Si Hyun Kim for his assistance with the *zmp-1::GFP* RNAi screen, Mihoko Kato, Steven Kuntz, Erich Schwarz, and Xiaodong Wang for comments on the manuscript, and Takao Inoue for discussions. T.O.R. was supported by T32-GM007616. P.W.S. is an investigator with the Howard Hughes Medical Institute.

References

1. Davidson EH (2001) Genomic regulatory systems: development and evolution (Academic Press, San Diego) pp. xii, 261.
2. Sternberg PW. Vulval development. (WormBook, Community TCeR, ed.).
3. Inoue T, Wang M, Ririe TO, Fernandes JS & Sternberg PW (2005) Transcriptional network underlying *Caenorhabditis elegans* vulval development. Proc Natl Acad Sci U S A 102(14):4972–4977.
4. Sherwood DR & Sternberg PW (2003) Anchor cell invasion into the vulval epithelium in *C. elegans*. Dev Cell 5(1):21–31.
5. Freyd G, Kim SK & Horvitz HR (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene *lin-11*. Nature 344(6269):876–879.
6. Bach I (2000) The LIM domain: regulation by association. Mech Dev 91(1–2):5–17.
7. Hobert O & Westphal H (2000) Functions of LIM-homeobox genes. Trends Genet 16(2):75–83.
8. Rougvie AE & Ambros V (1995) The heterochronic gene *lin-29* encodes a zinc finger protein that controls a terminal differentiation event in *Caenorhabditis elegans*. Development 121(8):2491–2500.
9. Euling S, Bettinger JC & Rougvie AE (1999) The LIN-29 transcription factor is required for proper morphogenesis of the *Caenorhabditis elegans* male tail. Dev Biol 206(2):142–156.

10. Bettinger JC, Euling S & Rougvie AE (1997) The terminal differentiation factor LIN-29 is required for proper vulval morphogenesis and egg laying in *Caenorhabditis elegans*. Development 124(21):4333–4342.

11. Newman AP, Inoue T, Wang M & Sternberg PW (2000) The *Caenorhabditis elegans* heterochronic gene *lin-29* coordinates the vulval-uterine-epidermal connections. Curr Biol 10(23):1479–1488.

12. Palmer RE, Inoue T, Sherwood DR, Jiang LI & Sternberg PW (2002) *Caenorhabditis elegans cog-1* locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. Dev Biol 252(2):202–213.

13. Habener JF, Kemp DM & Thomas MK (2005) Minireview: transcriptional regulation in pancreatic development. Endocrinology 146(3):1025–1034.

14. Vallstedt A et al. (2001) Different levels of repressor activity assign redundant and specific roles to *Nkx6* genes in motor neuron and interneuron specification. Neuron 31(5):743–755.

15. Schedl A & Hastie ND (2000) Cross-talk in kidney development. Curr Opin Genet Dev 10(5):543–549.

16. Chamberlin HM et al. (1997) The *PAX* gene *egl-38* mediates developmental patterning in *Caenorhabditis elegans*. Development 124(20):3919–3928.

17. Fernandes JS & Sternberg PW (2007) The tailless ortholog *nhr-67* regulates patterning of gene expression and morphogenesis in the *C. elegans* vulva. PLoS Genet 3(4):e69.

18. Hartmann B, Reichert H & Walldorf U (2001) Interaction of gap genes in the *Drosophila* head: *tailless* regulates expression of *empty spiracles* in early embryonic patterning and brain development. Mech Dev 109(2):161–172.

19. Burdine RD, Branda CS & Stern MJ (1998) EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in *C. elegans*. Development 125(6):1083–1093.

20. Chang C, Newman AP & Sternberg PW (1999) Reciprocal EGF signaling back to the uterus from the induced *C. elegans* vulva coordinates morphogenesis of epithelia. Curr Biol 9(5):237–246.

21. Colavita A & Tessier-Lavigne M (2003) A Neurexin-related protein, BAM-2, terminates axonal branches in *C. elegans*. Science 302(5643):293–296.

22. Shen K, Fetter RD & Bargmann CI (2004) Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. Cell 116(6):869–881.

23. Hwang HY & Horvitz HR (2002) The *Caenorhabditis elegans* vulval morphogenesis gene *sqv-4* encodes a UDP-glucose dehydrogenase that is temporally and spatially regulated. Proc Natl Acad Sci U S A 99(22):14224–14229.

24. Inoue T et al. (2002) Gene expression markers for *Caenorhabditis elegans* vulval cells. Mech Dev 119 Suppl 1:S203–209.

25. Pettitt J, Wood WB & Plasterk RH (1996) *cdh-3*, a gene encoding a member of the cadherin superfamily, functions in epithelial cell morphogenesis in *Caenorhabditis elegans*. Development 122(12):4149–4157.

26. Wang X, Greenberg JF & Chamberlin HM (2004) Evolution of regulatory elements producing a conserved gene expression pattern in *Caenorhabditis*. Evol Dev 6(4):237–245.

27. Clark SG, Chisholm AD & Horvitz HR (1993) Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*. Cell 74(1):43–55.

28. Wang BB et al. (1993) A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. Cell 74(1):29–42.

29. Maloof JN & Kenyon C (1998) The Hox gene *lin-39* is required during *C. elegans* vulval induction to select the outcome of Ras signaling. Development 125(2):181–190.

30. Clandinin TR, Katz WS & Sternberg PW (1997) *Caenorhabditis elegans* HOM-C genes regulate the response of vulval precursor cells to inductive signal. Dev Biol 182(1):150–161.

31. Wagmaister JA, Gleason JE & Eisenmann DM (2006) Transcriptional upregulation of the *C. elegans* Hox gene *lin-39* during vulval cell fate specification. Mech Dev 123(2):135–150.

32. Perens EA & Shaham S (2005) *C. elegans daf-6* encodes a patched-related protein required for lumen formation. Dev Cell 8(6):893–906.

33. Hao L, Johnsen R, Lauter G, Baillie D & Burglin TR (2006) Comprehensive analysis of gene expression patterns of *hedgehog*-related genes. BMC Genomics 7:280.

34. Davis MW, Birnie AJ, Chan AC, Page AP & Jorgensen EM (2004) A conserved metalloprotease mediates ecdysis in *Caenorhabditis elegans*. Development 131(23):6001–6008.

35. Rosenfeld N, Elowitz MB & Alon U (2002) Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323(5):785–793.

36. Wang M (2000) Pattern formation during *Caenorhabditis elegans* vulval development. (PhD, California Institute of Technology, Pasadena, CA).

37. Rajakumar V & Chamberlin HM (2007) The *Pax2/5/8* gene *egl-38* coordinates organogenesis of the *C. elegans* egg-laying system. Dev Biol 301(1):240–253.

38. Kirouac M & Sternberg PW (2003) *cis*-Regulatory control of three cell fate-specific genes in vulval organogenesis of *Caenorhabditis elegans* and *C. briggsae*. Dev Biol 257(1):85–103.

39. Mortazavi A, Leeper Thompson EC, Garcia ST, Myers RM & Wold B (2006) Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. Genome Res 16(10):1208–1221.

40. Johnston RJ, Jr., Chang S, Etchberger JF, Ortiz CO & Hobert O (2005) MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. Proc Natl Acad Sci U S A 102(35):12449–12454.

41. Cai CL et al. (2003) Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. Dev Cell 5(6):877–889.

42. Yu J, McMahon AP & Valerius MT (2004) Recent genetic studies of mouse kidney development. Curr Opin Genet Dev 14(5):550–557.

43. Stiernagle T. Maintenance of *C. elegans*. (WormBook, Community TCeR, ed.).

44. Huang L & Sternberg PW. Genetic dissection of developmental pathways. (WormBook, Community TCeR, ed.).

45. Hanna-Rose W & Han M (2002) The *Caenorhabditis elegans* EGL-26 protein mediates vulval cell morphogenesis. Dev Biol 241(2):247–258.

46. Hwang BJ & Sternberg PW (2004) A cell-specific enhancer that specifies *lin-3* expression in the *C. elegans* anchor cell for vulval development. Development 131(1):143–151.

47. Gupta BP & Sternberg PW (2002) Tissue-specific regulation of the LIM homeobox gene *lin-11* during development of the *Caenorhabditis elegans* egg-laying system. Dev Biol 247(1):102–115.

48. Stringham E, Pujol N, Vandekerckhove J & Bogaert T (2002) *unc-53* controls longitudinal migration in *C. elegans*. Development 129(14):3367–3379.

49. Mello CC, Kramer JM, Stinchcomb D & Ambros V (1991) Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J 10(12):3959–3970.

50. Maduro M & Pilgrim D (1995) Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics 141(3):977–988.

51. Longabaugh WJ, Davidson EH & Bolouri H (2005) Computational representation of developmental genetic regulatory networks. Dev Biol 283(1):1–16.

Figure Legends

Figure 1. *egl-38* expression in the vulF cells is dependent on several regulatory inputs. (A–C) Nomarski (left), fluorescence (center), and overlaid (right). All animals displayed carry the *egl-38::GFP* [*guEx877*] transgene in their background. (A) In wild type animals, *egl-38* is detected exclusively in the vulF cells (arrows). (B) *nhr-67* RNAi results in the abolition of *egl-38* expression in the vulF cells (arrows). (C) *egl-38(n578)* mutants lose the ability to positively autoregulate their expression levels in vulF (arrows).

Figure 2. Differentiation of vulE vs. vulF. The positions of the vulE and vulF cells relative to the anchor cell (AC) are shown. vulF is closer spatially than vulE to the AC. The network diagram was generated using BioTapestry Editor 2.1.0 (http://www.biotapestry.org) (51). Linkages with arrowheads represent positive inputs and linkages with bar-heads represent repressor inputs for target gene expression. Black font indicates detectable expression levels and gray font indicates no detectable expression. This model presumes that *nhr-67* acts in the AC to differentiate between vulE and vulF cells. Signal X could be Ras, Wnt or some other signaling pathway. The blue linkage in vulF is indicated by a thicker line than the blue linkage in vulE because it is hypothesized that vulF receives higher levels of signal X-mediated *nhr-67* signal from the AC. The thick purple and pink linkages highlight differences in the network architecture between the vulE and vulF cells, respectively.

Figure S1. Development of the vulva. The VPCs divide once during early L3 resulting in the VPC daughters. The P5.p, P6.p, and P7.p daughters then divide again during mid-L3 to form the VPC granddaughters. During late L3 each of these cells, except D, divide again to create the VPC great granddaughter cells which constitute the mature vulva. The vulval cell types vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF migrate during the L4 stage as the vulva assumes its adult morphology.

Figure S2. GRN for *C. elegans* vulval development. The vulval GRN is divided into the cell types that comprise the vulva: vulA, vulB1/B2, vulC, vulD, vulE, and vulF. Cells with blue backgrounds arise from the 1° lineage and cells with green backgrounds arise from 2° lineages. *cog-1, egl-38, lin-11, lin-29*, and *nhr-67* regulate gene expression in the vulval cells and are clustered on the left in each frame. The remaining genes, clustered to the right in each frame and highlighted by colored boxes, are the genes that are expressed in the vulva in defined spatial and temporal patterns. The colored boxes highlighting these genes indicate when each gene's expression is first visualized. With the exception of *egl-17* expression in vulE and vulF, the expression of all reporter genes persists into adulthood. *egl-17* reporter transgene expression in vulE and vulF appears during the L3 stage and turns off during the early L4 stage, while the vulC and vulD expression appears during the mid-L4 stage and persists into adulthood. Light yellow squares indicate genes expressed beginning at early L4, yellow squares at mid L4, light green squares at late L4, and brown squares at adulthood as indicated. Not all of the genes are expressed at the same time in each cell type. For example, *zmp-1* is expressed in vulD and vulE at late L4, but it is not expressed until adulthood in vulA. Black font

indicates genes with detectable expression in a given cell type. Gray font indicates genes without detectable expression in a given cell type but that have been shown to have function in the cell type. In vulA, for example, *cog-1* reporter gene expression has not been identified, but in the *cog-1* mutant background *cog-1* expression is visible. Linkages with arrowheads indicate positive regulatory activity and linkages with bar-heads indicate repressor activity. We do not know if any of these linkages are direct.

Figure S3. Mutation of conserved subregions of the *zmp-1* vulval expression module differentially effect its expression. Several transcriptional reporter constructs containing a 386-bp enhancer element (blue box) located 2 kb upstream of the presumptive *zmp-1* translational start site (black arrow) were generated. The black boxes indicate exons. The orange, red, and purple rectangles indicate conserved motifs identified by Cistematic. Construct mk50-51 includes the vulval expression module attached to minimal *Δpes-10::GFP*. The green arrow indicates the GFP translational start site. Construct 103/4Δ is the same as mk50-51 except the motif CGAGTACGTTTACAC was deleted. In construct 105/6Δ the motif GTACGTATTGCTT was deleted. In construct 107/8Δ the motif AGAAAAAGTAGAAGG was deleted. In each deletion construct the deleted motif was replaced by a *Sac*II restriction site.

Figure 1

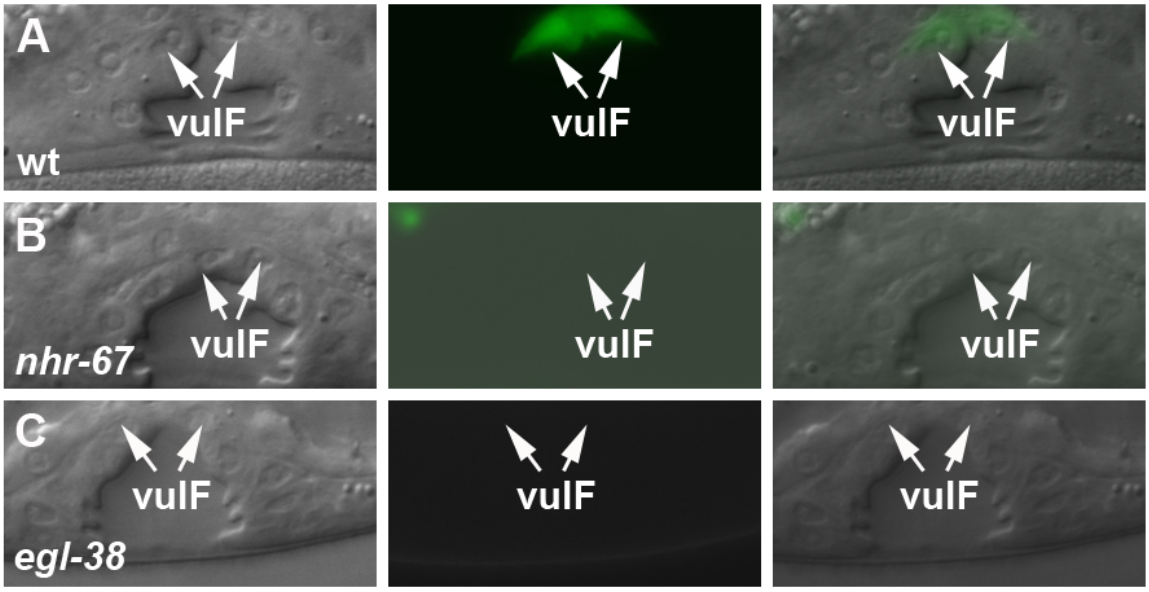

A
vulF
wt
vulF
vulF
B
vulF
nhr-67
vulF
vulF
C
vulF
egl-38
vulF
vulF

Figure 2

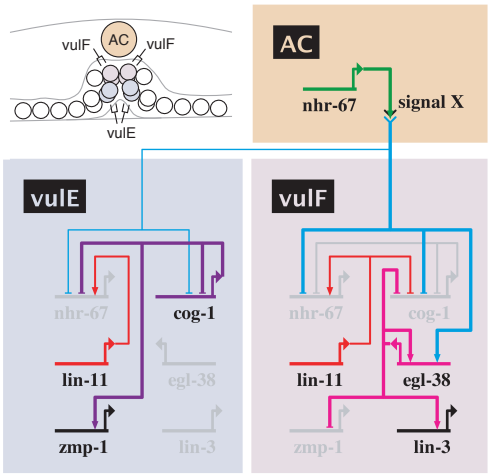

AC
vulF
vulF
vulE
AC
nhr-67
signal X
vulE
nhr-67
cog-1
lin-11
egl-38
zmp-1
lin-3
vulF
nhr-67
cog-1
lin-11
egl-38
zmp-1
lin-3

Figure S1

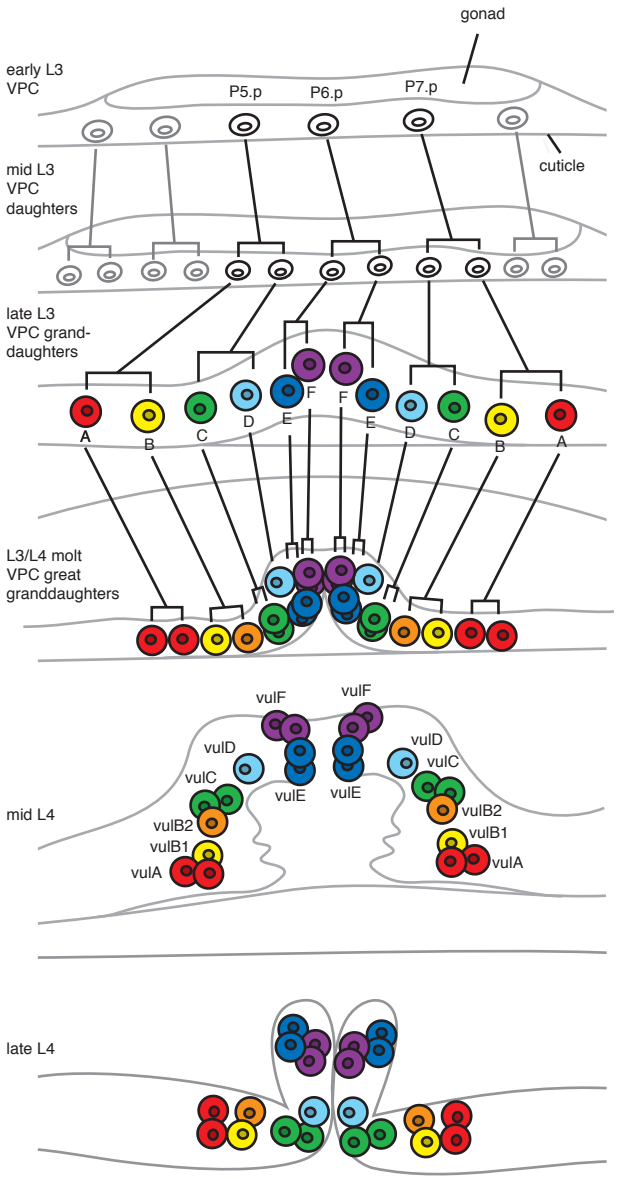

gonad
early L3
VPC
P5.p
P6.p
P7.p
cuticle
mid L3
VPC
daughters
late L3
VPC grand-
daughters
A
B
C
D
E
F
F
E
D
C
B
A
L3/L4 molt
VPC great
granddaughters
mid L4
vulF
vulF
vulD
vulC
vulE
vulE
vulD
vulC
vulB2
vulB1
vulA
vulB2
vulB1
vulA
late L4

Figure S2

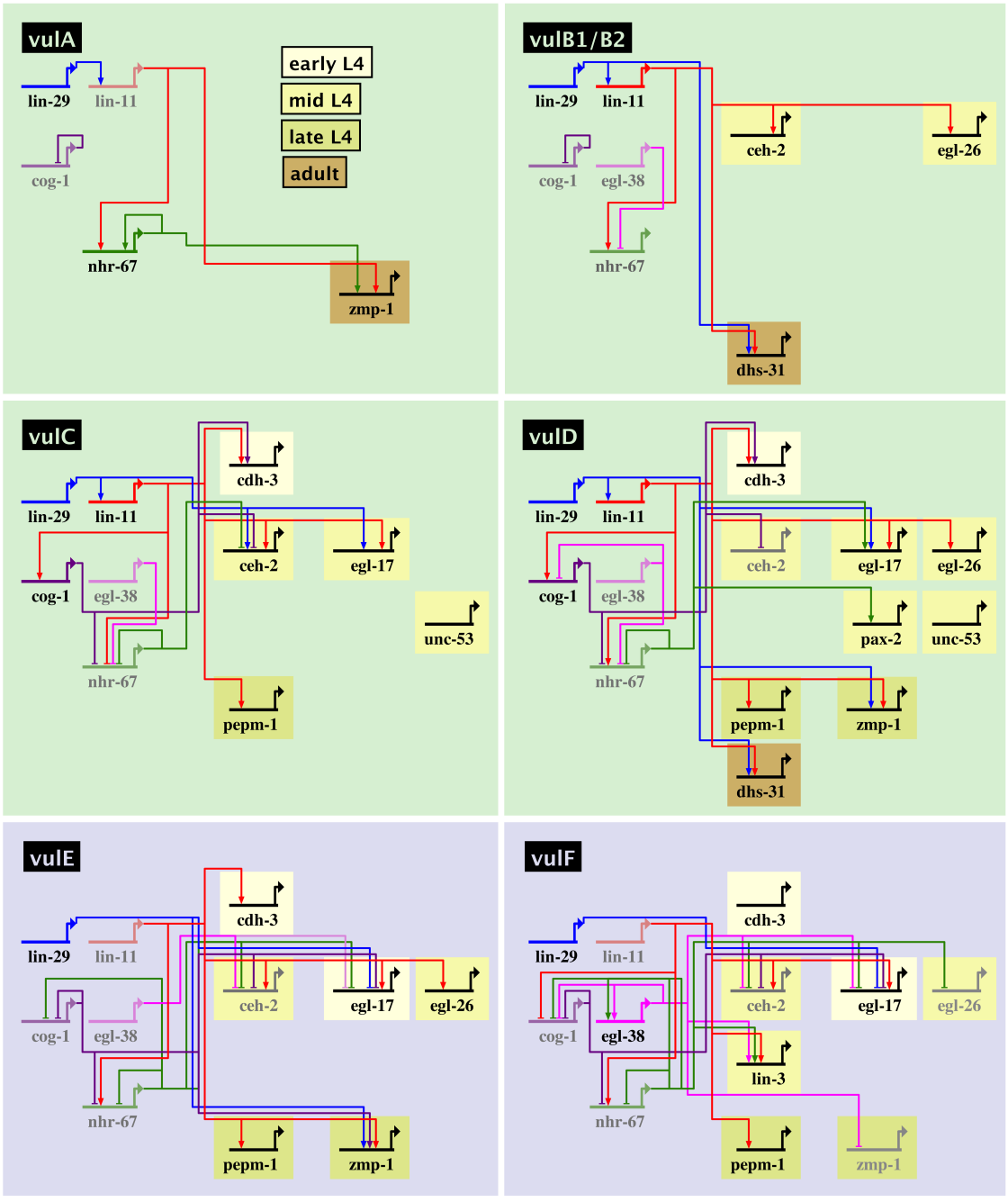

vulA
early L4
mid L4
late L4
adult
lin-29
lin-11
cog-1
nhr-67
zmp-1
vulB1/B2
lin-29
lin-11
ceh-2
egl-26
cog-1
egl-38
nhr-67
dhs-31
vulC
cdh-3
lin-29
lin-11
ceh-2
egl-17
cog-1
egl-38
unc-53
nhr-67
pepm-1
vulD
cdh-3
lin-29
lin-11
ceh-2
egl-17
egl-26
cog-1
egl-38
pax-2
unc-53
nhr-67
pepm-1
zmp-1
dhs-31
vulE
cdh-3
lin-29
lin-11
ceh-2
egl-17
egl-26
cog-1
egl-38
nhr-67
pepm-1
zmp-1
vulF
cdh-3
lin-29
lin-11
ceh-2
egl-17
egl-26
cog-1
egl-38
lin-3
nhr-67
pepm-1
zmp-1

Figure S3

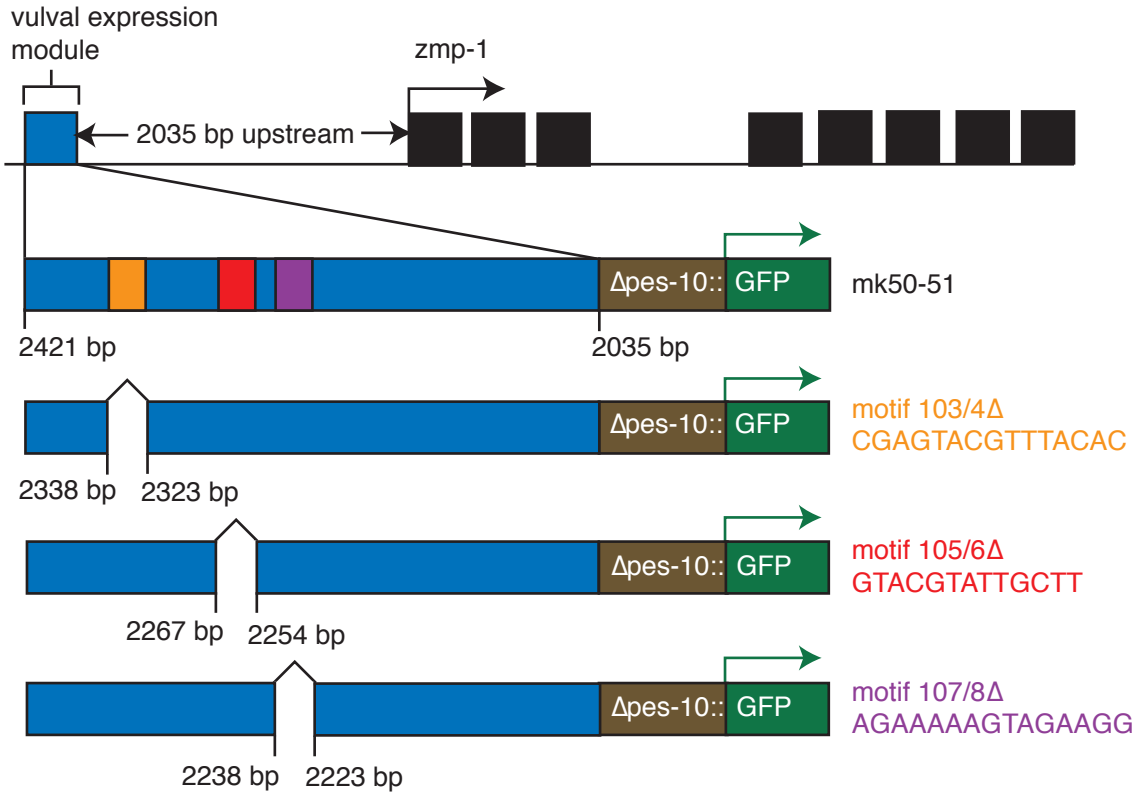

vulval expression module
zmp-1
2035 bp upstream
Δpes-10::
GFP
mk50-51
2421 bp
2035 bp
motif 103/4Δ
CGAGTACGTTTACAC
2338 bp
2323 bp
motif 105/6Δ
GTACGTATTGCTT
2267 bp
2254 bp
motif 107/8Δ
AGAAAAAGTAGAAGG
2238 bp
2223 bp

Table 1. Expression of *dhs-31*, *egl-26*, *lin-3*, *lin-11*, *pepm-1,* and *unc-53* in the vulva

Reporter	Mutation	vulA	vulB1	vulB2	vulC	vulD	vulE	vulF	*n*
dhs-31	+	0	100	100	0	100	0	0	31
dhs-31	*lin-11*(*n389*)	0	**0***	**0***	0	**0***	0	0	28
dhs-31	*lin-29*(*sy292*)	0	**0***	**0***	0	**0***	0	0	42
egl-26	+	0	61	30	0	61	30	0	37
egl-26	*lin-11*(*n389*)	0	**0***	**1†**	0	**0***	**1†**	0	39
lin-3	+	0	0	0	0	0	0	98	26
lin-3	*lin-11*(*n389*)	0	0	0	0	0	0	**7***	22
lin-3	*lin-29*(*n333*)	0	0	0	0	0	0	81	18
lin-11	+	41	92	90	81	100	0	0	38
lin-11	*lin-29*(*sy292*)	**5***	**43***	**59‡**	**30***	**58***	0	0	52
pepm-1	+	0	0	0	91	91	95	95	22
pepm-1	*lin-11*(*n389*)	0	0	0	**0***	**0***	**0***	**0***	50
pepm-1	*lin-29*(*sy292*)	0§	0§	0§	67§	67§	67§	67§	55
unc-53	+	0	0	0	100	0	0	0	many
unc-53	*lin-29*(*sy292*)	0	0	0	80	0	0	0	20

Percentages of cells that express *dhs-31::YFP*, *egl-26::GFP*, *lin-3::GFP*, *lin-11::GFP*, *pepm-1::GFP,* and *unc-53::GFP*. Boldface indicates *P* values are significantly different than wild type. *, $P < 0.0001$; †, $P = 0.0013$; ‡, $P = 0.0019$. §*pepm-1::GFP* is not detectable until the end of the fourth larval stage, at this point the vulva has already protruded in *lin-29(sy292)* mutants making it difficult to distinguish between the vulval cell types.

Table 2. Expression of mutated *zmp-1* enhancers in the vulva

Construct	vulA, % (*n*)	vulC, % (*n*)	vulD, % (*n*)	vulE, % (*n*)	vulF, % (*n*)
mk50-51	86 (28)	0 (35)	0 (35)	89 (35)	0 (35)
103/4Δ	**34* (29)**	0 (41)	0 (41)	**0† (41)**	0 (41)
105/6Δ	72 (32)	4 (46)	**22‡ (46)**	**30† (46)**	7 (46)
107/8Δ	**15† (27)**	0 (31)	0 (31)	**6† (31)**	0 (31)

Shown are percentage of cells that express the indicated constructs. Values in bold indicate *P* values that are significantly different than wild type. *, $P = 0.0001$; †, $P < 0.0001$; ‡, $P = 0.0040$

Table S1. Genes with detectable and dynamic expression in subsets of the cells that comprise the *C. elegans* vulva

Gene	Description
B0034.1	unfamiliar protein
bam-2	neurexin-related transmembrane protein
C55C3.5	unfamiliar protein
cdh-3	cadherin-related
ceh-2	*Drosophila empty spiracles* (*ems*)
cog-1	Nkx6.1/6.2
daf-6	Patched-related
dhs-31	predicted short chain-type dehydrogenase
egl-17	fibroblast growth factor (FGF)-related ligand
egl-26	H-box and NC domain
egl-38	Pax2/5/8
grd-5	*hedgehog*-related
grd-12	*hedgehog*-related
grl-4	*hedgehog*-related
grl-10	*hedgehog*-related
grl-15	*hedgehog*-related
grl-25	*hedgehog*-related
grl-31	*hedgehog*-related
lin-3	epidermal growth factor (EGF)
lin-11	LIM homeodomain
lin-29	C2H2 type zinc finger
lin-39	*Drosophila Sex combs reduced* (*Scr*)
nas-37	metalloprotease
nhr-67	*Drosophila tailless* (*tll*)
pax-2	Pax2/5/8
pepm-1	protein with M1 peptidase domain
sqv-4	UDP-glucose dehydrogenase-related
syg-2	Immunoglobulin superfamily transmembrane protein
unc-53	novel protein with SH3- and actin-binding sites
zmp-1	zinc metalloproteinase (MT4-MMP-related)

Table S2. Expression of *cdh-3*, *dhs-31*, and *unc-53* in the vulva

Reporter	RNAi	vulA	vulB1	vulB2	vulC	vulD	vulE	vulF	n
cdh-3	+	0	0	0	100	100	100	100	many
cdh-3	*nhr-67*	0	0	0	100	100	100	100	42
dhs-31	+	0	100	100	0	100	0	0	31
dhs-31	*nhr-67*	0	100	100	0	100	0	0	27
unc-53	+	0	0	0	84	20	0	0	39
unc-53	*nhr-67*	0	0	0	76	28	0	0	39

Percentages of cells that express *cdh-3::CFP, dhs-31::YFP,* and *unc-53::GFP*

Table S3: List of screened transcription factor RNAi clones in *ceh-2::YFP* screen

Gene	Coordinates	Gene	Coordinates	Gene	Coordinates
aha-1	I-5I09	C49C3.5	IV-8C12	*eat-1*	IV-4D13
ahr-1	I-4P19	C49D10.9	II-3H07	*egl-13*	X-3I21
ahr-1	I-4P21	C50B6.8	V-8F10	*egl-13*	X-3K13
arx-6	III-2B15	C52B9.2	X-2H15	*egl-18*	IV-1H19
asc-1	II-9K18	C54E10.5	V-12C10	*egl-27*	II-5A22
B0281.3	II-2O24	*ceh-10*	III-3E03	*egl-27*	II-5A24
B0336.13	III-3A19	*ceh-13*	III-4C05	*egl-38*	IV-5E04
B0379.4a	I-5E11	*ceh-14*	X-3N08	*egl-43*	II-7G07
B0496.7	IV-3P19	*ceh-16*	III-4H21	*egl-44*	II-4B10
B0496.8	IV-3P21	*ceh-17*	I-2A09	*egl-46*	V-4L02
brf-1	II-5F13	*ceh-18*	X-2O02	*egl-5*	III-4I05
C01B12.2	II-1A05	*ceh-2*	I-2HO8	*egl-9*	V-7A24
C02B8.4	X-4M03	*ceh-20*	III-3N09	*egr-1*	V-3C14
C02F12.5	X-2I14	*ceh-21*	X-1L01	*elt-1*	IV-5A05
C03D6.5	I-4F24	*ceh-22*	V-7I18	*elt-2*	X-5I11
C03G6.12	V-5E24	*ceh-23*	III-4I23	*elt-3*	X-6D13
C06E1.8	III-4H13	*ceh-24*	V-9G15	*elt-6*	IV-1J07
C07E3.6	II-7C14	*ceh-26*	III-4A18	*end-1*	V-9A08
C09F5.3	III-1M23	*ceh-28*	X-6K07	*eor-2*	X-6F06
C09G12.1	IV-1H14	*ceh-31*	X-2E23	F01G4.1	IV-5N06
C09H10.6	II-7L11	*ceh-32*	V-7N10	F10C1.5	II-4J17
C12D12.5	X-2C18	*ceh-33*	V-4M17	F16H9.2	X-6C21
C13C4.1	V-8K03	*ceh-34*	V-4M15	F17A2.3	X-5J12
C14C6.4	V-1C02	*ceh-35*	V-9O12	F17A9.6	V-4H19
C16A3.7	III-3K10	*ceh-36*	X-6L03	F19F10.1	V-5B09
C17A2.1	II-3D23	*ceh-37*	X-6L05	F19F10.5	V-5B17
C17H12.9	IV-3G22	*ceh-40*	X-5J16	F20D12.5	IV-4E09
C18B12.3	X-6P20	*ceh-41*	X-1J23	F22A3.1	X-3B11
C25E10.1	V-6I22	*ceh-43*	III-2A12	F22A3.5	X-3B19
C25G4.4	IV-6J23	*ceh-5*	I-4J02	F22D6.2	I-3O20
C26B2.4	IV-4G13	*ceh-6*	I-4M15	F22F1.3	X-4M15
C26C6.6	I-3P23	*ceh-7*	II-6G20	F25H5.1	I-4F03
C27C12.6	X-6L10	*ceh-8*	I-4E11	F28C6.2	II-6C06
C27D6.4	II-4A22	*ces-1*	I-4C10	F28F5.3	III-3H13
C28A5.4	III-2A12	*ces-2*	I-7M02	F29G9.4	V-4F03
C28H8.6	III-3I18	*che-1*	I-3K15	F33D11.1	I-2P03
C28H8.9	III-3I19	*cnd-1*	III-2B18	F36F12.8	V-2C13
C29H12.5	II-4H10	D1046.2	IV-4F21	F38C2.8	IV-7J16
C33A11.4	X-7G15	D1081.2	I-4M01	F38H12.3	V-3D21
C33G8.10	V-5G01	*daf-12*	X-5M07	F41D3.3	I-6I21
C33G8.12	V-5G05	*daf-12*	X-5M11	F41H10.4	IV-2F20
C33G8.8	V-5E21	*daf-14*	IV-5M16	F42G2.6	II-2F13
C34B2.4	I-5K14	*daf-16*	I-5M24	F42H10.4	III-4D19
C34B7.4	I-4G21	*daf-19*	II-7M11	F44A2.4	V-6F13
C34D1.5	V-8D10	*daf-3*	X-1M03	F44C4.2	V-4H24
C36C5.13	V-2J04	*dpl-1*	II-6H13	F44E7.8	V-4M16
C36F7.1	I-4B02	*dpy-20*	IV-6A20	F45C12.15	II-2E19
C42D8.4	X-2N18	*dpy-22*	X-4H18	F45H11.1	I-5O01
C44B9.4	III-5F02	*dro-1*	III-6J13	F47C10.1	V-3G04
C44F1.2	III-1J24	*eat-1*	IV-4D11	F47C10.3	V-3G08

Table S3 continued: List of screened transcription factor RNAi clones in *ceh-2::YFP* screen

Gene	Coordinates	Gene	Coordinates	Gene	Coordinates
F47C10.4	V-3G10	*hlh-8*	X-4M03	*mls-2*	X-2F16
F47C10.7	V-3G14	*hnd-1*	X-5J01	*mxl-1*	V-7N05
F47G9.4	V-7P21	K01H12.1	IV-5C17	*mxl-2*	III-1G23
F47H4.1	V-11J03	K03A11.5	X-6K11	*mxl-3*	X-6C16
F49E12.6	II-6K03	K04C1.2	X-6L15	*ncl-1*	III-4G07
F52C12.4	IV-1J05	K06A1.1	II-5A17	*ngn-1*	IV-8L18
F53B3.1	X-2A17	K06B9.5	IV-2K15	*nhl-2*	III-2F07
F53F8.1	V-13M01	K08F8.2	I-6I16	*nhl-3*	II-1E24
F53H1.4	IV-1O04	K11D2.4	I-6G22	*nhr-1*	X-7H18
F53H4.5	X-7C06	*let-381*	I-3C19	*nhr-10*	III-3D10
F55A3.3	I-5O08	*let-418*	V-4O16	*nhr-100*	IV-5E01
F55A3.3	I-5O14	*lim-4*	X-2O04	*nhr-103*	V-2I13
F56F11.3	III-1J15	*lim-7*	I-2L05	*nhr-104*	IV-2B09
F56F3.4	III-2A24	*lin-1*	IV-1N03	*nhr-105*	IV-3O12
F57B10.1	I-3M11	*lin-11*	I-5K09	*nhr-106*	V-1M23
F59E11.10	V-6G14	*lin-14*	X-5B19	*nhr-108*	V-10F13
F59E11.11	V-6G16	*lin-26*	II-5B12	*nhr-11*	IV-4J17
fax-1	X-2K03	*lin-29*	II-8G07	*nhr-111*	V-10B16
fkh-10	I-5G15	*lin-29*	II-8G15	*nhr-112*	V-11O01
fkh-2	X-2K08	*lin-31*	II-3J21	*nhr-113*	I-5B18
fkh-3 & 4	X-6E18	*lin-32*	X-1D20	*nhr-115*	V-2M13
fkh-4	X-6E20	*lin-35*	I-2N15	*nhr-116*	V-11E14
fkh-5	III-2O14	*lin-36*	III-4A10	*nhr-116*	V-11E16
fkh-6	II-3F08	*lin-39*	III-4A21	*nhr-118*	V-3N21
fkh-7	IV-2L18	*lin-41*	I-4J11	*nhr-119*	II-2F08
fkh-7	IV-2L20	*lin-42*	II-1D18	*nhr-119*	II-2F12
fkh-8	II-4H24	*lin-48*	III-1J18	*nhr-120*	X-3D21
ham-2	X-2P09	*lin-49*	IV-4I16	*nhr-121*	III-1H19
hbl-1	X-3O03	*lin-59*	I-1H05	*nhr-123*	V-1M07
hif-1	V-13O05	*lir-1*	II-5B14	*nhr-124*	V-3I08
hlh-1	II-3J04	*lir-2*	II-5B16	*nhr-125*	V-3L07
hlh-10	V-6F01	*lir-3*	II-7N15	*nhr-126*	V-2K01
hlh-11	III-5K17	*lpd-2*	I-2P22	*nhr-127*	V-10J08
hlh-12	IV-4G20	*ltd-1*	II-5P20	*nhr-128*	V-2I15
hlh-13	X-2H09	M01E5.6	I-6B22	*nhr-13*	V-2G19
hlh-14	II-4J11	*mab-18*	X-5I17	*nhr-13*	V-2I03
hlh-15	X-1J06	*mab-3*	II-6P10	*nhr-130*	V-1O07
hlh-16	I-4I02	*mab-5*	III-4I09	*nhr-132*	V-1O13
hlh-17	IV-7J16	*mab-9*	II-1C14	*nhr-134*	V-2I09
hlh-19	X-1E19	*mdl-1*	X-3H11	*nhr-135*	V-5I11
hlh-2	I-3D23	*mec-3*	IV-5F05	*nhr-136*	V-8K07
hlh-21	III-1G09	*med-1*	X-5N14	*nhr-137*	X-3F05
hlh-25	II-4B13	*mef-2*	I-4D05	*nhr-14*	X-4G22
hlh-26	II-4B15	*mes-6*	IV-4I10	*nhr-15*	V-1G09
hlh-27	II-4B19	*mes-6*	IV-4I12	*nhr-17*	X-6C05
hlh-28	X-7H22	*mgl-2*	I-5O07	*nhr-18*	V-2I11
hlh-29	X-7J02	*mix-1*	II-7D13	*nhr-19*	II-6J12
hlh-3	II-6F03	*mix-1*	II-7D19	*nhr-2*	I-2N21
hlh-4	III-5O11	*mls-1*	III-4A01	*nhr-21*	II-5F23
hlh-6	II-6J18	*mls-2*	X-2F14	*nhr-22*	II-5A23

Table S3 continued: List of screened transcription factor RNAi clones in *ceh-2::YFP* screen

Gene	Coordinates	Gene	Coordinates	Gene	Coordinates
nhr-23	I-3F11	*nhr-84*	V-10B23	*sop-3*	I-8G11
nhr-25	X-6I19	*nhr-85*	I-6J15	*sop-3*	I-8G19
nhr-3	X-6D16	*nhr-87*	IV-8F19	*sop-3*	I-8G23
nhr-31	IV-4G11	*nhr-88*	II-2J14	*sop-3*	I-8I11
nhr-35	X-2P13	*nhr-88*	II-2J16	*sox-2*	X-3L06
nhr-38	IV-5C21	*nhr-88*	II-2J20	*sur-2*	I-7G14
nhr-4	IV-5K05	*nhr-89*	I-6E04	T01D3.2	V-9C07
nhr-41	IV-1B01	*nhr-9*	III-3B02	T01G6.6	V-1O03
nhr-42	V-5E17	*nhr-90*	V-1E12	T03E6.3	V-11G23
nhr-43	IV-6G18	*nhr-91*	X-7E21	T04C10.4	X-6J22
nhr-44	V-6E18	*nhr-91*	X-7E23	T05C1.4	II-3H22
nhr-45	X-2D08	*nhr-92*	IV-8F23	T05G11.1	V-10F17
nhr-46	IV-2P24	*nhr-94*	V-5J01	T05G11.1	V-10F19
nhr-47	V-4C18	*nhr-95*	V-12I08	T06A10.4	IV-8E05
nhr-49	I-4N14	*nhr-96*	V-2K03	T08D10.1	X-5P15
nhr-50	V-10J23	*nhr-97*	IV-5E11	T09D3.4	V-4O03
nhr-51	V-10E18	*nhr-98*	V-1M09	T10B11.9	I-3M02
nhr-54	V-11E13	*nob-1*	III-6O01	T13C5.4	X-3I10
nhr-55	V-1O05	*oma-2*	V-5D12	T21B10.5	II-6O20
nhr-56	V-2I17	*osm-1*	X-7M24	T22C8.5	II-6E08
nhr-57	V-3F03	*pal-1*	III-2O04	T23B12.1	V-6C23
nhr-58	V-1O15	*pax-3*	II-7M03	T24H10.2	II-6F13
nhr-59	V-2M05	*pes-1*	IV-2H21	*tab-1*	II-5A12
nhr-59	V-2M07	*pha-4*	V-13O01	*taf-10*	V-3D22
nhr-6	III-1F16	*php-3*	III-6M23	*taf-11.2*	I-3B21
nhr-60	V-10K18	*pie-1*	III-6E08	*taf-3*	X-7K02
nhr-61	II-9E18	*pop-1*	I-1K04	*taf-4*	I-1E17
nhr-62	I-4N02	*pos-1*	V-6A23	*taf-5*	I-3J16
nhr-62	I-4N04	R03E1.3	X-6J19	*taf-6.1*	II-1B18
nhr-63	V-10J21	R04A9.5	X-1A23	*taf-9*	III-6D24
nhr-64	I-1M14	R05D11.1	I-4O23	*tbp-1*	III-3N22
nhr-65	V-12C18	R06C7.9	I-3H13	*tbx-11*	III-3O01
nhr-66	IV-4M05	R06F6.6	II-7D05	*tbx-18*	II-3K10
nhr-67	IV-5P08	R07E5.3	III-2O05	*tbx-2*	III-2N13
nhr-68	V-7J14	R11G11.12	V-1A12	*tbx-30*	IV-8B10
nhr-69	I-4N18	R11H6.5	V-9H21	*tbx-31*	X-1F03
nhr-7	IV-6C03	R13D11.8	V-1B09	*tbx-32*	X-1D11
nhr-70	V-12M01	*ref-1*	II-7O13	*tbx-32*	X-1D17
nhr-71	X-6O14	*ref-2*	X-4C05	*tbx-33*	III-6C21
nhr-72	II-3F13	*rpm-1*	V-6C14	*tbx-33*	III-6J06
nhr-74	I-5P17	*sdc-1*	X-7I12	*tbx-35*	II-4O24
nhr-75	II-3H01	*sdc-2*	X-5D07	*tbx-36*	IV-6I18
nhr-76	IV-1A08	*sdc-3*	V-9F23	*tbx-38*	III-6G07
nhr-77	I-6A16	*sel-8*	III-1F20	*tbx-41*	X-1J15
nhr-78	IV-2O03	*sem-4*	I-3O02	*tbx-8*	III-5F01
nhr-8	IV-3J14	*sex-1*	X-5C07	*tbx-9*	III-5F09
nhr-80	III-1A07	*skn-1*	IV-2N18	*tlf-1*	I-4E14
nhr-81	I-6O05	*sop-2*	II-8A20	*tlp-1*	IV-7C16
nhr-82	I-6I17	*sop-2*	II-8E04	*tra-1*	III-5N20
nhr-83	V-1G08	*sop-3*	I-8E05	*ttx-3*	X-5L01

Table S3 continued: List of screened transcription factor RNAi clones in *ceh-2::YFP* screen

Gene	Coordinates	Gene	Coordinates	
unc-115	X-5A09	ZK337.2	I-7M08	
unc-130	II-7B04	ZK381.5	IV-3M22	
unc-3	X-6J06	ZK455.6	X-5O14	
unc-3	X-6J10	ZK488.4	V-1E14	
unc-30	IV-7A23	ZK697.2	V-1J06	
unc-37	I-3E10	ZK856.9	V-7G21	
unc-4	II-7E11	ZK945.5	II-7K13	
unc-42	V-6H06	*zyx-1*	II-8H13	
unc-55	I-3L14			
unc-55	I-3L18			
unc-62	V-3B18			
unc-86	III-4I08			
unc-97	X-3I11			
unc-98	X-3P06			
vab-15	X-4J02			
vab-3	X-5I17			
vab-7	III-5F10			
W02C12.3	IV-2C11			
W03H1.2	X-1H15			
Y113G7B.14	V-12N02			
Y116A8C.18	IV-8G13			
Y17D7A.1	V-12C14			
Y17D7B.1	V-12A16			
Y1A5A.1	III-2A03			
Y22F5A.1	V-7I23			
Y37E11B.1	IV-1J24			
Y38E10A.6	II-8M18			
Y39A3CR.6	III-7I11			
Y39A3CR.6	III-7I13			
Y40B1A.4	I-6F02			
Y51H4A.12	IV-7P18			
Y53C12C.1	II-7A15			
Y53H1A.2	I-5J09			
Y55F3AM.7	IV-8J23			
Y55F3AM.7	IV-8L01			
Y57G11A.3	IV-7D23			
Y5F2A.4	IV-5J12			
Y65B4A.7	I-7P14			
Y6G8.3	V-11D16			
Y75B8A.6	III-6O09			
Y77E11A.6	IV-1D09			
Y80D3A.3	V-12E10			
zag-1	IV-1P04			
ZC123.3	I-1I01			
ZC123.3	I-1I05			
ZC123.3	I-1I07			
ZC328.2	I-3E21			
ZK1193.5	X-1C09			
ZK1240.1	II-2B09			
ZK1240.3	II-2B13			

Table S4. List of screened transcription factor RNAi clones in *zmp-1::GFP* screen

Gene name	Gene name	Gene name	Gene name	Gene name	Gene name
2L52.1	C07A12.1	C18B12.3	C33H5.17	C50E10.4	F16B12.6
AH6.5	C07A12.3	C18D1.1	C34B2.4	C50F4.12	F16B12.8
B0019.2	C07A12.5	C24G6.4	C34B7.1	C52B9.2	F16B4.9
B0035.1	C07E3.5	C24H11.3	C34C6.8	C53D5.4	F16H11.5
B0250.4	C07E3.6	C25A1.11	C34D1.1	C54C8.1	F16H9.2
B0261.1	C07H6.7	C25A1.2	C34D1.2	C54E10.5	F17A2.3
B0280.4	C08B11.3	C25B8.6	C34D1.5	C54F6.8	F17A2.5
B0280.8	C08C3.1	C25D7.3	C34E10.7	C54F6.9	F17A9.3
B0281.3	C08C3.3	C25E10.1	C34H3.2	C55B7.12	F17A9.6
B0286.5	C08F8.8	C25G4.4	C35C5.1	C56C10.8	F18A1.2
B0304.1	C08G9.2	C25H3.6	C35D10.16	C56E10.1	F18A1.3
B0310.2	C09F5.3	C26B2.3	C35D6.4	C56E10.4	F18A1.4
B0336.13	C09G12.1	C26B2.4	C36C5.13	D1005.3	F19B10.9
B0336.3	C09G4.5	C26C6.1	C36C9.2	D1007.1	F19B2.6
B0336.7	C09G9.6	C26C6.6	C36E8.1	D1014.8	F19F10.1
B0379.4a	C09H10.6	C26D10.5	C36F7.1	D1046.2	F19F10.5
B0412.1	C09H6.1	C26E1.3	C37A2.5	D1081.2	F20D12.5
B0414.2	C10A4.8	C26E6.2	C37E2.4	D2030.7	F21A10.2
B0496.7	C10G8.6	C27A12.2	C37E2.5	DY3.3	F21D12.1
B0496.8	C10G8.7	C27A12.3	C37F5.1	E02H1.7	F21D12.5
B0564.10	C11G6.1	C27A12.5	C38D4.3	E02H9.8	F21D5.9
C01B12.2	C12C8.3	C27B7.4	C38D4.6	E03H4.13	F21G4.5
C01B7.1	C12D12.4	C27C12.2	C38D4.7	E03H4.6	F21H11.3
C01B7.6	C12D12.5	C27C12.6	C39E6.4	F01D4.6	F22A3.1
C01F1.1	C12D5.2	C27C7.3	C40H5.5	F01G10.8	F22A3.5
C01F6.9	C12D5.8	C27D6.4	C41G7.5	F01G4.1	F22D6.2
C01G12.1	C12D8.1	C28A5.4	C41G7.5	F02E9.4	F22E12.4
C01H6.5	C13C4.1	C28C12.8	C42D8.4	F08C6.7	F22F1.3
C02B4.2	C13C4.2	C28D4.1	C43H6.7	F09B12.2	F23A7.6
C02B8.4	C13C4.3	C28G1.4	C43H6.8	F09B9.2	F23B12.7
C02F12.5	C13G5.1	C28H8.6	C44B9.4	F09C6.8	F23C8.4
C02F5.12	C14B9.6	C28H8.9	C44C10.8	F09C6.9	F23F1.1
C03D6.5	C14C6.4	C29E6.5	C44F1.2	F09F3.10	F23F12.9
C03G6.10	C15H11.8	C29F7.4	C44H4.7	F09G2.9	F23H11.1
C03G6.12	C16A3.4	C29F7.5	C45E1.1	F10C1.5	F23H11.5
C03G6.8	C16A3.7	C29G2.5	C45E5.6	F11A1.1	F25D7.3
C04A2.3	C16C2.1	C29H12.5	C46E10.8	F11A1.3	F25E2.5
C04F1.3	C17A2.1	C30A5.7	C46E10.9	F11A10.2	F25E5.6
C04F5.5	C17A2.8	C32A3.1	C46F11.3	F11C1.6	F25H5.1
C04F5.9	C17C3.10	C32F10.2	C47C12.3	F13A2.8	F26A1.2
C04G2.7	C17C3.7	C32F10.5	C47F8.2	F13D11.2	F26A10.2
C05D10.1	C17C3.8	C32F10.6	C47F8.8	F13G11.1	F26B1.7
C05G6.1	C17E4.6	C33A11.4	C47G2.2	F13G3.1	F26C11.2
C05G6.2	C17E7.1	C33D12.1	C48D5.1	F13H6.1	F26D12.1
C06B8.1	C17E7.5	C33D3.1	C48E7.3	F14A5.1	F26F12.7
C06C6.4	C17E7.6	C33G8.10	C49C3.5	F14D12.2	F26F4.7
C06C6.5	C17E7.7	C33G8.12	C49D10.6	F14F3.1	F26F4.8
C06E1.8	C17E7.8	C33G8.6	C49D10.9	F14H3.11	F27D4.6
C06E2.1	C17H12.9	C33G8.8	C49F5.4	F15C11.1	F27E5.2
C06G3.1	C18A3.8	C33G8.9	C50B6.8	F15E6.1	F28B12.2

Table S4 continued. List of screened transcription factor RNAi clones in *zmp-1::GFP* screen

Gene name	Gene name	Gene name	Gene name	Gene name	Gene name
F28C6.1	F40G9.11	F47H4.1	F58A4.7	K10D3.3	R11E3.5
F28C6.2	F40G9.14	F48D6.3	F58E10.2	K10G6.1	R11G11.1
F28F5.3	F40G9.4	F48G7.11	F58E10.5	K10G6.3	R11G11.12
F28F9.1	F40H3.4	F48G7.3	F58E6.10	K11D2.4	R11G11.2
F28H6.2	F40H6.4	F49E10.5	F58H1.2	K11E4.5	R11H6.5
F29F11.5	F41B5.10	F49E12.6	F59B1.7	K11G9.4	R13.1
F29G9.4	F41B5.9	F49E8.2	F59B10.1	K11H12.8	R13.2
F30F8.8	F41D3.1	F49H12.1	F59E11.10	K12H4.1	R13A5.5
F31A3.2	F41D3.3	F52B5.5	F59E11.11	K12H6.1	R13D11.8
F31A3.4	F41H10.4	F52C12.4	H01A20.1	K12H6.12	R13F6.5
F31E3.2	F42A9.2	F52C12.5	H10E21.3	M01A10.1	R13H8.1
F31E8.3	F42G2.6	F52E1.1	H12C20.3	M01E11.5	R144.3
F31F4.12	F42G4.3	F52E10.1	H14A12.4	M01E5.6	R53.3
F31F6.1	F42H10.4	F52E4.8	H21P03.1	M02H5.3	R74.3
F31F7.3	F43D9.5	F52F12.4	H22D14.1	M02H5.6	T01B10.4
F32A5.1	F43G9.11	F53A2.5	H25P06.2	M02H5.7	T01C1.3
F32B6.1	F44A2.4	F53B2.6	H27C11.1	M03D4.4	T01D3.2
F32H2.1	F44A6.2	F53B3.1	K01G5.1	M04B2.1	T01E8.2
F33A8.3	F44B9.6	F53F8.1	K01H12.1	M04B2.3	T01G6.4
F33D11.1	F44C4.2	F53G12.5	K01H12.3	M04G12.4	T01G6.6
F33D4.1	F44C8.10	F53H1.4	K02B12.1	M05B5.5	T01G6.7
F33E11.1	F44C8.11	F53H4.5	K02B9.4	M106.1	T01G6.8
F33E11.2	F44C8.2	F54A5.1	K02C4.4	M142.4	T03E6.3
F33H1.1	F44C8.3	F54D1.4	K02F3.4	R02C2.4	T03F6.2
F33H1.4	F44C8.4	F54D5.11	K02F3.8	R02D1.1	T04C10.4
F34D10.5	F44C8.5	F54F2.5	K02G10.1	R02D3.7	T05A7.4
F35E8.12	F44C8.6	F54F2.9	K03A11.3	R02E12.4	T05B4.2
F35H8.3	F44C8.9	F54F7.1	K03A11.5	R03D7.4	T05C1.4
F36A4.14	F44D12.10	F55A3.3	K03B4.3	R03E1.3	T05G11.1
F36D1.1	F44E2.7	F55A3.7	K04C1.2	R03E9.1	T05G5.2
F36D3.2	F44E7.8	F55A8.1	K04C1.3	R04A9.5	T06A10.4
F36F12.8	F44G3.9	F55B11.4	K06A1.1	R04B5.3	T06C12.13
F37A4.8	F45C12.15	F55B12.1	K06A1.4	R05D11.1	T06C12.7
F37B4.10	F45E12.2	F55D12.3	K06B4.1	R05D11.1	T07C12.11
F37D6.6	F45E4.9	F55D12.4	K06B4.10	R05D3.3	T07C4.2
F37H8.1	F45E6.2	F55D12.4	K06B4.5	R06C7.9	T07C4.6
F38A5.13	F45H11.1	F56A12.1	K06B4.6	R06F6.6	T07D10.3
F38A6.1	F45H11.4	F56E3.4	K06B4.7	R07B1.1	T07F8.4
F38A6.3	F46C8.5	F56F11.3	K06B9.5	R07B7.13	T07G12.10
F38B7.1	F46F11.2	F56F3.1	K07C11.1	R07B7.14	T07G12.11
F38C2.2	F46G10.6	F56F3.4	K08A2.5	R07B7.15	T07G12.12
F38C2.5	F47A4.2	F56H1.2	K08A8.2	R07C12.4	T08D10.1
F38C2.7	F47C10.1	F57A10.5	K08B4.1	R07E5.3	T08G5.7
F38C2.8	F47C10.3	F57A8.1	K08B5.2	R07H5.10	T09A12.4
F38H12.3	F47C10.4	F57A8.5	K08E4.1	R08H2.9	T09D3.4
F39B2.4	F47C10.7	F57B10.1	K08F8.2	R09G11.2	T09F3.1
F39H11.2	F47D12.4	F57C12.3	K08F8.2	R10D12.2	T10B11.9
F39H11.3	F47E1.3	F57C9.4	K09H9.7	R10E11.1	T10D4.6
F40E10.2	F47F6.1	F57G4.6	K10C3.6	R10E4.11	T10G3.5
F40F8.7	F47G9.4	F57G8.6	K10D2.3	R119.6	T11A5.1

Table S4 continued. List of screened transcription factor RNAi clones in *zmp-1::GFP* screen

Gene name	Gene name	Gene name	Gene name	Gene name	
T11B7.4	T27A8.2	Y16B4A.1	Y53G8AR.9	ZK1025.9	
T12D8.7	T27B1.1	Y17D7A.1	Y53H1A.2	ZK1037.4	
T12F5.4	T27B1.2	Y17D7A.3	Y54E10BR.8	ZK1037.5	
T13C5.4	T27B7.1	Y17D7B.1	Y55F3AM.7	ZK112.2	
T13F3.2	T27B7.2	Y17G7A.1	Y56A3A.18	ZK1128.4	
T13F3.3	T27B7.3	Y18D10A.1	Y56A3A.28	ZK1193.5	
T14F9.4	T27B7.4	Y1A5A.1	Y56A3A.4	ZK1240.1	
T14F9.5	T27B7.5	Y22D7AL.16	Y57A10A.31	ZK1240.3	
T14G12.4	T27C4.4	Y22F5A.1	Y57G11A.3	ZK177.10	
T15D6.6	T27F2.4	Y32B12B.6	Y57G11C.25	ZK185.1	
T15H9.3	T28D9.9	Y37E11B.1	Y59E9AR.3	ZK218.6	
T16H12.4	T28F12.2	Y37E11B.2	Y5F2A.4	ZK265.4	
T18D3.7	T28H11.4	Y38E10A.1	Y5H2B.2	ZK337.2	
T19A5.4	VC5.5	Y38E10A.6	Y60A9.3	ZK370.2	
T19A5.5	W01D2.2	Y38F2AR.13	Y62E10A.17	ZK380.1	
T19B10.11	W02A2.7	Y38H8A.5	Y65B4A.7	ZK381.5	
T19E7.2	W02C12.3	Y39A1C.3	Y66A7A.5	ZK418.1	
T19H12.8	W02D3.9	Y39A3CR.6	Y66A7A.8	ZK418.9	
T20B12.2	W02D7.6	Y39B6A.12	Y67A6A.2	ZK455.6	
T20B12.6	W02D9.3	Y39B6A.17	Y67A6A.2	ZK488.1	
T20B12.8	W02H5.7	Y39B6A.36	Y67D8A.3	ZK488.2	
T20F7.1	W03A3.1	Y39B6A.47	Y69A2AR.26	ZK488.4	
T20H4.2	W03C9.4	Y40B1A.4	Y69A2AR.29	ZK6.1	
T21B10.5	W03D2.1	Y41D4B.7	Y6G8.3	ZK6.2	
T22B3.1	W03F9.2	Y41D4B.8	Y70C5C.6	ZK6.5	
T22B7.1	W04A8.7	Y43F8C.10	Y71F9B.10	ZK652.5	
T22C8.3	W04D2.4	Y44E3B.1	Y71H2AM.17	ZK652.6	
T22C8.4	W04H10.3	Y46H3D.6	Y75B8A.1	ZK682.4	
T22C8.5	W05B5.3	Y47D3A.6	Y75B8A.2	ZK697.2	
T22E7.2	W05E10.3	Y47D3B.7	Y75B8A.29	ZK829.5	
T22H6.6	W05H7.4	Y47D3B.9	Y75B8A.6	ZK856.13	
T23B12.1	W06H12.1	Y48A6C.1	Y77E11A_3443.a	ZK856.9	
T23G4.1	W09B6.2	Y48A6C.3	Y77E11A.6	ZK892.7	
T23G5.6	W09C2.1	Y48A6C.5	Y80D3A.3	ZK909.4	
T23G7.1	W10C8.2	Y48B6A.14	Y82E9BR.1	ZK945.5	
T23H4.2	W10D5.1	Y48G8AR.1	Y82E9BR.15	ZK993.1	
T24A6.8	W10D5.3	Y48G9A.11	Y95B8A.7		
T24B8.6	Y104H12A.1	Y49E10.14	ZC123.3		
T24C4.7	Y105E8A.17	Y51A2D.17	ZC132.2		
T24D3.1	Y105E8C.s	Y51B9A.5	ZC204.2		
T24H10.1	Y106G6H.4	Y51H1A.5	ZC247.3		
T24H10.2	Y113G7B.14	Y51H4A.12	ZC328.2		
T25C12.1	Y116A8C.17	Y51H4A.17	ZC376.4		
T26A5.8	Y116A8C.18	Y51H4A.4	ZC395.8		
T26A8.4	Y116A8C.20	Y53C10A.12	ZC410.1		
T26C11.1	Y116A8C.22	Y53C10A.3	ZC416.1		
T26C11.5	Y116A8C.35	Y53C12B.5a	ZC513.6		
T26C11.6	Y11D7A.12	Y53C12C.1	ZC64.3		
T26E4.8	Y11D7A.13	Y53F4B.3	ZC64.4		
T27A1.6	Y15E3A.1	Y53F4B.5	ZC8.4		

Appendix I:

The *Caenorhabditis elegans* vulva: A post-embryonic gene regulatory network controlling organogenesis

Ted O. Ririe, Jolene S. Fernandes, and Paul W. Sternberg

Published in *PNAS*, **105**(51):20095-20099 (2008).

The *Caenorhabditis elegans* vulva: A post-embryonic gene regulatory network controlling organogenesis

Ted O. Ririe, Jolene S. Fernandes, and Paul W. Sternberg[1]

Howard Hughes Medical Institute and Division of Biology, California Institute of Technology, Pasadena, CA 91125

Edited by Michael S. Levine, University of California, Berkeley, CA, and approved August 25, 2008 (received for review July 10, 2008)

The *Caenorhabditis elegans* vulva is an elegant model for dissecting a gene regulatory network (GRN) that directs postembryonic organogenesis. The mature vulva comprises seven cell types (vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF), each with its own unique pattern of spatial and temporal gene expression. The mechanisms that specify these cell types in a precise spatial pattern are not well understood. Using reverse genetic screens, we identified novel components of the vulval GRN, including *nhr-113* in vulA. Several transcription factors (*lin-11*, *lin-29*, *cog-1*, *egl-38*, and *nhr-67*) interact with each other and act in concert to regulate target gene expression in the diverse vulval cell types. For example, *egl-38* (Pax2/5/8) stabilizes the vulF fate by positively regulating vulF characteristics and by inhibiting characteristics associated with the neighboring vulE cells. *nhr-67* and *egl-38* regulate *cog-1*, helping restrict its expression to vulE. Computational approaches have been successfully used to identify functional *cis*-regulatory motifs in the *zmp-1* (zinc metalloproteinase) promoter. These results provide an overview of the regulatory network architecture for each vulval cell type.

genetic regulatory networks | nematode | transcriptional regulation

Complex interactions of signaling molecules, transcription factors, and effector genes direct spatial and temporal patterning during organogenesis (1). The differentiation and morphogenesis of the *Caenorhabditis elegans* vulva is useful for studying the gene regulatory network (GRN) of larval stage organogenesis due to its invariant cell lineage, its amenability to genetic manipulation, and the availability of reporter genes with many spatial and temporal expression patterns in the seven vulval cell types (2). In ref. 3, we described a regulatory network of interactions between a set of evolutionarily conserved transcription factors and an array of genes expressed in the differentiated cells of the *C. elegans* vulva. Here, we briefly review vulval development and aspects of the provisional GRN directing its organogenesis. We then describe additional pairwise *trans*-regulatory interactions, including the results of RNAi screens and a *cis* regulatory analysis of *zmp-1* that together help refine our network model, and infer common network themes, such as boundary formation, combinatorial control, and stable feedback loops. These additional data support the hypothesis that overall network architecture is unique for each of the vulval cell types.

The life cycle of *C. elegans* consists of four larval stages (L1–4) and an adult stage, with each stage separated by a molt (2). The *C. elegans* vulva is derived postembryonically from six vulval precursor cells (VPCs) termed P3.p-P8.p. All VPCs are competent to receive an inductive signal from a specialized somatic gonadal cell, the anchor cell (AC), during the L2 stage. P6.p, which is closest to the AC, is induced to generate the 1° vulval lineage, producing the inner cells of the vulva. The P5.p and P7.p cells generate 2° vulval lineages, producing the outer cells of the vulva. P3.p, P4.p, and P8.p are uninduced and adopt the 3° fate, and fuse to the hypodermal syncytium hyp7. The L4 stage vulva comprises 22 differentiated cells that are descendents of P5.p, P6.p, and P7.p, and that are of seven different types: vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF [supporting information (SI) Fig. S1]. EGF, Notch, and Wnt signaling pathways specify which VPCs generate the 1° and 2° lineages, but we are just now identifying the network of transcription factors that control cell-fate differentiation in the seven vulval cell types (2).

The *C. elegans* vulva allows passage of sperm and eggs by connecting the uterus to the outside environment (2). Each vulval cell type has specialized roles that contribute to vulval function and morphology. For example, vulF cells, the innermost vulval cells, contact the AC and the uterus; they are the target for AC invasion, thus creating the vulval-uterine connection required for egg-laying (4). Vulval muscles that regulate egg-laying connect to the vulva between vulC and vulD, and the outermost portion of the vulva comprises the vulA cells, which attach the vulva to the hypodermis (2). The unique patterns of gene expression in each of the vulval cell types are likely responsible for their individual properties. Comparison of the vulval GRN to those in other organisms is necessary for expanding our knowledge of organ development.

Results and Discussion

Functional Roles of Gene Expression During Vulval Differentiation. Much is known about the signaling network that establishes the pattern of vulval cell differentiation, but our understanding of the GRN that specifies the terminal seven vulval cell types is limited (2). Five transcription factors (*lin-11*, *lin-29*, *cog-1*, *egl-38*, and *nhr-67*) are major regulators of cell-fate determination and morphogenesis in the vulva. *lin-11* encodes a LIM homeodomain protein, consisting of a homeodomain and two specialized LIM-type zinc-fingers (5). LIM homeodomain family members play roles in differentiation and pattern formation in arthropods and vertebrates (6, 7). *lin-29*, a C2H2-type zinc finger, plays a role in many events occurring at the larva to adult transition, including terminal differentiation of the seam cells (8), morphogenesis (9), and formation of the vulval-uterine-seam cell connection (10). *cog-1* encodes a Nkx6.1/6.2 homeoprotein transcription factor (11); vertebrate Nkx6.1 proteins are involved in neuronal and pancreatic endocrine cell formation (12, 13). *egl-38* encodes a Pax2/5/8 protein, which are known to be involved in organogenesis; e.g., mouse *Pax2* mediates nephrogenesis (14, 15). *nhr-67* is an ortholog of *Drosophila melanogaster tailless* (*tll*) (16), a conserved nuclear hormone receptor necessary for *Drosophila* embryogenesis and neuronal development (17). The

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Gene Networks in Animal Development and Evolution," held February 15–16, 2008, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS web site at http://www.nasonline.org/SACKLER_Gene_Networks.

Author contributions: T.O.R., J.S.F., and P.W.S. designed research; T.O.R. and J.S.F. performed research; T.O.R., J.S.F., and P.W.S. analyzed data; and T.O.R., J.S.F., and P.W.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: pws@caltech.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806377105/DCSupplemental.

Table 1. Expression of *dhs-31*, *egl-26*, *lin-3*, *lin-11*, *pepm-1* and *unc-53* in the vulva.

Reporter	Mutation	vulA	vulB1	vulB2	vulC	vulD	vulE	vulF	*n*
dhs-31	+	0	100	100	0	100	0	0	31
dhs-31	*lin-11*(*n389*)	0	**0***	**0***	0	**0***	0	0	28
dhs-31	*lin-29*(*sy292*)	0	**0***	**0***	0	**0***	0	0	42
egl-26	+	0	61	30	0	61	30	0	37
egl-26	*lin-11*(*n389*)	0	**0***	**1***	0	**0***	**1***	0	39
lin-3	+	0	0	0	0	0	0	98	26
lin-3	*lin-11*(*n389*)	0	0	0	0	0	0	**7***	22
lin-3	*lin-29*(*n333*)	0	0	0	0	0	0	81	18
lin-11	+	41	92	90	81	100	0	0	38
lin-11	*lin-29*(*sy292*)	**5***	**43***	**59†**	**30***	**58***	0	0	52
pepm-1	+	0	0	0	91	91	95	95	22
pepm-1	*lin-11*(*n389*)	0	0	0	**0***	**0***	**0***	**0***	50
pepm-1	*lin-29*(*sy292*)	0‡	0‡	0‡	67‡	67‡	67‡	67‡	55
unc-53	+	0	0	0	100	0	0	0	Many
unc-53	*lin-29*(*sy292*)	0	0	0	80	0	0	0	20

Percentages of cells that express *dhs-31*::*YFP*, *egl-26*::*GFP*, *lin-3*::*GFP*, *lin-11*::*GFP*, *pepm-1*::*GFP* and *unc-53*::*GFP*. Boldface indicates *P* values are significantly different than wild type. *, $P = 0.000$; †, $P = 0.002$.
‡*pepm-1*::*GFP* is not detectable until the end of the fourth larval stage, at which point the vulva has already protruded in *lin-29(sy292)* mutants, making it difficult to distinguish between the vulval cell types.

pattern of vulval cell types is specified by the differential interactions of the transcription factors that operate within each cell (2).

We have identified 30 genes dynamically expressed in specific subsets of the cells of the mature *C. elegans* vulva (Table S1 has a complete list). These genes encode transcription factors, guidance cues, proteases, structural proteins, signaling molecules, and novel proteins with unknown function. The physiological relevance is known for several genes. For example, *egl-17*, which encodes a fibroblast growth factor (FGF), is necessary for migration of the sex myoblasts to the vulva (18). Induction of the uterine uv1 cells depends on the epidermal growth factor (EGF) family member *lin-3* (19). *bam-2* (neurexin-related transmembrane protein) and *syg-2* (transmembrane Ig superfamily protein) are required for vulval innervation (20, 21). *sqv-4*, which encodes an UDP-glucose dehydrogenase-related protein, is involved in the structural integrity and morphology of the vulva (22). The significance of some of the genes expressed in the vulva is not yet known, including the *Drosophila empty spiracles* (*ems*) homolog *ceh-2* (23), a cadherin-related protein encoded by *cdh-3* (24), and *zmp-1*, which encodes a MT4-MMP-related zinc metalloproteinase (23, 25). Genetic perturbations that result in altered expression patterns of these effector genes are helpful in elucidating the regulatory network.

Several additional genes with detectable expression in the mature vulva, including two putative transcription factors, have been identified since we last described the vulval GRN (3). The Pax2/5/8 gene *pax-2* is expressed exclusively in the vulD cells (16). *pax-2* is the result of a recent duplication of *egl-38* (26). *egl-38* has been identified as a regulator of cell-fate specification in the *C. elegans* vulva (3, 15, 19). *lin-39* encodes a Hox protein that is an ortholog of *Drosophila Sex combs reduced* (*Scr*) (27, 28). During the late L3 stage, *lin-39* expression increases in the vulA precursor cells; this expression persists in vulA until late L4 (29). The Patched-related protein DAF-6 is expressed in vulE and vulF (30). Hao *et al.* (31) reported the expression of seven *hedgehog*-related genes in the vulva: *grd-5* in vulB and vulD, *grd-12* in vulC, *grl-4* in vulA and vulB, *grl-10* in vulA and vulB, *grl-15* in vulB, vulC, vulD and vulE, *grl-25* in vulA, and *grl-31* in vulF. Last, *nas-37* (32), which encodes a metalloprotease, is expressed in vulB (data not shown).

Conserved Regulatory Strategies in the Vulval GRN. In ref. 3, we described how the expression patterns of a subset of reporter constructs are affected in transcription factor mutant backgrounds. We have increased the number of known interactions more than twofold, from 15 to 36, and identified many of the regulatory relationships among the transcription factors (Table 1, Fig. S2, and Table S2). This network includes strategies that are shared by other GRNs. For example, the cell-type specific expression of *cog-1* appears to be restricted to the vulC and vulD cells by a variety of mutual and autoregulatory controls (16). In a second example, *lin-11* is necessary for vulA-specific expression of *nhr-67* (16). In turn, *nhr-67* is necessary for the expression of the vulA effector gene *zmp-1*. The differentiated state of vulA may be further stabilized, because *nhr-67* is positively autoregulated in vulA. This is an instance where multiple positive inputs, including feedback loops, ensure the maintenance of a terminal cell fate.

Other network strategies that are present in the vulva include combinatorial control circuits. *egl-17* expression in vulF is perturbed by neither *cog-1* nor *egl-38* mutations (16). However, *egl-17* expression is derepressed in vulF in *cog-1; egl-38* double mutants. This redundancy could ensure proper execution of cell fate. Finally, negative autoregulation, as in *cog-1* in vulA, vulB, vulE, and vulF and *nhr-67* in vulC, vulD, vulE, and vulF, appears to be a fundamental strategy used in the vulval GRN (16). Negative autoregulation has been found to accelerate gene circuit response time and assists in making quick cell fate decisions (33).

LIN-29 and LIN-11 Interact to Determine Vulval Cell Fate. By examining the effects of specific transcription factor mutations on a wider array of genes expressed in the mature vulva, we have found increased complexity in some of their roles. We had hypothesized that *lin-29* is a temporal regulator of gene expression, because it is required for vulval gene expression at the mid to late L4 stage (*egl-17* in vulC and vulD, *ceh-2* in vulC, and *zmp-1* in vulD and vulE) (3, 23), and functions in several developmental processes at the L4 stage to adult transition (8–10, 34). However, in a *lin-29* mutant background the mid-L4 expression of *lin-3* in vulF cells is not abolished, whereas expression of *dhs-31*, which is initiated in gravid adults, is abolished (Table 1 and Fig. S2).

LIN-29 is necessary for wild-type levels of *lin-11* transgene expression (Table 1) and is thus a key regulator of *lin-11*. In addition to genes described in ref. 3, *lin-11* is required for *dhs-31*, *egl-26*, *lin-3* and *pepm-1* expression (Table 1). Thus, of the known vulval transcriptional regulators, *lin-11* exhibits the broadest effect (3). *egl-17* during L3 and *cdh-3* in vulF are the only vulva

expression patterns not abolished in a *lin-11* background. Reporter constructs are powerful tools for identifying genes' spatial and temporal expression patterns; however, reporter gene constructs might well lack relevant regulatory motifs. Consequently, we might be missing some regulatory connections.

The interplay between *lin-11*, *lin-29* and their downstream targets is complex. *lin-11*::*GFP* expression is not abolished in a *lin-29* mutant background, and several of the gene expression patterns perturbed by loss of *lin-11* are not affected in a *lin-29* deficient background (Fig. S2). LIN-29 might act in concert with another, as yet unidentified, factor to ensure the proper temporal and spatial expression of the general cell-fate regulator LIN-11. Considering LIN-29's role as a regulator of developmental timing, it could serve as a temporal input, whereas another factor serves as the spatial input.

Identification of Additional Components of the Vulval GRN. Transcription factors identified by forward genetic screens are biased by ascertainment toward those with strong, nonredundant effects, and those that affect multiple aspects of vulval development. Two categories of transcription factors are relatively more difficult to identify in genetic screens. Mutations that result in lethality or other severe defects will not be identified for roles that a partial loss of function mutation might reveal in postembryonic development. Conversely, genes with more subtle developmental phenotypes are also more likely to be overlooked. We identified genes in both these categories by RNA interference (RNAi) screens of 508 transcription factors (Table S3). One of these screens was conducted in a *ceh-2*::*YFP* background, a readout for vulB fate during the L4 stage, and identified *nhr-67* (*tailless*) (16). Because *nhr-67* deletion mutants die as young larvae, its role in vulval development was not identified. RNAi often causes a partial-loss-of-function phenotype instead of a null phenotype. It can also serve as a temporal or conditional downregulator of gene function. In the case of *nhr-67*, the larval lethality phenotype was bypassed, because RNAi was administered to L1 larvae, thus making its vulval phenotypes visible.

The second screen was conducted in a *zmp-1*::*GFP* background, focusing on perturbations of vulA expression. This screen also identified *nhr-67*, and the orphan nuclear hormone receptor *nhr-113* as positive regulators of *zmp-1* expression in vulA cells. *nhr-113* might have a narrow role in vulval organogenesis, because *nhr-113* RNAi has no effect on the regulation of several other genes: *cdh-3, ceh-2, dhs-31, lin-3*, or *pepm-1* (data not shown). These results, however, show that RNAi screens can identify new components of GRNs.

Differentiation of Discrete Fates in the 1° Vulval Lineage. The 1° lineage of the vulva generates four vulE and four vulF cells. Signals from the AC and Wnt are required for proper specification of these cell fates (25). The GRN, however, which acts downstream of these intercellular signals to guide differentiation of vulE and vulF fates, is not known. *egl-38*::*GFP* is detectable solely in the vulF cells (35); however, analysis of the mutant *egl-38* with *nhr-67* RNAi and mutant *cog-1* shows that *egl-38* functions in vulF cells and inhibits both *ceh-2* and *egl-17* expression in vulE (16) and vulF (3). Consequently, *egl-38*, *cog-1*, and *nhr-67* enforce spatial boundaries by preventing the 2° cell-fate associated genes *ceh-2* and *egl-17* from being expressed in 1° cells (Fig. S2, vulE and vulF).

It was speculated that the vulF cells, which are nearer the AC, have higher levels of *nhr-67* activity versus *cog-1*, whereas vulE cells have higher levels of *cog-1* than *nhr-67* (16). The recent availability of an *egl-38*::*GFP* reporter has allowed us to dissect further the GRN controlling the vulE versus vulF fates. *egl-38* expression in the vulF cells is positively regulated by *nhr-67* (Fig. 1). This *egl-38* expression is necessary for specification of the uv1 fate via regulation of the vulF specific gene *lin-3*, thus allowing

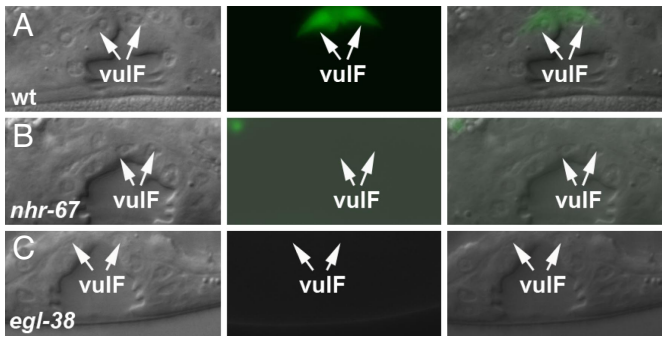


Fig. 1. *egl-38* expression in the vulF cells is dependent on several regulatory inputs. (*A–C*) Nomarski (*Left*), fluorescence (*Center*) and overlaid (*Right*). All animals displayed carry the *egl-38*::*GFP* [*guEx877*] transgene in their background. (*A*) In wild-type animals, *egl-38* is detected exclusively in the vulF cells (arrows). (*B*) *nhr-67* RNAi results in the abolition of *egl-38* expression in the vulF cells (arrows). (*C*) *egl-38(n578)* mutants lose the ability to positively autoregulate their expression levels in vulF (arrows).

for the proper development of a vulval-uterine connection (19). Conversely, *zmp-1* is expressed in vulE but not vulF, and *egl-38* is required for inhibiting *zmp-1* expression in the vulF cells (3). In vulE cells, where *egl-38* expression is absent, *lin-3* is not expressed and *cog-1* activated expression of *zmp-1* is observed (19). *egl-38* expression is unaffected in a *cog-1(sy275)* background, but in an *egl-38(n578)* background *egl-38*::*GFP* expression is decreased, suggesting that *egl-38* positively autoregulates in vulF (Fig. 1). Therefore, *egl-38* appears to stabilize vulF fate by repressing vulE characteristics and by reinforcing its own expression (Fig. 2).

In the vulE and vulF cells, *cog-1* and *nhr-67* negatively regulate both each other and themselves (16). We speculated that this

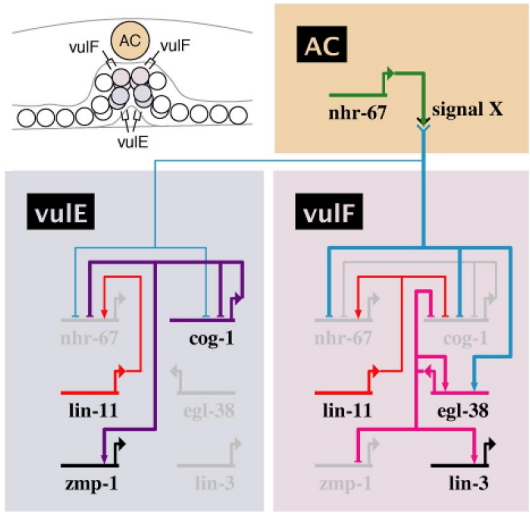


Fig. 2. Differentiation of vulE vs. vulF. The positions of the vulE and vulF cells relative to the anchor cell (AC) are shown. vulF is closer spatially than vulE to the AC. The network diagram was generated using BioTapestry Editor, Version 2.1.0 (www.biotapestry.org) (44). Linkages with arrowheads represent positive inputs and linkages with bar-heads represent repressor inputs for target gene expression. Black font indicates detectable expression levels and gray font indicates no detectable expression. This model presumes that *nhr-67* acts in the AC to differentiate between vulE and vulF cells. Signal X could be Ras, Wnt, or some other signaling pathway. The blue linkage in vulF is indicated by a thicker line than the blue linkage in vulE because it is hypothesized that vulF receives higher levels of signal X-mediated *nhr-67* signal from the AC. The thick purple and pink linkages highlight differences in the network architecture between the vulE and vulF cells, respectively.

Table 2. Expression of mutated *zmp-1* enhancers in the vulva

Construct	vulA, % (*n*)	vulC, % (*n*)	vulD, % (*n*)	vulE, % (*n*)	vulF, % (*n*)
mk50–51	86 (28)	0 (35)	0 (35)	89 (35)	0 (35)
103/4Δ	**34* (29)**	0 (41)	0 (41)	**0* (41)**	0 (41)
105/6Δ	72 (32)	4 (46)	**22† (46)**	**30* (46)**	7 (46)
107/8Δ	**15* (27)**	0 (31)	0 (31)	**6* (31)**	0 (31)

Shown are percentage of cells that express the indicated constructs. Values in bold indicate *P* values that are significantly different than wild type. *, $P = 0.000$; †, $P = 0.004$.

might allow 1° cells to rapidly switch their fates upon altered intra- and intercellular inputs. Presumptive vulE cells can induce uterine uv1 fate specification in the absence of vulF cells, thus ensuring the establishment of a proper vulval-uterine connection (19). These observations fit a model in which presumptive vulE and vulF cells are bipotential and positional cues help specify their fates. In this model, vulE, would then be biased for increased *nhr-67* activity in the absence of vulF. *egl-38* levels would thus increase, and vulF characteristics would be activated, whereas vulE characteristics would be inhibited.

Dissection of *zmp-1* Regulatory Elements. The vulA cells occupy a unique position as they form the outermost ring of cells and fuse to the surrounding hypodermal syncytium. *zmp-1* (zinc metalloproteinase) is first expressed in vulD and vulE cells beginning at the late L4 stage and in the vulA cells at the L4 to adult transition (23). vulA-specific expression of *zmp-1* is initiated in a different temporal window than its expression in vulD and vulE, which is particularly interesting because, in a *lin-29* background, vulD and vulE expression is abolished, but vulA expression is unaffected. Because *lin-29* affects early *zmp-1* expression (vulD and vulE) but not late expression (vulA), and because *lin-29* temporally regulates gene expression, it seems likely that modular *cis*-regulatory elements drive *zmp-1* expression.

A 386-bp fragment of upstream sequence (mk50–51) is sufficient to drive vulA- and vulE- but not vulD-specific expression of *zmp-1* (36). We analyzed sequences upstream of *zmp-1* and its orthologs in *C. briggsae* (CBG09053) and *C. remanei* (CRE04503), using Cistematic (37), which carries out motif-finding and phylogenetic footprinting (Table 2 and Fig. S3) and identified three motifs within the mk50–51 *zmp-1* enhancer element and a fourth motif five basepairs upstream of the 5′ end. We deleted the instances of the motifs and analyzed their effects on *zmp-1* reporter expression. Deletion of element 103/4 decreased vulA expression and abolished vulE expression. Previous *cis*-regulatory studies did not identify an element competent to drive vulA expression in the absence of vulE. Deletion of element 107/8 decreased vulA and vulE expression. Deletion of element 105/6, however, resulted in ectopic expression of mk50–51 *zmp-1*::*GFP* in vulC, vulD, and vulF. Thus, these elements likely act as positive (103/4 and 107/8) and negative (105/6) regulatory sites for controlling of *zmp-1* expression.

Conclusion

The regulatory architecture of the vulval GRN differs in all seven cell types. We postulate that this accounts for the differences in vulval cell fate, function, and morphology. Development of the *C. elegans* vulva utilizes several types of gene regulatory strategies that have been identified in other networks. For example, COG-1 participates in a network of mutual inhibitions with NHR-67 in the 1°-lineage derived vulval cells to differentiate between the vulE and vulF fates. This is reminiscent of the cross-inhibitory interaction of COG-1 and DIE-1 in the *C. elegans* ASE neurons (38). Vertebrate COG-1 homologs, the homeodomain proteins Nkx6.1 and Nkx6.2, might act in a similar manner (3). These proteins interact with the transcription factors Dbx1 and Dbx2 in a network of mutual inhibitions to specify motor neuron and interneuron fates during neural tube development (13). In another example, *lin-11* function is necessary for EGL-17 (FGF) expression in the vulva. An analogous network interaction is present during heart development in mice. Isl1, which like *lin-11* is a LIM homeodomain transcription factor, is required for the expression of FGFs (39).

We also describe new interactions within the vulva that may aid in the understanding of analogous regulatory interactions in other transcriptional networks, because the majority of the transcription factors and effectors present in this GRN have relatives in a diverse array of organisms. These interactions include positive regulation of *lin-11* by the heterochronic transcription factor LIN-29, *lin-3* (EGF) by LIN-11, and *egl-38* (Pax2/5/8) by NHR-67 (*tailless*). Our increased knowledge of the roles of transcription factors, such as *lin-11* (LIM homeodomain) and *egl-38* (Pax2/5/8) may help to further characterize other GRNs. For example, a LIM homeodomain protein (*Lim1*) and a Pax2/5/8 protein (*Pax2*) are involved in murine kidney development (40). Thus, the GRN that directs nephrogenesis in mice appears to share at least two components with the vulval network in *C. elegans*. Our understanding of the kidney morphogenesis GRN in mammals may be enhanced by investigating the mouse orthologs of other components of the vulval GRN.

New approaches are needed to elucidate all of the constituents of the complex regulatory architecture that directs organogenesis. Until now, most of the key players in vulval organogenesis have been isolated using traditional mutagenesis screens. These types of screens, however, are often limited to identifying only those factors with severe phenotypes. To account for the regulation of all genes that are expressed and function during vulva development it is apparent that other members of the vulval GRN are yet to be identified. There are transcription factors with major and minor effects. For example, LIN-11 is required in two tissues for egg-laying and regulates gene expression in multiple vulval cell types. By contrast, using an RNAi screen we were able to identify NHR-113, a factor that is possibly only required for fine tuning gene expression in the vulA cells.

Analysis of the effects of various genetic perturbations on vulval organogenesis has revealed detailed spatial and temporal distinctions in the regulation of diverse yet related cell types. Our approach provides for precise and accurate study of gene expression in an intact organism and unveils the distinct network architecture in the different cell types. Further dissection of the genomic network within the differentiated cell types would extend our knowledge of vulval organogenesis and could also provide further insights into organogenesis in other systems.

Materials and Methods

Genetics and RNAi. *C. elegans* strains were grown and constructed using standard protocols. The *nhr-67* RNAi feeding protocol was described in ref. 16.

Generation of Reporter Transgenes. *syEx1009, syEx1091* and *syEx1018* were generated by deleting one motif each, 103/4 (5′-CGAGTACGTTTACAC-3′), 105/6 (5′-GTACGTATTGCTT-3′), or 107/8 (5′-AGAAAAAGTAGAAGG-3′), respectively, from the mk50–51 (36) construct and replacing the motif with a *Sac*II restriction site. These constructs were then individually injected into the gonads of *unc-119(ed4); him-5(e1490)* animals (41), using *unc-119*(+) (42) and pBSK+ (Stratagene). *syEx756* and *syEx1012* were generated by injecting the constructs pNP10 (43) and mk50–51 (36), respectively, into a *unc-119(ed4); him-5(e1490)* background.

Microscopy. 3 mM levamisole was used to anesthetize transgenic animals for observation using Nomarski DIC optics.

ACKNOWLEDGMENTS. We thank A. Mortazavi for help with comparative genomics, H. Chamberlin (Ohio State University, Columbus, OH) for

guEx877; A. Fire (Stanford University, Palo Alto, CA) for GFP vectors; S.-H. Kim for his assistance with the *zmp-1*::*GFP* RNAi screen; M. Kato, S. Kuntz, E. Schwarz and X. Wang for comments on the manuscript; and T. Inoue for discussions. T.O.R. was supported by National Institutes of Health Grant T32-GM007616. P.W.S. is an investigator with the Howard Hughes Medical Institute.

1. Davidson EH (2001) Genomic Regulatory Systems: Development and Evolution (Academic, San Diego).
2. Sternberg PW (2005) Vulval development. In *WormBook*, ed Community TCeR. Available at www.WormBook.org. Accessed June 30, 2008.
3. Inoue T, Wang M, Ririe TO, Fernandes JS, Sternberg PW (2005) Transcriptional network underlying *Caenorhabditis elegans* vulval development. *Proc Natl Acad Sci USA* 102:4972–4977.
4. Sherwood DR Sternberg PW (2003) Anchor cell invasion into the vulval epithelium in *C. elegans*. *Dev Cell* 5:21–31.
5. Freyd G, Kim SK, Horvitz HR (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene *lin-11*. *Nature* 344:876–879.
6. Bach I (2000) The LIM domain: Regulation by association. *Mech Dev* 91:5–17.
7. Hobert O, Westphal H (2000) Functions of LIM-homeobox genes. *Trends Genet* 16:75–83.
8. Rougvie AE, Ambros V (1995) The heterochronic gene *lin-29* encodes a zinc finger protein that controls a terminal differentiation event in *Caenorhabditis elegans*. *Development (Cambridge, UK)* 121:2491–2500.
9. Bettinger JC, Euling S, Rougvie AE (1997) The terminal differentiation factor LIN-29 is required for proper vulval morphogenesis and egg laying in *Caenorhabditis elegans*. *Development (Cambridge, UK)* 124:4333–4342.
10. Newman AP, Inoue T, Wang M, Sternberg PW (2000) The *Caenorhabditis elegans* heterochronic gene *lin-29* coordinates the vulval-uterine-epidermal connections. *Curr Biol* 10:1479–1488.
11. Palmer RE, Inoue T, Sherwood DR, Jiang LI, Sternberg PW (2002) *Caenorhabditis elegans cog-1* locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. *Dev Biol* 252:202–213.
12. Habener JF, Kemp DM, Thomas MK (2005) Minireview: Transcriptional regulation in pancreatic development. *Endocrinology* 146:1025–1034.
13. Vallstedt A, *et al.* (2001) Different levels of repressor activity assign redundant and specific roles to *Nkx6* genes in motor neuron and interneuron specification. *Neuron* 31:743–755.
14. Schedl A, Hastie ND (2000) Cross-talk in kidney development. *Curr Opin Genet Dev* 10:543–549.
15. Chamberlin HM, *et al.* (1997) The *PAX* gene *egl-38* mediates developmental patterning in *Caenorhabditis elegans*. *Development (Cambridge, UK)*124:3919–3928.
16. Fernandes JS, Sternberg PW (2007) The *tailless* ortholog *nhr-67* regulates patterning of gene expression and morphogenesis in the *C. elegans* vulva. *PLoS Genet* 3:e69.
17. Hartmann B, Reichert H, Walldorf U (2001) Interaction of gap genes in the *Drosophila* head: *Tailless* regulates expression of *empty spiracles* in early embryonic patterning and brain development. *Mech Dev* 109:161–172.
18. Burdine RD, Branda CS, Stern MJ (1998) EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in *C. elegans*. *Development (Cambridge, UK)* 125:1083–1093.
19. Chang C, Newman AP, Sternberg PW (1999) Reciprocal EGF signaling back to the uterus from the induced *C. elegans* vulva coordinates morphogenesis of epithelia. *Curr Biol* 9:237–246.
20. Colavita A, Tessier-Lavigne M (2003) A Neurexin-related protein, BAM-2, terminates axonal branches in *C. elegans*. *Science* 302:293–296.
21. Shen K, Fetter RD, Bargmann CI (2004) Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. *Cell* 116:869–881.
22. Hwang HY, Horvitz HR (2002) The *Caenorhabditis elegans* vulval morphogenesis gene *sqv-4* encodes a UDP-glucose dehydrogenase that is temporally and spatially regulated. *Proc Natl Acad Sci USA* 99:14224–14229.
23. Inoue T, *et al.* (2002) Gene expression markers for *Caenorhabditis elegans* vulval cells. *Mech Dev* 119:S203–S209.
24. Pettitt J, Wood WB, Plasterk RH (1996) *cdh-3*, a gene encoding a member of the cadherin superfamily, functions in epithelial cell morphogenesis in *Caenorhabditis elegans*. *Development (Cambridge, UK)* 122:4149–4157.
25. Wang M, Sternberg PW (2000) Patterning of the *C. elegans* 1° vulval lineage by RAS and Wnt pathways. *Development (Cambridge, UK)* 127:5047–5058.
26. Wang X, Greenberg JF, Chamberlin HM (2004) Evolution of regulatory elements producing a conserved gene expression pattern in *Caenorhabditis*. *Evol Dev* 6:237–245.
27. Clark SG, Chisholm AD, Horvitz HR (1993) Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*. *Cell* 74:43–55.
28. Wang BB, *et al.* (1993) A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. *Cell* 74(1):29–42.
29. Wagmaister JA, Gleason JE, Eisenmann DM (2006) Transcriptional upregulation of the *C. elegans* Hox gene *lin-39* during vulval cell fate specification. *Mech Dev* 123:135–150.
30. Perens EA, Shaham S (2005) *C. elegans daf-6* encodes a Patched-related protein required for lumen formation. *Dev Cell* 8:893–906.
31. Hao L, Johnsen R, Lauter G, Baillie D, Burglin TR (2006) Comprehensive analysis of gene expression patterns of *hedgehog*-related genes. *BMC Genomics* 7:280.
32. Davis MW, Birnie AJ, Chan AC, Page AP, Jorgensen EM (2004) A conserved metalloprotease mediates ecdysis in *Caenorhabditis elegans*. *Development (Cambridge, UK)* 131:6001–6008.
33. Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* 323:785–793.
34. Euling S, Bettinger JC, Rougvie AE (1999) The LIN-29 transcription factor is required for proper morphogenesis of the *Caenorhabditis elegans* male tail. *Dev Biol* 206:142–156.
35. Rajakumar V, Chamberlin HM (2007) The *Pax2/5/8* gene *egl-38* coordinates organogenesis of the *C. elegans* egg-laying system. *Dev Biol* 301:240–253.
36. Kirouac M, Sternberg PW (2003) *cis*-Regulatory control of three cell fate-specific genes in vulval organogenesis of *Caenorhabditis elegans* and *C. briggsae*. *Dev Biol* 257:85–103.
37. Mortazavi A, Leeper Thompson EC, Garcia ST, Myers RM, Wold B (2006) Comparative genomics modeling of the NRSF/REST repressor network: From single conserved sites to genome-wide repertoire. *Genome Res* 16:1208–1221.
38. Johnston RJ, Jr, Chang S, Etchberger JF, Ortiz CO, Hobert O (2005) MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc Natl Acad Sci USA* 102:12449–12454.
39. Cai CL, *et al.* (2003) Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev Cell* 5:877–889.
40. Yu J, McMahon AP, Valerius MT (2004) Recent genetic studies of mouse kidney development. *Curr Opin Genet Dev* 14:550–557.
41. Mello CC, Kramer JM, Stinchcomb D, Ambros V (1991) Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences *EMBO J* 10:3959–3970.
42. Maduro M, Pilgrim D (1995) Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* 141:977–988.
43. Stringham E, Pujol N, Vandekerckhove J, Bogaert T (2002) *unc-53* controls longitudinal migration in *C. elegans*. *Development (Cambridge, UK)* 129:3367–3379.
44. Longabaugh WJ, Davidson EH, Bolouri H (2005) Computational representation of developmental genetic regulatory networks. *Dev Biol* 283:1–16.

Appendix II:

Transcriptional network underlying *Caenorhabditis elegans* vulval development

Takao Inoue, Minqin Wang, Ted O. Ririe, Jolene S. Fernandes, and Paul W. Sternberg

Published in *PNAS*, **102**(14):4972-4977 (2005).

Transcriptional network underlying *Caenorhabditis elegans* vulval development

Takao Inoue, Minqin Wang, Ted O. Ririe, Jolene S. Fernandes, and Paul W. Sternberg*

Division of Biology and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125

Edited by Eric H. Davidson, California Institute of Technology, Pasadena, CA, and approved February 1, 2005 (received for review November 11, 2004)

The vulval development of *Caenorhabditis elegans* provides an opportunity to investigate genetic networks that control gene expression during organogenesis. During the fourth larval stage (L4), seven vulval cell types are produced, each of which executes a distinct gene expression program. We analyze how the expression of cell-type-specific genes is regulated. Ras and Wnt signaling pathways play major roles in generating the spatial pattern of cell types and regulate gene expression through a network of transcription factors. One transcription factor (*lin-29*) primarily controls the temporal expression pattern. Other transcription factors (*lin-11*, *cog-1*, and *egl-38*) act in combination to control cell-type-specific gene expression. The complexity of the network arises in part because of the dynamic nature of gene expression, in part because of the presence of seven cell types, and also because there are multiple regulatory paths for gene expression within each cell type.

organogenesis | signaling pathways | transcription

Developmental events are driven by spatially and temporally regulated gene expression. Understanding how complex patterns of expression are produced is therefore a critical part of deciphering mechanisms of development. In general, intercellular signaling mechanisms interact with a network of transcription factors to generate cell-type-specific patterns of gene expression. The late stage of *Caenorhabditis elegans* vulval development offers a useful model in which to study this process. During this period of vulval development, seven distinct cell types are produced that express unique combinations of genes. Over the last several years, a number of genes were discovered that are expressed in cell-type and stage-specific patterns in the vulva, and several transcription factors were found to regulate these genes. In this paper, we synthesize and extend our current knowledge of this genetic network.

The *C. elegans* vulva connects the uterine lumen to the outside, allowing for passage of sperm and fertilized eggs (1). Vulval cells are generated postembryonically from precursor cells P3.p P4.p, P5.p, P6.p, P7.p, and P8.p [also called vulval precursor cells (VPC)]. During the mid-third larval (L3) stage, EGF and Notch signaling induces the middle three VPCs (P5.p, P6.p, and P7.p) to adopt vulval fates, whereas P3.p, P4.p, and P8.p fuse with the hypodermal syncytium, hyp7 (2–6).

During the late-L3 to the early-L4 stage, P5.p, P6.p, and P7.p undergo two or three rounds of cell division to produce 22 nuclei (7) (Fig. 1*A*). These nuclei are in cells of seven types (vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF), as evidenced by subsequent morphogenetic movements and by the pattern of gene expression (8, 9) (Fig. 1*B*). The seven cell types that are present in the adult vulva represent specializations within the general epithelial cell class. These cells exhibit cell-type general features; for example, each expresses *ajm-1*, a component of the apical junction that connects neighboring cells in epithelial tissues (8). However, in addition, each cell type exhibits functional specializations: vulF cells, which form the innermost section of the vulva, connect directly with cells of the uterus. vulE cells form structural attachments to lateral hypodermal (seam) cells. vulC and vulD cells attach to vulval muscles that open the vulva for the passage of eggs. vulA cells form attachment to the hyp7 syncytium. It is expected that gene expression differences underlie these specializations.

Here, we are concerned with the execution of cell-type-specific gene expression programs during the late L3 and L4 stages, mostly after the terminal division of vulval cells. During this period, each cell type exhibits a cell-type-specific pattern of gene expression, and several transcription factors are known that regulate the expression of these cell-type-specific genes. We bring together our current knowledge of this system to produce the framework in which to investigate the gene regulatory network controlling vulval organogenesis.

Materials and Methods

Determination of Gene Expression Patterns. Essentially all gene expression analyses described in this paper (including data from other papers) were carried out by using *gfp* reporter transgenes. For all results, it is possible that reporter expression does not accurately reflect the expression pattern of the endogenous gene. For simplicity, we refer to the reporter by the corresponding gene name.

The expression pattern of C55C3.5 was determined by using *gfp* reporter clone pUL#G221N (I. Hope, personal communication). This plasmid was injected into *unc-119*(*ed4*) animals by using the plasmid pDP#MM016B [*unc-119*(+)] as a coinjection marker (10). Of genes listed in Fig. 1*B* and in the main text, we have not examined the expression pattern of *syg-2*, *bam-2*, and *sqv-4*. Because GFP is likely to be stable for many hours, the time at which expression is turned off is not reliably indicated by decreased GFP expression. For most genes we analyzed, GFP fluorescence persists into the adult stage.

Genotypes. For Tables 1–3, *gfp* reporter transgenes used were *ayIs4[egl-17::gfp]*, *syIs50[cdh-3::gfp]*, *syIs49[zmp-1::gfp]*, and *syIs54[ceh-2::gfp]* (9). The *egl-26::gfp* transgenic line analyzed was *kuIs36* (11). Mutations used are; *cog-1*(*sy275*), *cog-1*(*sy607*), *lin-29*(*sy292*), *lin-11*(*n389*), and *egl-38*(*n578*). Of two *cog-1* transcripts, the longer *cog-1A* transcript contains a corepressor-binding domain, whereas the shorter *cog-1B* transcript does not (12). *sy275* is a missense mutation predicted to affect both transcripts. *sy607* is a deletion that eliminates the *cog-1A* transcript. The two alleles exhibit complementary defects in vulval development (13). Although both alleles are recessive, it is not known whether the loss of *cog-1* function causes observed phenotypes. *lin-29*(*sy292*) and *lin-11*(*n389*) are strong loss-of-function alleles, and *egl-38*(*n578*) is a reduction-of-function allele. Strains were constructed by using standard methods.

Results and Discussion

Vulval Cell-Type-Specific Gene Expression. A number of genes are expressed in specific subsets of vulval cells (Fig. 1*B*). Previously

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: VPC, vulval precursor cell; L*n* stage, larval *n* stage.

*To whom correspondence should be addressed. E-mail: pws@caltech.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0408122102

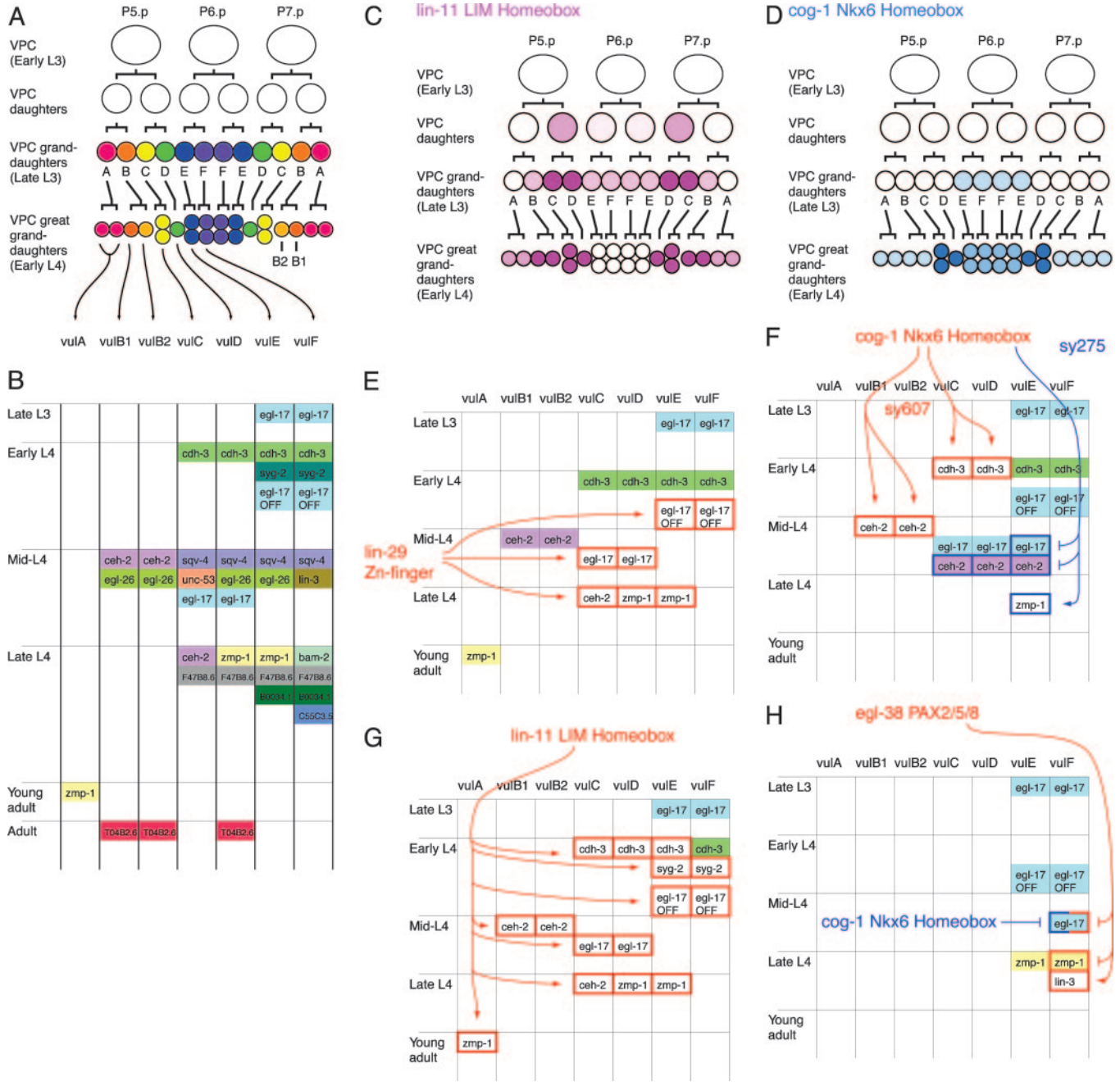


Fig. 1. The pattern of gene expression during late stages of vulval development. (*A*) An overview of vulval development. Lineal origins of 22 vulval nuclei are indicated. ''ABCDEFFEDCBA'' refer to vulval cell types vulA, vulB1, vulB2, vulC, vulD, vulE, and vulF. vulB is the only case in which a single VPC granddaughter gives rise to two cell types (8). Vulval cell nuclei at each stage are positioned as indicated (left, anterior; right, posterior). (*B*) A summary of cell-type specificity and timing of expression in the wild type (*Materials and Methods*) (9, 11, 13–21, 24–26). Boxes indicate stages at which gene expression is activated. The vertical order of events within each time block is arbitrary. For *egl-17*, vulE/vulF expression begins in P6.p (early L3) and persists in their descendants (vulE and vulF) until turned off in the early L4 stage. This inactivation, which is regulated by *lin-29* and *lin-11*, is indicated by the box marked ''*egl-17* OFF.'' *ceh-2* is expressed at a higher level in vulB1 compared with vulB2. (*C*) Expression pattern of *lin-11*. The diagrammed pattern is based on the *lin-11*::*gfp* transgene *syIs80* (26). (*D*) Expression pattern of *cog-1*. The pattern is based on the *cog-1*::*gfp* transgene *syIs63* (13). (*E*) The altered pattern of gene expression in *lin-29* mutants (9, 23) (Tables 1 and 2). White boxes with the red outlines indicate loss of expression and loss of *egl-17* down-regulation in the *lin-29* mutant. *lin-29* appears to regulate events that occur during the mid-L4 to the late L4 stage. (*F*) The altered pattern of gene expression in *cog-1* mutants (Table 3 and Fig. 2). Arrows are drawn with the assumption that both *sy607* and *sy275* phenotypes are caused by different reduction of function of the *cog-1* gene. Filled boxes with red or blue outline indicate ectopic expression. (*G*) Altered pattern of gene expression in *lin-11* mutants (26). (*H*) Altered pattern of gene expression in the *egl-38* mutant (16) (Tables 1 and 2). *egl-17* expression in vulF is observed in the *cog-1*(*sy275*); *egl-38* double mutant, suggesting a redundant repression mechanism.

described genes of this type include *cdh-3* (14), *egl-17* (15), *lin-3* (16), *zmp-1* (9, 17), *ceh-2* (9), T04B2.6 (9), F47B8.6 (9), B0034.1 (9), *unc-53* (18), *egl-26* (11), *sqv-4* (19), *bam-2* (20), and *syg-2* (21). *egl-26* was previously reported to express in vulE and vulB2 cells (11). We found that a nuclear-localized *egl-26*::*gfp* transcriptional fusion expressed in vulB1, vulB2, vulD, and vulE cells (*Materials and Methods*). The expression was somewhat variable and was observed starting from the mid-L4 stage and continuing

Table 1. Expression of *zmp-1* in vulE and vulF cells

Genotype	vulE	vulF	No. of animals
Wild type	+	−	80
lin-11	−	−	55*
lin-29	−	−	50†
cog-1 (sy275)	−	−	52
egl-38	+	+	48
lin-11; egl-38	−	−	52
lin-29; egl-38	−	−	56
cog-1; egl-38	−	−	56

*Ref. 26.
†Ref. 9.

into the adult stage. The C55C3.5 gene encoding a novel protein was previously found to express in vulval cells (I. Hope, personal communication). We found that *C55C3.5*::*gfp* was expressed in vulF cells, starting from the late-L4 and continuing into the adult stage.

Several conclusions can be drawn from Fig. 1*B*. First, all seven cell types exhibit distinct programs of gene expression, despite the fact that these cells are related by cell lineage and function. [vulB1 and vulB2 differ in the level of *ceh-2* expression but otherwise have similar expression profiles (9)]. Distinct expression profiles likely underlie distinct functions of vulval cell types. For example, *lin-3*, which encodes an EGF-related signaling protein, is expressed in vulF cells in the mid-L4 stage (16). This signal is required for a vulva-to-uterus signaling that induces a specific fate, uv1, in uterine cells adjacent to vulF.

The pattern of marker expression also reveals a strict temporal regulation of gene expression (Fig. 1*B*). For example, *cdh-3* is expressed in early L4, F47B8.6 is expressed in late L4, and T04B2.6 is expressed ≈1 day after the L4-to-adult molt (9). For *egl-17*, *ceh-2*, *zmp-1*, and *sqv-4*, the timing of gene expression is different for different vulval cells (9, 15, 19). For example, *egl-17* is expressed in vulE and vulF cells in the L3 stage and in vulC and vulD in the L4 stage.

Trans-Regulation of Vulva Gene Expression. The analysis of the regulatory network controlling the pattern of gene expression in the vulva has focused primarily on the effect of transcription factor mutations on gene expression reporter transgenes. In most cases, a direct transcriptional regulation of the target has not been demonstrated. Key results are summarized in Fig. 1 *E–H*. So far, important regulators are *lin-29* (encoding Zn-finger transcription factor; Fig. 1*E*) (9, 22, 23), *cog-1* (Nkx6 homeodomain; Fig. 1 *D* and *F*) (13), *lin-11* (LIM homeodomain; Fig. 1 *C* and *G*) (24–26), and *egl-38* (PAX 2/5/8; Fig. 1*H*) (16, 27).

Table 2. Expression of *egl-17* in vulE and vulF cells (L4)

Genotype	vulE	vulF	No. of animals
Wild type	−	−	59
cog-1 (sy275)	+++	−	46
egl-38	−	−	38
cog-1; egl-38	+++	+++	37
lin-11	—*	—*	45†
lin-29	—*	—*	43‡
lin-11; lin-29	—*	—*	38
lin-11; cog-1	—*	—*	43
lin-29; cog-1	—*	—*	40
lin-11; cog-1; egl-38	—*	—*	35
lin-29; cog-1; egl-38	—*	—*	36

*These cells express *egl-17*::*gfp* at a low level. We interpret these as the persistence of L3 expression.
†Ref. 26.
‡Ref. 23.

Table 3. Expression of *egl-17, ceh-2,* and *cdh-3* in *cog-1* mutants

Reporter	Mutations	vulA	vulB1 and vulB2	vulC	vulD	vulE	vulF
egl-17	+	0	0	100	100	0	0
egl-17	*cog-1 (sy275)*	0	0	100	92	92	0
egl-17	*cog-1 (sy607)*	0	0	93	100	0	0
ceh-2	+	0	100	0	0	0	0
ceh-2	*cog-1 (sy275)*	20	90	80	80	88	0
ceh-2	*cog-1 (sy607)*	0	0	0	0	0	0
cdh-3	+	0	0	100	100	100	100
cdh-3	*cog-1 (sy275)*	0	0	100	100	100	100
cdh-3	*cog-1 (sy607)*	0	0	14	14	71	94

Percentages of cells in mid-L4 animals that expressed *egl-17*::*gfp*, *ceh-2*::*gfp* and *cdh-3*::*gfp*. See Table 4, which is published as supporting information on the PNAS web site, for number of cells scored.

A Temporal Regulator of Gene Expression. *lin-29* is required for the expression of *egl-17* in vulC and vulD (23), *ceh-2* in vulC (9), and *zmp-1* in vulD and vulE (Fig. 1*E*, Tables 1 and 2, and Fig. 5, which is published as supporting information on the PNAS web site) (9). By contrast, *lin-29* in not required for the expression of *cdh-3* in vulC, vulD, vulE, vulF (9), *ceh-2* in vulB (9), *egl-17* in vulE and vulF (23), and *zmp-1* in vulA (9). Moreover, the expression of *egl-17* in vulE and vulF is observed during the L4 stage (23), suggesting that the mechanism that turns off *egl-17* expression in these cells is compromised (Fig. 5). These *lin-29* phenotypes are not easily explained by cell fate changes between vulval cell types but suggest a temporal regulatory defect: *lin-29* mutations cause loss of events associated with the mid-to-late L4 time points. This interpretation of these data is particularly attractive, because *lin-29* mutations are known to cause heterochronic defects in other tissues, specifically in the L4-to-adult transition in the lateral hypodermis (22, 28, 29). *lin-29* is expressed in all vulval cells, starting in the mid-L3 stage and continuing through the L4 stage (30).

Cell-Type-Specific Regulators of Gene Expression. We analyzed the effect of two *cog-1* (Nkx6.1/6.2 homeodomain) mutations on the expression of vulval-cell-specific gene expression reporters (Fig. 1*F*, Table 3, and *Materials and Methods*). *cog-1*(*sy275*) is a missense mutation in the homeodomain, and *cog-1*(*sy607*) is a small deletion that eliminates one of two *cog-1* transcripts (13). We found that in the mid-L4 stage, *cog-1*(*sy275*) caused ectopic expression of *egl-17* in vulE cells (Fig. 2) and ectopic expression of *ceh-2* in vulC, vulD, and vulE cells and loss of *zmp-1* expression in vulE cells. In contrast, *cog-1*(*sy607*) caused loss of *cdh-3* expression in vulC, vulD, and vulE cells and loss of *ceh-2* expression in vulB. These results indicate that *egl-17*, *cdh-3*, *ceh-2*, and *zmp-1* are regulated by the *cog-1* gene. Although some *cog-1* expression is observed in all vulval cells, *gfp* reporters suggest that *cog-1* is most strongly expressed in vulC and vulD and weakly in vulE and vulF, implying a cell-type-specific function (13) (Fig. 1*D*).

A somewhat similar situation is presented with *lin-11* (LIM-homeodomain) (Fig. 1 *C* and *G*). During the L4 stage, *lin-11* is expressed strongly in vulB, vulC, and vulD and weakly in other vulval cells, suggesting that *lin-11* is involved in the specification of these cell types (24, 26). However, unexpectedly, *lin-11* is cell-autonomously required for expression of most vulval genes tested, including in cells where the *lin-11* level is low (26).

egl-38 is a PAX2/5/8 transcription factor required for expres-

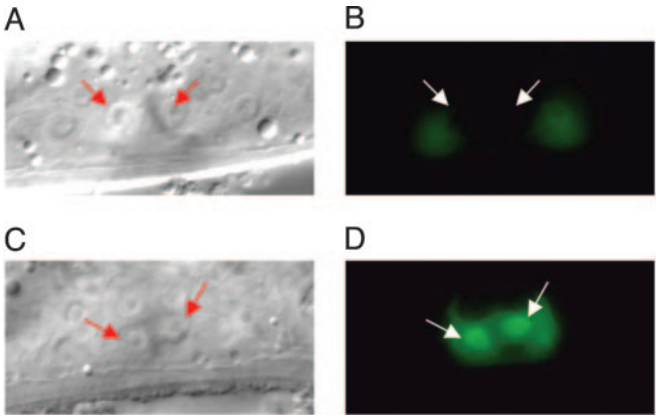


Fig. 2. Regulation of *egl-17* by *cog-1*. (*A* and *B*) Nomarski and epifluorescence images of wild-type mid-L4 animal carrying the *egl-17*::*gfp* transgene. Arrows point to vulE nuclei. vulE cells are not fluorescent. (*C* and *D*) *cog-1*(*sy275*) animals at the same stage carrying the *egl-17*::*gfp* transgene. vulE cells are fluorescent.

sion of the *lin-3* gene in vulF cells (16, 27). We found that *egl-38* represses expression of *zmp-1* in vulF cells, indicated by ectopic *zmp-1* expression in *egl-38* mutants (Fig. 1*G* and Table 2). In addition, in an *egl-38*; *cog-1* double mutant, *egl-17* is expressed in both vulE and vulF cells. Thus, *egl-38* is also capable of repressing *egl-17* expression in vulF cells, although in the wild type, this function is redundant with the *cog-1*-dependent mechanism that restricts *egl-17* expression to vulC and vulD. *egl-38* is currently the best example of cell-type-specific factors, promoting expression of some genes (*lin-3*) and repressing expression of others (*zmp-1*, *egl-17*) in a single cell type, vulF.

Regulators of the Transcription Factor Network. The transcription factor network that regulates gene expression in individual cell types must be regulated by the cell-fate-patterning mechanism that specifies each cell to a specific fate and does so in a spatially precise pattern. In the vulva, the cell types occur in a specific ABCD-EFFE-DCBA pattern (Fig. 1*A*). Although the full mechanism that establishes this pattern is not known, Wnt signals, mediated by *lin-17* (Frizzled-type Wnt receptor) and *lin-18* (Ryk-type Wnt receptor), control the anterior/posterior order of cell types among P7.p descendants (31, 32) (Fig. 3). Analysis of *cog-1* (31) and *lin-11* (25) expression in *lin-17* and *lin-18* mutants indicates that Wnt signaling establishes the correct spatial pattern of transcription factor expression. As described above (Tables 1 and 3 and Fig. 1 *F* and *G*) (26), *cog-1* and *lin-11*, in turn, control the expression pattern of *egl-17* and *cdh-3*. Patterns of *egl-17* and *cdh-3* expression observed in *lin-17* and *lin-18* mutants are consistent with high levels of *cog-1* and *lin-11* turning on the expression of these genes (31, 32). Another set of cell-fate-patterning mechanisms controlling gene expression was revealed by the analysis of vulE vs. vulF fate specification using the *zmp-1* reporter. A dominant-negative Ras or the ablation of the anchor cell disrupts the pattern of *zmp-1* expression in presumptive vulE and vulF cells, indicating that a Ras-mediated signal, probably from the anchor cell, establishes the spatial pattern of cell fates (17).

These results confirm that cell–cell communication is important in patterning cell fates, and that signaling pathways operate through the transcription factor network to control the pattern of gene expression. Expression patterns of various genes (Fig. 1 *B–D*) suggest that transcription factors are expressed in all vulval cells at different levels, whereas genes regulated by them have relatively simple on/off patterns of expression. This difference suggests that the spatial pattern becomes progressively more refined as the information is passed through the regulatory network. This progressive refinement of pattern is likely a consequence of integration of information from multiple regulatory mechanisms, such as intercellular communication and feedback regulation. Many of these disparate data inputs are likely processed at the level of cis-regulatory modules. Thus, the

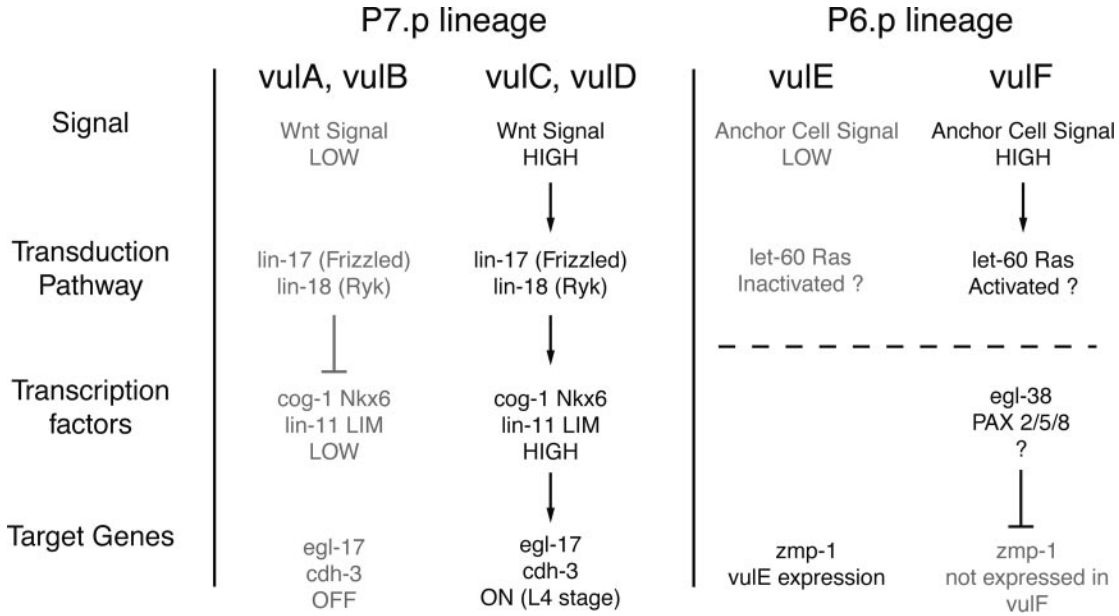


Fig. 3. Link between cell fate patterning mechanisms and gene expression. In general, inductive signals regulate transcription factor networks to regulate gene expression. In the P7.p (but not P5.p) lineage, Wnt signals transduced by *lin-17* and *lin-18* control the pattern of *cog-1* and *lin-11* expression (25, 31). *cog-1* and *lin-11* in turn regulate *egl-17* and *cdh-3* expression (Table 3) (26). It has not been determined whether *cog-1* and *lin-11* regulate each other. In the P6.p lineage, an anchor cell signal and a *let-60* Ras signal transduction pathway are required to establish the correct pattern of *zmp-1* expression pattern (17). *zmp-1* expression is also repressed in vulF by *egl-38* PAX2/5/8 (Table 1). It is not known whether the patterning mechanism acts through *egl-38*. The expression pattern of *egl-38* is also not known.

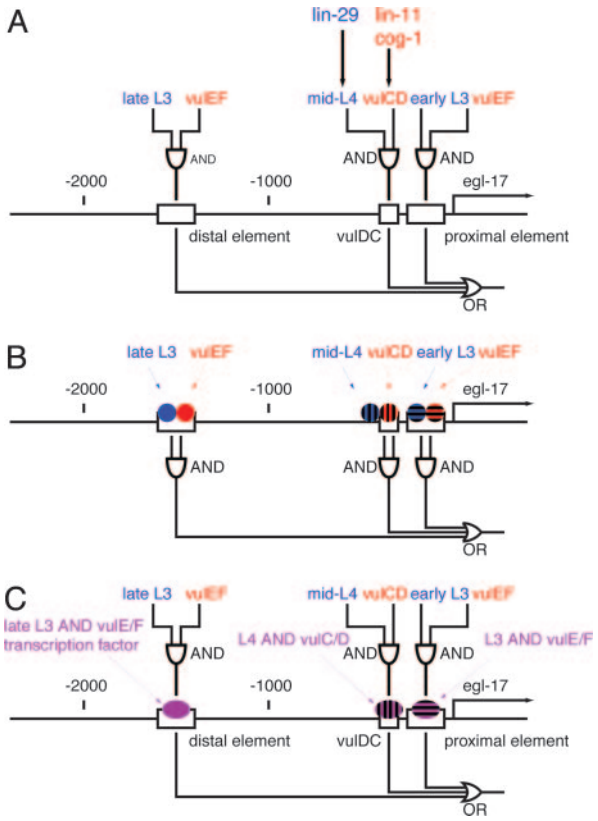


Fig. 4. cis-regulatory elements of *egl-17*. (*A*) A map of the *egl-17* 5′ regulatory region. Boxes indicate enhancer elements defined by Cui and Han (34) and Kirouac and Sternberg (33). ''AND'' and ''OR'' logic gate symbols indicate sites and logic of information integration. Temporal (blue) and spatial (red) information is integrated as indicated by the logic circuit diagram to produce the complete *egl-17* expression pattern. In one model (*B*), spatially and temporally regulated transcription factors each bind directly to the *egl-17* cis-regulatory region. The integration of information takes place on enhancer elements. In the alternative model (*C*), spatial and temporal cues are integrated at the transcription factor level. These transcription factors (purple) with both spatially and temporally restricted activity regulate each enhancer element.

spatial pattern of transcription factor effects becomes more restricted than the spatial pattern of transcription factor expression. This hypothesis is consistent with the observation that cells affected by *lin-11* and *cog-1* mutations do not correspond directly to cells that express high levels of *lin-11* and *cog-1*.

cis-Regulation of Vulva Gene Expression. cis-regulatory elements (e.g., enhancers) have been analyzed in detail for several genes expressed in the vulva, most notably *egl-17*, *cdh-3*, and *zmp-1*, using transgenic assays (33, 34). A comparative genomics analysis of the regulatory region of orthologs from *C. elegans* and *Caenorhabditis briggsae* has also proved useful.

Here, we focus on the analysis of the *egl-17* gene. As shown in Fig. 1*B*, this gene is expressed in vulE and vulF cells during the L3 stage and in vulC and vulD cells during the L4 stage. Dissection of the 5′ regulatory region revealed that there are three separable enhancer elements, two driving expression in vulE and vulF and one driving expression in vulC and vulD (33, 34) (Fig. 4*A*). Notably, each of these elements drives expression at different times. The distal vulE/vulF element drives expression in the mid-L3 to early-L4 and the proximal vulE/vulF element drives expression in the early to mid-L3 stage (34). The vulDC element drives expression in the mid-L4 stage. Thus, the expression of *egl-17* is produced by the composite activity of three discrete enhancers, each of which drives both spatially and temporally restricted pattern of expression. We propose two models for how the information that operates on these enhancers is integrated. In one model (Fig. 4*B*), temporal (blue) and spatial (red) regulators both bind directly to the *egl-17* promoter, and information integration is achieved directly on the cis-regulatory element. Alternatively, transcription factors that bind to each of these promoters may already combine temporal and spatial information (Fig. 4*C*). Our results indicate that the vulDC element regulating mid-L4 expression is likely regulated by *lin-29*, *lin-11*, *cog-1*, and *egl-38*. Additional experiments are necessary to determine the molecular mechanism of information integration.

Conclusion

The late vulval development of *C. elegans* offers an excellent system in which to investigate cell fate determination and regulation of cell-type-specific gene expression. In particular, this system combines single-cell resolution with a high degree of temporal resolution in an easily manipulated model organism. In many respects, vulval development is reminiscent of other systems in that transcription factors are expressed in overlapping domains, and the identity of each domain is established combinatorially by the presence or absence of specific subsets of these transcription factors. One interesting example with possible parallels to the vulva is the fate-specification mechanism in the vertebrate ventral neural tube (35). In this system, Nkx6.1 and Nkx6.2 homeodomain proteins (homologs of *cog-1*) interact with transcription factors Dbx1 and Dbx2 in a mutually repressive network, and different activities of repressor proteins help establish the spatial pattern of cell fates (36, 37). It is possible that *C. elegans cog-1* functions in a similar manner in the vulva.

Analysis of vulval development also highlights several features that are not necessarily evident in other systems. First, analysis of vulval development has revealed a highly complex pattern of temporal regulation, which is undoubtedly a feature of most organogenetic processes (for example, see refs. 38 and 39). The involvement of *lin-29*, a known regulator of stage-specific development in *C. elegans*, suggests that the global mechanism of temporal regulation feeds into the development of this particular organ. Additional mechanisms probably exist that control expression at other time points. Whether these other time points are regulated by a global mechanism or in an organ-autonomous manner is not yet clear.

One concept that has been invoked in analyses of cell or organ fate specification is that of ground state and selector genes. For example, in *Drosophila* appendage development, it has been proposed that a default "ground state" exists and is modified by "selector" genes to produce an antenna or a leg (40). The concept can be applied to the level of individual cell types as well (for example, ref. 41). From this point of view, the cell-type-specific transcription factors *cog-1*, *lin-11*, and *egl-38* can be thought of as selector genes for subsets of vulval cell types. What is the ground state of vulval cells in the absence of selector genes? A cell in such a state presumably will not express the cell-type-specific genes described in Fig. 1 but will retain the epithelial identity common to all vulval cells. It is unclear whether such a state has been observed in any of the mutants. Vulval cells in *lin-11* mutants lack most cell-type-specific expression but retain the ability to undergo some morphogenetic movements characteristic of vulval cells and thus may most closely resemble the ground state.

In other systems, analyses of coregulated genes have successfully identified "gene batteries" (42), sets of genes with common cis-regulatory elements that are coexpressed (for example, ref. 43). However, our understanding of vulval development is still limited, relative to the number of cell types

and the number of distinct stages that require different gene expression patterns. Consequently, within the relatively small number of functionally unrelated genes analyzed so far, genes are more likely to be regulated by distinct mechanisms. Thus, although gene batteries with multiple genes probably exist in this system, their analysis requires knowledge of more genes and a detailed understanding of which transcription factors regulate their expression.

We thank A. Fire (Stanford University, Stanford, CA) for providing GFP vectors, and I. Hope (University of Leeds, Leeds, U.K.) for providing the *C55C3.5*::*gfp* construct. We thank Ryan Baugh, Jennifer Sanders, Mihoko Kato, Steven Kuntz, Alok Saldanha, and reviewers for comments on the manuscript. We thank the *Caenorhabditis* Genetics Center (University of Minnesota, Minneapolis) for *C. elegans* strains. T.I. was supported by fellowship DRG-1646 from the Damon Runyon Cancer Research Foundation. Research was supported by the Howard Hughes Medical Institute, with which P.W.S. is an investigator.

1. Wood, W. B. (1988) *The Nematode Caenorhabditis elegans* (Cold Spring Harbor Lab. Press, Plainview, NY).
2. Kimble, J. (1981) *Dev. Biol.* **87,** 286–300.
3. Sternberg, P. W. & Horvitz, H. R. (1986) *Cell* **44,** 761–772.
4. Sternberg, P. W. (1988) *Nature* **335,** 551–554.
5. Yochem, J., Weston, K. & Greenwald, I. (1988) *Nature* **335,** 547–550.
6. Hill, R. & Sternberg, P. (1992) *Nature* **358,** 470–476.
7. Sulston, J. & Horvitz, H. R. (1977) *Dev. Biol.* **56,** 110–156.
8. Sharma-Kishore, R., White, J., Southgate, E. & Podbilewicz, B. (1999) *Development (Cambridge, U.K.)* **126,** 691–699.
9. Inoue, T., Sherwood, D., Aspöck, G., Butler, J., Gupta, B. P., Kirouac, M., Wang, M., Lee, P.-Y., Kramer, J. M., Bürglin, T., *et al*. (2002) *Mech. Dev.* **119S,** S203–S209.
10. Maduro, M. & Pilgrim, D. (1995) *Genetics* **141,** 977–988.
11. Hanna-Rose, W. & Han, M. (2002) *Dev. Biol.* **241,** 247–258.
12. Chang, S., Johnston, R. J., Jr. & Hobert, O. (2003) *Genes Dev.* **17,** 2123–2137.
13. Palmer, R. E., Inoue, T., Sherwood, D. R., Jiang, L. I. & Sternberg, P. W. (2002) *Dev. Biol.* **252,** 202–213.
14. Pettitt, J., Wood, W. & Plasterk, R. (1996) *Development (Cambridge, U.K.)* **122,** 4149–4157.
15. Burdine, R. D., Branda, C. S. & Stern, M. J. (1998) *Development (Cambridge, U.K.)* **125,** 1083–1093.
16. Chang, C., Newman, A. P. & Sternberg, P. W. (1999) *Curr. Biol.* **9,** 237–246.
17. Wang, M. & Sternberg, P. W. (2000) *Development (Cambridge, U.K.)* **127,** 5047–5058.
18. Stringham, E., Pujol, N., Vandekerckhove, J. & Bogaert, T. (2002) *Development (Cambridge, U.K.)* **129,** 3367.
19. Hwang, H. Y. & Horvitz, H. R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 14224–14229.
20. Colavita, A. & Tessier-Lavigne, M. (2003) *Science* **302,** 293–296.
21. Shen, K., Fetter, R. D. & Bargmann, C. I. (2004) *Cell* **116,** 869–881.
22. Rougvie, A. E. & Ambros, V. (1995) *Development (Cambridge, U.K.)* **121,** 2491–2500.
23. Newman, A. P., Inoue, T., Wang, M. & Sternberg, P. W. (2000) *Curr. Biol.* **10,** 1479–1488.
24. Freyd, G., Kim, S. & Horvitz, H. R. (1990) *Nature* **344,** 876–879.
25. Gupta, B. P. & Sternberg, P. W. (2002) *Dev. Biol.* **247,** 102–115.
26. Gupta, B. P., Wang, M. & Sternberg, P. W. (2003) *Development (Cambridge, U.K.)* **130,** 2589–2601.
27. Chamberlin, H. M., Palmer, R. E., Newman, A. P., Sternberg, P. W., Baillie, D. L. & Thomas, J. H. (1997) *Development (Cambridge, U.K.)* **124,** 3919–3928.
28. Ambros, V. & Horvitz, H. R. (1984) *Science* **226,** 409–416.
29. Ambros, V. (1989) *Cell* **57,** 49–57.
30. Bettinger, J. C., Euling, S. & Rougvie, A. E. (1997) *Development (Cambridge, U.K.)* **124,** 4333–4342.
31. Inoue, T., Oz, H. S., Wiland, D., Gharib, S., Deshpande, R., Hill, R. J., Katz, W. S. & Sternberg, P. W. (2004) *Cell* **118,** 795–806.
32. Deshpande, R., Inoue, T., Priess, J. R. & Hill, R. J. (2005) *Dev. Biol.* **278,** 118–129.
33. Kirouac, M. & Sternberg, P. W. (2003) *Dev. Biol.* **257,** 85–103.
34. Cui, M. & Han, M. (2003) *Dev. Biol.* **257,** 104–116.
35. Briscoe, J., Pierani, A., Jessell, T. M. & Ericson, J. (2000) *Cell* **101,** 435–445.
36. Muhr, J., Andersson, E., Persson, M., Jessell, T. M. & Ericson, J. (2001) *Cell* **104,** 861–873.
37. Vallstedt, A., Muhr, J., Pattyn, A., Pierani, A., Mendelsohn, M., Sander, M., Jessell, T. M. & Ericson, J. (2001) *Neuron* **31,** 743–755.
38. Aurelio, O., Boulin, T. & Hobert, O. (2003) *Development (Cambridge, U.K.)* **130,** 599–610.
39. Bateman, J. M. & McNeill, H. (2004) *Cell* **119,** 87–96.
40. Casares, F. & Mann, R. S. (2001) *Science* **293,** 1477–1480.
41. Lints, R., Jia, L., Kim, K., Li, C. & Emmons, S. W. (2004) *Dev. Biol.* **269,** 137–151.
42. Davidson, E. H. (2001) *Genomic Regulatory Systems: Development and Evolution* (Academic, San Diego).
43. Wenick, A. S. & Hobert, O. (2004) *Dev. Cell* **6,** 757–770.

CHAPTER III:

Rapid Identification of Transcriptional Regulatory Interactions Using Homology-Based Enhancer Identification and Yeast One-Hybrid Analysis

Ted O. Ririe*, Takao Inoue*, Shahla Gharib*, Vanessa Vermeirssen, Albertha J.M. Walhout, and Paul W. Sternberg

* Equal contribution

(Prepared for submission)

Abstract

Transcriptional regulation is a major mechanism directing development. However, demonstration of direct transcriptional regulation through identification of *cis*-regulatory elements and corresponding *trans*-acting factors is a significant bottleneck in the analysis of developmental processes. During the vulval development of *Caenorhabditis elegans*, seven terminally differentiated cell types are produced that express distinct sets of genes. A network of mutually interacting transcription factors controls the expression of these genes. While the identity of some transcription factors are known (COG-1, EGL-38, LIN-11, LIN-29, and NHR-67), genes and *cis*-regulatory elements directly regulated by them have not been identified. To identify such interactions, we carried out a screen for *cis*-regulatory elements that direct gene expression in vulval cell types. First, we identified candidate ~ 200 base pair (bp) enhancer regions from genomic sequences upstream of 17 vulva-expressing genes using sequence conservation between *C. elegans* and *C. briggsae*. Second, we tested the ~ 200-bp regions *in vivo* for enhancer activity. Of the 32 regions tested, 22 exhibited expression in some cell or tissue, nine of which showed vulval activity. Third, by using six of these regions as DNA "baits" in a yeast one-hybrid (Y1H) system, we identified transcription factors that bind three of these regions. Characterization of these factors indicated that the zinc finger transcription factor ZTF-16 plays a role in both vulval induction and vulval terminal cell-fate specification. These results demonstrate a set of procedures that can be used to efficiently identify small enhancers and regulators that interact with them.

Introduction

Tissue- and stage-specific gene activity is controlled through a variety of mechanisms, however, during development transcriptional regulation is the most common mechanism (1). Transcriptional regulation is achieved through *cis*-regulatory elements encoded in the genome in close proximity to the protein-coding region. These sites are bound by factors that determine the transcriptional state of the gene in a cell-type and developmental stage-specific manner. In general, *cis*-regulatory elements controlling transcription are organized into promoters and enhancers. Promoters reside immediately upstream of the transcription start site and the loading site for RNA polymerases. Enhancers reside at various distances from the promoter and activate transcription from the promoter. Although enhancers are most commonly located to the 5' of the promoter region, they are also found to the 3' of the gene, and in introns (1, 2).

Identification of *cis*-regulatory elements remains a major challenge. Classically, such elements were identified through the dissection of regulatory genomic sequences in successive steps using an *in vivo* (transgenic) assay system at every step to follow the regulatory activity (3, 4). Although powerful, such analyses are arduous and not easily amenable to a large-scale effort. The availability of genomic sequences from a number of closely related species provides a potential shortcut. Since genomic sequences in different species diverge over time due to mutation and genetic drift, any conservation suggests a purifying selection, which in turn indicates that that sequence has a function. Conserved sequences outside of the protein-coding region often correspond to *cis*-regulatory elements (1, 5, 6).

C. elegans is an excellent model system for analysis of development (7). Because nematode development is mostly invariant at the cellular level, this system offers single-cell resolution for investigating developmental processes (8). In particular, we have been analyzing late-stage development of the vulva as a model in which to study differentiation and gene expression control during organogenesis (9–11). The *C. elegans* vulva is derived from three precursor cells, P5.p, P6.p, and P7.p, which are induced to adopt vulval fates by a combination of EGF, Wnt, and Notch signaling. These three cells undergo three rounds of division near the end of the third larval (L3) stage to produce cells of seven distinct terminally differentiated cell types, each of which expresses a unique combination of genes. Several transcription factors are known that are required for the correct pattern of gene expression in these cells. These include COG-1 (Nkx6 homeodomain) (12), EGL-38 (PAX 2/5/8) (13, 14), LIN-11 (LIM homeodomain) (15, 16), LIN-29 (zinc finger) (18), and NHR-67 (nuclear hormone receptor) (19). These genes are linked by a complex network of mutual- and auto-regulatory interactions and constitute a gene regulatory network (GRN), which is required to establish cellular identity and direct cell-type-specific gene expression. Thus, the full understanding of the system requires the identification of multiple regulatory interactions.

Here we describe a series of screens to identify *cis*-regulatory interactions in the vulval GRN. Because previous analyses indicated that most enhancer elements are conserved between *C. elegans* and *C. briggsae* (5, 20, 21), we used conservation to identify 200-bp candidate enhancer regions, which we tested for activity *in vivo* using a fusion-PCR/biolistic transformation pipeline. Starting from seventeen genes, we tested 32 regions and found 22 enhancer regions, nine of which displayed vulval enhancer activity.

Next we subjected six vulval enhancers to yeast one-hybrid (Y1H) analysis. Using a library of known and putative *C. elegans* transcription factors we identified five transcription factors that bind three of these regions in yeast. We then determined if these factors play roles in the vulval GRN by screening the available mutants for possible vulval phenotypes and for effects on enhancer and reporter gene expression. One of the factors identified by the screen, ZTF-16, is expressed in all vulval cells, and a *ztf-16* deletion mutant exhibits defects in vulval morphology and gene expression. Together, our analysis suggests a set of parameters for single-step identification of *C. elegans* gene enhancer elements that can be successfully used in Y1H analysis.

Results

Selection of Candidate Enhancer Elements. We began our analysis with a set of 31 genes that are expressed in the vulval cells (Table 1). One additional gene, *grl-1*, was included because preliminary reports had suggested vulva expression. From these 32 genes, we selected 17 (boldface in Table 1) that have a probable ortholog in *C. briggsae*, and whose upstream regions have not been previously investigated. For each gene from *C. elegans* and *C. briggsae*, we took the sequence of the region 2-kb (kilobase = 1000 bases) upstream of the ATG start codon, or in cases where the upstream gene was closer than 2-kb from the ATG, up to the next gene. Upstream regulatory regions from *C. elegans* and *C. briggsae* were aligned using the clustalw program under default parameters.

To make later stages of the screen more uniform, we focused our initial efforts exclusively on ~ 200-bp candidate regions. Previous analyses of vulval enhancers found

that minimal functional enhancers are typically in the range of 50 to 150-bp (20, 21). To select 200-bp candidate regions from these alignments, we selected 150-bp windows that showed high degrees of conservation. High-scoring regions were "padded" on each end with 25 additional bases to generate 200-bp regions. We also allowed a maximum of 100-bp overlap between adjacent regions to ensure that highly conserved sections of the genome are always contained in one or more candidate regions (see Materials and Methods for full description of the algorithm). Some genes, including *pax-2*, returned no candidate regions of high score by this analysis, even though a visual scan of the alignment suggested that these promoters contained regions smaller than 150-bp that show a high degree of sequence similarity. To include *pax-2* in our analysis, we used a modified set of parameters in which 50-bp regions with a high degree of conservation were identified and padded with 75 bases on either end to generate 200-bp regions. Candidate regions identified from the 16 genes and *pax-2* were merged into a single list and sorted by the score (score = identical bases out of 150). We normalized the scores for *pax-2* (identical bases out of 50) by multiplying the score by three. From the merged list, we selected the 48 highest-scoring regions (named RGN7 through RGN54) for *in vivo* analysis of enhancer activities (Table 2).

***In Vivo* Analysis of Enhancer Activity.** To assay enhancer activity of candidate regions, we utilized the *delta-pes-10::GFP* construct that contains a naive promoter upstream of the *gfp* (green fluorescent protein) coding region (22–24). The *Δ-pes-10* promoter displays little to no transcriptional activity on its own, but can be enhanced in a tissue- and stage- specific manner when an enhancer is linked in *cis* to the promoter. One

caveat of this assay system is that some enhancers that activate their native promoters may not act on the *Δ-pes-10* promoter (25). However, previous analyses suggested that the *Δ-pes-10* system could detect most of the enhancer activity present upstream of *egl-17, cdh-3*, and *zmp-1* (20, 21). Thus, this system is likely to detect a significant fraction of enhancer activity present in these fragments.

To link our candidate enhancers to *Δ-pes-10::GFP*, we used fusion-PCR (Materials and Methods). Fusion-PCR products were then introduced into *C. elegans* by microparticle bombardment using *unc-119(ed4)* mutants and *unc-119(+)* plasmid DNA as a co-transformation marker. *unc-119* mutants display severe morphological and behavioral defects, and transformed (rescued) animals are easy to find and select for (26). Although recommended in some published protocols, we did not find it necessary to covalently link the transformation marker to the fusion-PCR product, possibly because we used a larger quantity of DNA per bombardment. In a pilot experiment, we found that nine out of ten independent transgenic lines made from co-bombardment of *unc-119* and an *enhancer::Δ-pes-10::GFP* fusion PCR product expressed GFP in a near-identical pattern (see Materials and Methods). Nevertheless, during our analysis of candidate regions, we found multiple cases where non-expressing and expressing lines were obtained using the same pool of fusion-PCR products. To reduce error from such non-expressing lines, multiple lines were analyzed for each region tested (typically two lines). Finally, we tested for and found no difference in expression pattern between transgenic lines made by bombardment and lines made by microinjection (see Materials and Methods).

Expression Patterns. Expression of GFP in transformed animals was examined first on a dissecting microscope equipped for epifluorescence (up to 66x magnification). In cases where vulval expression was observed or was suspected, animals were examined on a compound microscope (typically 1000x magnification).

We generated lines for 32 of 48 candidate regions selected, and found that nine of 32 candidate regions (RGN7, 12, 14, 16, 17, 21, 28, 30, and 44) drove expression in vulval cells (Table 2). Expression in other tissues was observed in 13 additional lines. This was expected, since the full-length upstream regulatory regions for some of the genes drive expression in tissues other than the vulva, and we had no way of determining *a priori* which conserved regions are likely to be vulva-specific. Altogether, about two thirds of the regions tested showed enhancer activity in some tissue. Regions with scores above 95 (= 95 identical bases out of 150) typically expressed GFP at a high level in a distinct spatial pattern (RGN7, 14, 16, 17, also non-vulval enhancers including RGN9, 11). In contrast, regions with lower scores (RGN21, 28, 30, 44) often expressed GFP at a low level (Figure 1). Although the spatial pattern was often difficult to discern because of the low level of expression, many of these lower-scoring lines appear to express GFP in all or a large number of vulval cell types. Because regions with scores above 95 often give stronger enhancer activity, we concentrated our efforts on regions with scores above 85 (RGN7-RGN28), and transgenes were not generated for some of the low-scoring regions. RGN12, which was identified using the 50-bp window, displayed a bright and distinct expression pattern (score of 34/50-bp identity = 102 normalized). This is significant in that the scan of the same alignment using a 150-bp window did not return a high-scoring candidate region; the highest scoring candidate region using the 150-bp

window was 78/150. Thus, in some cases, using a smaller window can identify regions missed using a larger window size. Finally, some regions with high scores (e.g., RGN8) did not exhibit obvious enhancer activity in multiple transgenic lines we examined.

Spatial and temporal patterns of vulval enhancer activity were examined in detail. Expression patterns observed varied in their similarity to the patterns generated by more-complete regulatory regions (compare Table 1 and Table 2). Some regions (e.g., RGN7 from *grl-4*) faithfully reproduced the expression of the more complete upstream regulatory region. Some regions had a more restricted expression pattern than the more complete regulatory region (e.g., RGN17 from *sqv-4*). This was expected, since multiple enhancer elements typically contribute to the total transcriptional activity of an upstream regulatory region. Similar restricted expression patterns have been observed for enhancers from *egl-17, cdh-3*, and *zmp-1*, in comparison to the complete upstream regulatory region (21). Finally, in some cases expression in unexpected vulval cell types was observed (e.g., RGN12 from *pax-2*). The observation of GFP expression in unexpected cell types suggests the presence of repressor elements in the regulatory region, although it is also possible that *Δ-pes-10* is failing to interact correctly with the enhancer.

Screen for 80-bp Regions with Enhancer Activity. Some of the vulval enhancers previously described are as small as 60-bp. To attempt to determine the minimal element required for enhancer activity, we subdivided each of five regions with vulval enhancer activity (RGN7, 12, 14, 16, 28) into three equal sub-regions of 80-bp. The three sub-

regions together cover the original ~ 200-bp region with about 20-bp of overlap between. These sub-regions were fused to *Δ-pes-10::GFP* and were tested *in vivo* as with larger fragments. We found that four sub-regions from three regions (RGN7, RGN14, RGN28) exhibited enhancer activity in vulval cells (Table 3). Sub-regions of RGN12 and RGN16 failed to drive expression, perhaps indicating that these ~ 200-bp regions were close to the minimal functional enhancer.

Yeast One-Hybrid Analysis. Using six of the RGNs with vulval enhancer activity as DNA "baits," we conducted yeast one-hybrid analysis (27). Briefly, each DNA "bait" was cloned upstream of two reporter genes, *HIS3* and *lacZ*, to create DNAbait::reporter constructs, and these were subsequently mated to an array of ~ 750 *C. elegans* transcription factors (80% of all predicted) (28). Two reporter constructs are used for each "bait" to reduce the rate of isolating false positives. Generally, only clones that score positively in both reporter assays (double positives) are considered high-confidence interactions. However, as some clones rarely interact with DNA, positive results for such clones (rare hits) were also pursued further. The Y1H screen identified six protein-DNA interactions involving five transcription factors and three RGNs (Table 4). That the Y1H analysis did not identify interactors for all six of the RGNs tested is not unexpected since this system exhibits a false negative rate of up to 67%, consequently, RGNs that did not return positive interactions may still interact with known transcription factors *in vivo* (27).

Among five transcription factors that were either double-positives or rare hits, we were able to obtain deletion alleles of *cey-1* (cold-shock/Y-box), the orphan nuclear hormone receptors *nhr-43* and *nhr-102,* and *ztf-16* (zinc finger). These mutations allowed us to test if these transcription factors affect transgene expression in *C. elegans*. We conducted two tests with each mutation, one with the relevant ~ 200-bp RGN reporter transgene and one with a more-complete reporter transgene. For example, to test the interactions between *cey-1* and RGN16 we constructed and scored the vulval expression of two strains, *cey-1(ok1805); RGN16::Δ-pes-10::GFP* [*syEx824*] and *cey-1(ok1805); F48B9.5::GFP* [*leEx1519*]. The results of the RGN strains are summarized in Table 5, the *ztf-16* strains are described below and in Figure 2 and Table 5, and all other strains are summarized in Supplemental Table 1. Mutations in three of the five Y1H-identified protein-DNA interactions we tested caused either up- or down-regulation of enhancer element reporter construct expression in vulval cell subsets. Both the *ztf-16* and *nhr-102* mutations affected expression of the *pax-2::GFP* [*guEx64*] reporter, and *ztf-16* affected expression of the *sqv-4::GFP* [*nEx851*] translational reporter. These results indicate that many protein-DNA interactions identified in yeast exhibit biological relevance in the worm.

ZTF-16 and the Vulval GRN. Of the transcription factors identified by the Y1H analysis, *ztf-16* possesses the most characteristics commensurate with a role as a component of the vulval GRN. *ztf*-16 encodes a DNA-binding protein with eight C2H2 zinc fingers with reported similarities to the Ikaros family of transcription factors (29). *ztf-16(ok1916)* is a 1005-bp deletion that disrupts predicted zinc-fingers three and four.

The vulvae of 25% of *ztf-16(ok1916)* mutant animals are under-induced (*n*=59) (Figure 3). An expression reporter for the *ztf-16* gene, *ztf-16::GFP* [*sIs12144*], was previously reported to be expressed in the pharynx, intestine, seam cells, and unidentified cells in the tail (30, 31). We found that this reporter was also expressed in vulval cells. In the wild type, all cells present in the adult vulva are generated by the onset of the L4 stage. Vulval morphogenesis and differentiation occur during the L4 stage which spans approximately 12 hours at 20˚C. In order to better describe dynamic expression patterns during vulval morphogenesis, we divided this timeframe into 10 sub-stages that we have termed L4.0–L4.9 (Supplemental Figure 1). Further analysis of *ztf-16::GFP* revealed expression in the VPCs and 2˚ vulval cells (vulA, vulB1, vulB2, vulC, and vulD), and faint expression in the 1˚ cells (vulE and vulF) (Figure 3). Expression in the 2˚ cells was present from the onset of L4 and its intensity decreases beginning at L4.4; expression in the 1˚ cells is weak and difficult to quantify, but it appears most strongly during later L4 (L4.6–L4.8).

ZTF-16 Regulates Expression of Vulval Genes. ZTF-16 is a rare interactor in the Y1H system, but it was isolated as both a *HIS*-only positive with RGN12 and a "double positive" with RGN17. RGN12 is a 199-bp putative *pax-2* expression enhancer element just 40-bp upstream of the *pax-2* ATG start site (Figure 4). The Pax2/5/8 transcription factor *pax-2* is the result of a recent duplication of *egl-38*; however, the two genes possess nonredundant characteristics (32, 33). Due to its proximity to the ATG start site, the RGN12 enhancer may also exhibit promoter activity.

In the vulva, the ~ 3.5-kb *pax-2::GFP* [*guEx64*] reporter consisting of the entire 5' upstream region, the first intron, and the majority of the second intron (Figure 4) is expressed in the vulE cells during very early L4 stage (L4.0–L4.1). While vulE expression disappears during L4.1, expression is initiated in the vulD cells, and this expression steadily increases throughout L4 (Figure 2). Vulval expression of the RGN12 reporter does not appear to be as tightly regulated as the 3.5-kb *pax-2::GFP*. Its expression is observed at a high frequency in vulC and vulD, and less frequently in vulB2, vulE, and vulF throughout the L4 stage (Table 5). The difference in these expression patterns points toward a modular system of gene regulation including both enhancers and repressors.

The intensity of the vulD expression of 3.5-kb *pax-2::GFP*, however, is enhanced more than threefold during L4.2 and L4.3 in a *ztf-16(ok1916)* mutant background (Figure 2). The spatiotemporal patterning of *ztf-16::GFP* expression appears to coincide with the effect of the *ztf-16* mutation on *pax-2::GFP* expression, indicating that ZTF-16 is a likely repressor of *pax-2* transcription (compare Figures 2 and 3).

RGN17 is a 174-bp fragment containing an enhancer of *sqv-4* expression, a gene that encodes a UDP-glucose 6-dehydrogenase (*Drosophila* Sugarless, Human UGDH). Both antibody staining and the translational *sqv-4::GFP* [*nEx851*] construct show expression in vulC, vulD, vulE, and vulF (34). The RGN17 enhancer element drives expression primarily in vulC and vulD. Another *sqv-4* enhancer element RGN28 drives expression in vulE and vulF. Consequently, these two enhancers likely represent two modules that work in concert to direct complete *sqv-4* expression. RGN17 expression was not affected by the *ztf-16(ok1916)* mutation (Table 5), but the frequency of *sqv-*

4::GFP expression in the vulE cells was up-regulated (Figure 2). These phenotypes suggest that in addition to a vulC- and vulD-specific enhancer, the RGN17 fragment also contains a vulE-specific repressor element responsive to ZTF-16. These results suggest dual roles for ZTF-16: initially during vulval induction and later during cell-fate specification.

Discussion

Using a combination of computational searches for conserved sequences and fast *in vivo* enhancer assays using fusion-PCR/microparticle bombardment, we tested 32 candidate enhancer regions from 17 different genes and found nine ~ 200-bp fragments that exhibited enhancer activity in our tissue of interest, the vulva. Furthermore, by subdividing five of these enhancers, we were able to identify four 80-bp fragments that exhibited enhancer activity in the vulva. These newly discovered enhancer elements, which originate from regulatory regions of *col-48*, *daf-6*, F48B9.5, *grl-4*, *pax-2*, and *sqv-4*, add to previously characterized vulval enhancers from *egl-17, cdh-3*, and *zmp-1*.

Currently, a basic blueprint of the gene regulatory network that directs vulval organogenesis is known. A set of transcription factors involved in the network have been identified, and many of their linkages with target genes have been mapped. There are at least two gaps, however, in our current understanding of this process. First, there are likely additional transcription factors involved in the network that have not yet been identified. Second, the directness of any of these interactions has not yet been demonstrated. Utilizing a subset of the enhancers identified through our approach, in

combination with a robust Y1H system, we were able to identify four transcription factors that may be involved in the network. In particular, we identified the zinc finger transcription factor ZTF-16 as a possible regulator of both vulval induction and terminal cell-fate specification. Also, the ZTF-16-*pax-2* protein-DNA interaction is the first direct linkage in the vulval gene regulatory network. Other methods, including chromatin immunoprecipitation (ChIP) or ChIP coupled with microarray analysis (ChIP-chip) may reveal others.

The localization of the vulva-specific enhancer activity from a large (> 1-kb) fragment to an 80 or 200-bp fragment is highly significant. From the computational side, this represents a major reduction in the amount of noise while searching for candidate transcription factor binding sites. Experimentally, mutational analysis is facilitated by the small size, since small enhancers with wild-type or mutated sequences can be synthesized *in vitro* (our results not shown).

Various aspects of the candidate regions we tested suggest that, as in other systems, *C. elegans* vulva gene regulation is modular and complex. As observed previously, enhancers we discovered reproduce the expression patterns of the complete regulatory region to varying degrees. Interestingly, in addition to high-scoring enhancers that exhibited specific bright patterns of expression, we also found enhancers that were generally less conserved (i.e., low-scoring) and appear to exhibit low-level enhancer activity without strong cell-type specificity. Some of these weak enhancers are from the same gene as strong and specific enhancers. For example, RGN30, which exhibited weak enhancer activity, is derived from the same gene (*sqv-4*) as RGN17, which had a strong cell-type-specific expression pattern. An interesting possibility is that these enhancers

interact in the endogenous *sqv-4* promoter, perhaps with RGN17 providing cell-type specificity and RGN30 interacting synergistically to increase expression level. Such examples are known from other systems (35).

Another interesting phenomenon is the case in which the enhancer activity is observed in an unexpected cell type. For example, RGN16 showed expression in vulA, vulB, and vulC cell types, even though the F48B9.5 upstream regulatory region from which this fragment is derived drives expression in vulC and vulD cell types only. Such cases suggest a possible repressor activity in the full-length region, and indeed, F48B9.5 regulatory region harbors RGN8, which is highly conserved between *C. elegans* and *C. briggsae*, but showed no detectable enhancer activity in multiple independent transgenic lines. Further experiments investigating the patterns of interactions among enhancer and conserved elements are needed to investigate and detect these phenomena.

Our work demonstrates a method by which small enhancer elements can be identified rapidly from a set of genes of interest. About 70% of the candidate regions we tested displayed enhancer activity in some tissue. Thus, *C. elegans*/*C. briggsae* sequence conservation is an excellent computational filter by which potentially interesting regions can be identified. Furthermore, our analysis suggests that regions with a score over 95/150 are likely to harbor strong cell-type-specific enhancers. Nevertheless, our approach is incomplete in that we found no vulval enhancers for the genes B0034.1, C55C3.5, *egl-26, grd-5, grl-1, grl-10, grl-12, grl-15,* and *lin-11*. There are a number of factors that might contribute to this problem: (1.) Enhancers for these genes may reside outside the 2-kb upstream region we analyzed. In particular we ignored potential intronic and 3' regions. Inclusion of these regions would not be difficult and we can use a set of

procedures essentially identical to the one we used to search these regions for candidate enhancer regions. (2.) These enhancers may not function in conjunction with *Δ-pes-10* naive promoter. (3.) These genes may contain enhancers whose minimal extent is over 200-bp, or our division of regions into candidate regions may have subdivided some of these into fragments too small to be active. (4.) The fusion-PCR bombardment protocol may be inefficient. (5.) The scoring scheme may need further optimization. For example, our result with the *pax-2* enhancer suggests that such further optimization is useful.

Acknowledgments

sqv-4::GFP [*nEx851*] was generated by HY Hwang and was a gift of HR Horvitz. *pax-2::GFP* [*guEx64*] was a gift of HM Chamberlin. Construction of *ztf-16::GFP* [*sIs12144*], a gift of DL Baillie, was supported by Genome British Columbia and Genome Canada (30). The *nhr-43(tm1381)* and *nhr-102(tm1542)* alleles were generated and kindly provided by S Mitani and the National Bioresource Project for the Nematode (Tokyo Women's Medical University). The *cey-1(ok1805)* allele was generated by the *C. elegans* Reverse Genetics Core Facility at the University of British Columbia, and the *ztf-16(ok1916)* allele was generated by the *C. elegans* Gene Knockout Project at the Oklahoma Medical Research Foundation, both of which are members of the International *C. elegans* Gene Knockout Consortium. These strains were provided by the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health—National Center for Research Resources.

T.O.R. was supported by National Institutes of Health Grant T32-GM007616. P.W.S. is an investigator with the Howard Hughes Medical Institute

Materials and Methods

General Methods and Strains. *C. elegans* strains were cultured on *E. coli* (OP50) at 20˚C using standard protocols (36). Alleles used in this study: LGII, *cey-1(ok1805),* a 539-bp deletion that disrupts the coding region of the predicted cold-shock DNA binding domain; LGIV, *nhr-43(tm1381),* a 5-bp insertion and 542-bp deletion that disrupts the putative nuclear hormone receptor domain; LGV, *nhr-102(tm1542),* a 382-bp deletion that removes the putative ATG start site*;* and LGX, *ztf-16(ok1916),* a 1005-bp deletion that disrupts zinc-fingers three and four. Transgenes used in this study: *pax-2::GFP* [*guEx64*], *sqv-4::GFP* [*nEx851*], *F48B9.5::GFP* [*leEx1519*], *ztf-16::GFP* [*sIs12144*], *RGN7::Δ-pes-10::GFP* [*syEx764*], *RGN7a::Δ-pes-10::GFP* [*syEx967*], *RGN9::Δ-pes-10::GFP* [*syEx765*], *RGN12::Δ-pes-10::GFP* [*syEx825*], *RGN14::Δ-pes-10::GFP* [*syEx834*], *RGN14b::Δ-pes-10::GFP* [*syEx968*], *RGN15::Δ-pes-10::GFP* [*syEx970*], *RGN16::Δ-pes-10::GFP* [*syEx824*], *RGN17::Δ-pes-10::GFP* [*syEx889*], *RGN21::Δ-pes-10::GFP* [*syEx807*], *RGN25::Δ-pes-10::GFP* [*syEx766*], *RGN27::Δ-pes-10::GFP* [*syEx877*], *RGN28::Δ-pes-10::GFP* [*syEx808*], *RGN28a::Δ-pes-10::GFP* [*syEx969*], *RGN30::Δ-pes-10::GFP* [*syEx826*], *RGN32::Δ-pes-10::GFP* [*syEx878*], and *RGN44::Δ-pes-10::GFP* [*syEx833*].

We reasoned that because vulval cells are specialized hypodermal cells, some of them may express collagen, a component of the *C. elegans* cuticle. We therefore

searched the Nematode Expression Pattern Database (http://nematode.lab.nig.ac.jp/) for *in situ* hybridization results of collagen genes and found that *col-7* and *col-48* may be expressed in the vulva. To test this, we constructed GFP transcriptional fusions of *col-7* and *col-48*, and found that they are expressed in subsets of the vulval cells.

Identification of Candidate Regions. *C. elegans* and *C. briggsae* genome sequences and orthology information were obtained from WormBase (data freeze WS140). Downloads of upstream regulatory sequence and alignment by clustalw (http://www.ebi.ac.uk/Tools/clustalw2/index.html) were done manually using a standard web browser and the MacVector software (http://www.macvector.com/). We wrote a short perl script to identify candidate regions and design primers. Sorting of candidate regions was done using Microsoft Excel (http://www.microsoft.com).

Candidate regions were selected as follows. First, the alignment was scanned with a 150-bp window and a score was assigned for each position of the window. The score was determined as the number of matched (identical) bases out of 150. Gaps (in *C. elegans* or *C. briggsae* sequences) were considered to be nonmatching. Next, the highest-scoring window was selected and extended by 25 bases on either side. This defined the first region for the given alignment. Note that 200-bp regions were first defined with respect to the coordinate system of the alignment, and subsequently translated to the equivalent *C. elegans* genomic sequence. Thus, actual regions tested are shorter than 200-bp in cases where the alignment contains a gap in the *C. elegans* sequence. After the first region was picked, window positions shifted from the first window by up to 100-bp on

either side were masked such that they could not be picked in subsequent scans. Then, the highest scoring unmasked region was picked as the second candidate region. This process was repeated until no regions with the score higher than 50 could be found.

The masking of 100-bp on either side of an already selected region ensures that no high-scoring region would be missed. If we mask out a wider region, two non-overlapping regions in close proximity can block a region in the middle from ever being selected. By using a 100-bp blocking radius, we ensure that in order to block a region in between two regions already picked, the two regions would need to be touching. Thus, all high-scoring areas of the DNA would be contained in one or more candidate regions according to this scheme.

The primers were designed such that the 5' end corresponded to the outer edges of the candidate regions. Lengths (or 3' ends) of the primers were adjusted by the program to generate primers with an approximate T_m of 60°C. The primer sequences were checked manually and some were edited by hand prior to ordering. The complementary sequence to the end of *Δ-pes-10::GFP* fragment was added to downstream (3') primer to allow fusion PCR to occur.

Fusion PCR. Primers were obtained from IDT (http://www.idtdna.com). To minimize the number of primers that need to be synthesized, we used the following scheme (Supplemental Figure 2). The *Δ-pes-10::GFP* fragment was amplified from a vector containing the fusion (pP97.78) using primers PGdo and PGup3 (see Supplemental Materials for primer sequences). The candidate enhancer region was amplified from *C.*

elegans genomic DNA using region-specific primers RGNxf and RGNxr. The 5' end of RGNxr primers contained a segment complementary to the PGup3 primer, allowing for fusion PCR. The second round of amplification was done with primers RGNxf and PGdi3, using a mixture of the *Δ-pes-10::GFP* fragment and the candidate enhancer region fragment as the template. The primer PGdi3 is inset by 40 bases from the PGdo primer. This scheme minimizes the number of region-specific primers needed (two per candidate region) while reducing the possibility of misprimed product by the use of partially nested reaction (by the use of PGdo/PGdi3 pair).

For amplification and fusion of ~ 200-bp fragments, we used the Expand Long Template PCR kit (Roche) for all amplifications. For amplification and fusion of 80-bp fragments, we used ExTaq polymerase (TaKaRa) for the amplification of enhancer fragments, and Expand Long Template PCR kit (Roche) for other reactions. 30 to 40 rounds of amplification were done for each step of the PCR fusion. Annealing temperatures of 50° to 60° were used for amplification of enhancer regions and adjusted when necessary.

Microparticle Bombardment. Microparticle bombardment was done as described with a few modifications (37). Some previously published protocols for transformation of *C. elegans* by microparticle bombardment call for covalent linkage of reporter fragment with a co-transformation marker. Because this involves ligation and cloning, this requirement imposed a potential bottleneck for the type of analysis we wished to carry out. To determine whether this was an absolute requirement, we carried out a pilot study

in which we co-transformed, using microparticle bombardment, fusion PCR products with the *unc-119(+)* co-transformation marker. For this analysis, we used fusions of enhancer regions RGN1 and RGN4 (see Supplemental Materials), derived from *nas-37* and *pax-2* upstream regions, respectively. We coated gold particles with the mixture of fusion PCR product and *unc-119(+)* plasmid DNA and bombarded them into *unc-119(ed4)* mutant animals. Animals rescued for the *unc-119* mutation were picked to establish lines and examined for GFP expression. We found that four of five lines generated from co-transformation of RGN1 PCR product and *unc-119(+)* plasmid expressed GFP in a similar pattern. The one line that did not express was derived from bombardment using 0.25 micrograms (μg) of *unc-119(+)* and RGN1 DNA per bombardment. In contrast, when 1 μg or greater quantity of *unc-119(+)* and PCR product (each) was used per bombardment, GFP expression was consistently observed in *unc-119* rescued lines. To test whether there were differences between lines generated by microparticle bombardment and microinjection, we generated 2 lines by coinjection of *unc-119(+)* and RGN4 PCR products into *unc-119(ed4)* animals. We found that the expression pattern was identical in 3 of 4 lines made by microparticle bombardment and the two lines made by injection. One line made from microparticle bombardment exhibited a slight difference in expression, with expression in vulC cells. Based on the results of these pilot experiments, we carried out most of our experiments by bombarding mixtures of about 1 μg of PCR products and 1 μg of *unc-119(+)* plasmids. In order to reduce the problem of nonlinkage, we typically analyzed at least two lines for each region to be tested. Each population of worms was bombarded 2 to 4 times with separate populations of DNA coated beads.

Yeast One-hybrid Analysis. RGN7, RGN12, RGN14, RGN16, RGN17, and RGN30 (for chromosomal coordinates see Table 2) were cloned upstream of two reporter genes, *HIS3* and *lacZ*, to create DNAbait::reporter constructs, and these were subsequently mated to an array of ~ 750 *C. elegans* transcription factors (80% of all predicted) using a protocol described previously (27).

Microscopy and Expression Analysis. Worms were anesthetized using 3 mM levamisole and observed using Nomarski optics (http://www.nomarski.com/). Images were taken with a monochrome Hamamatsu digital camera (http://www.hamamatsu.com/) and Improvision Openlab software (http://www.improvision.com/). GFP reporter expression was quantified using Openlab. Mean pixel intensity was measured in the brightest region of the cell of interest, normalized for exposure time, and then this value was divided by the normalized mean pixel intensity of a non-expressing cell to derive the relative GFP intensity.

References

1. Davidson EH (2006) The regulatory genome: gene regulatory networks in development and evolution (Academic, Burlington, MA ; San Diego) pp. xi, 289.
2. Lee TI & Young RA (2000) Transcription of eukaryotic protein-coding genes. Annu Rev Genet 34:77–137.
3. Hoch M, Schroder C, Seifert E & Jackle H (1990) *cis*-acting control elements for *Krüppel* expression in the *Drosophila* embryo. EMBO J 9(8):2587–2595.
4. Schier AF & Gehring WJ (1993) Analysis of a *fushi tarazu* autoregulatory element: multiple sequence elements contribute to enhancer activity. EMBO J 12(3):1111–1119.
5. Kuntz SG et al. (2008) Multigenome DNA sequence conservation identifies *Hox* cis-regulatory elements. Genome Res 18(12):1955–1968.
6. Li L, Zhu Q, He X, Sinha S & Halfon MS (2007) Large-scale analysis of transcriptional *cis*-regulatory modules reveals both common features and distinct subclasses. Genome Biol 8(6):R101.
7. Sternberg PW. Vulval development. (WormBook, Community TCeR, ed.).
8. Sulston JE & Horvitz HR (1977) Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol 56(1):110–156.
9. Inoue T et al. (2002) Gene expression markers for *Caenorhabditis elegans* vulval cells. Mech Dev 119 Suppl 1:S203–209.
10. Inoue T, Wang M, Ririe TO, Fernandes JS & Sternberg PW (2005) Transcriptional network underlying *Caenorhabditis elegans* vulval development. Proc Natl Acad Sci U S A 102(14):4972–4977.

11. Ririe TO, Fernandes JS & Sternberg PW (2008) The *Caenorhabditis elegans* vulva: a post-embryonic gene regulatory network controlling organogenesis. Proc Natl Acad Sci U S A 105(51):20095–20099.

12. Palmer RE, Inoue T, Sherwood DR, Jiang LI & Sternberg PW (2002) *Caenorhabditis elegans cog-1* locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. Dev Biol 252(2):202–213.

13. Chang C, Newman AP & Sternberg PW (1999) Reciprocal EGF signaling back to the uterus from the induced *C. elegans* vulva coordinates morphogenesis of epithelia. Curr Biol 9(5):237–246.

14. Rajakumar V & Chamberlin HM (2007) The *Pax2/5/8* gene *egl-38* coordinates organogenesis of the *C. elegans* egg-laying system. Dev Biol 301(1):240–253.

15. Gupta BP, Wang M & Sternberg PW (2003) The *C. elegans* LIM homeobox gene *lin-11* specifies multiple cell fates during vulval development. Development 130(12):2589–2601.

16. Freyd G, Kim SK & Horvitz HR (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene *lin-11*. Nature 344(6269):876–879.

17. Eisenmann DM & Kim SK (2000) Protruding vulva mutants identify novel loci and Wnt signaling factors that function during *Caenorhabditis elegans* vulva development. Genetics 156(3):1097–1116.

18. Newman AP, Inoue T, Wang M & Sternberg PW (2000) The *Caenorhabditis elegans* heterochronic gene *lin-29* coordinates the vulval-uterine-epidermal connections. Curr Biol 10(23):1479–1488.

19. Fernandes JS & Sternberg PW (2007) The tailless ortholog *nhr-67* regulates patterning of gene expression and morphogenesis in the *C. elegans* vulva. PLoS Genet 3(4):e69.

20. Cui M & Han M (2003) *Cis* regulatory requirements for vulval cell-specific expression of the *Caenorhabditis elegans* fibroblast growth factor gene *egl-17*. Dev Biol 257(1):104–116.

21. Kirouac M & Sternberg PW (2003) *cis*-Regulatory control of three cell fate-specific genes in vulval organogenesis of *Caenorhabditis elegans* and *C. briggsae*. Dev Biol 257(1):85–103.

22. Jantsch-Plunger V & Fire A (1994) Combinatorial structure of a body muscle-specific transcriptional enhancer in *Caenorhabditis elegans*. J Biol Chem 269(43):27021–27028.

23. Okkema PG & Fire A (1994) The *Caenorhabditis elegans* NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. Development 120(8):2175–2186.

24. Okkema PG, Harrison SW, Plunger V, Aryana A & Fire A (1993) Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. Genetics 135(2):385–404.

25. Okkema PG & Krause M. Transcriptional regulation. (WormBook, Community TCeR, ed.).

26. Maduro M & Pilgrim D (1995) Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics 141(3):977-988.

27. Deplancke B, Dupuy D, Vidal M & Walhout AJ (2004) A gateway-compatible yeast one-hybrid system. Genome Res 14(10B):2093–2101.

28. Vermeirssen V et al. (2007) Matrix and Steiner-triple-system smart pooling assays for high-performance transcription regulatory network mapping. Nat Methods 4(8):659–664.

29. Kwan CS, Vazquez-Manrique RP, Ly S, Goyal K & Baylis HA (2008) TRPM channels are required for rhythmicity in the ultradian defecation rhythm of *C. elegans*. BMC Physiol 8:11.

30. McKay SJ et al. (2003) Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. Cold Spring Harb Symp Quant Biol 68:159–169.

31. Hunt-Newbury R et al. (2007) High-throughput in vivo analysis of gene expression in *Caenorhabditis elegans*. PLoS Biol 5(9):e237.

32. Wang X, Greenberg JF & Chamberlin HM (2004) Evolution of regulatory elements producing a conserved gene expression pattern in *Caenorhabditis*. Evol Dev 6(4):237–245.

33. Park D, Jia H, Rajakumar V & Chamberlin HM (2006) Pax2/5/8 proteins promote cell survival in C. elegans. Development 133(21):4193–4202.

34. Hwang HY & Horvitz HR (2002) The *Caenorhabditis elegans* vulval morphogenesis gene *sqv-4* encodes a UDP-glucose dehydrogenase that is

temporally and spatially regulated. Proc Natl Acad Sci U S A 99(22):14224–14229.

35. Yuh CH, Li X, Davidson EH & Klein WH (2001) Correct expression of *spec2a* in the sea urchin embryo requires both Otx and other *cis*-regulatory elements. Dev Biol 232(2):424–438.

36. Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77(1):71–94.

37. Praitis V, Casey E, Collar D, & Austin J (2001) Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. Genetics 157(3):1217–1226.

38. Colavita A & Tessier-Lavigne M (2003) A Neurexin-related protein, BAM-2, terminates axonal branches in *C. elegans*. Science 302(5643):293–296.

39. Perens EA & Shaham S (2005) *C. elegans daf-6* encodes a patched-related protein required for lumen formation. Dev Cell 8(6):893–906.

40. Hanna-Rose W & Han M (2002) The *Caenorhabditis elegans* EGL-26 protein mediates vulval cell morphogenesis. Dev Biol 241(2):247–258.

41. Hao L, Johnsen R, Lauter G, Baillie D & Burglin TR (2006) Comprehensive analysis of gene expression patterns of *hedgehog*-related genes. BMC Genomics 7:280.

42. Shen K, Fetter RD & Bargmann CI (2004) Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. Cell 116(6):869–881.

43. Pettitt J, Wood WB & Plasterk RH (1996) *cdh-3*, a gene encoding a member of the cadherin superfamily, functions in epithelial cell morphogenesis in *Caenorhabditis elegans*. Development 122(12):4149–4157.

44. Burdine RD, Branda CS & Stern MJ (1998) EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in *C. elegans*. Development 125(6):1083–1093.

45. Bettinger JC, Euling S & Rougvie AE (1997) The terminal differentiation factor LIN-29 is required for proper vulval morphogenesis and egg laying in *Caenorhabditis elegans*. Development 124(21):4333–4342.

46. Bettinger JC, Lee K & Rougvie AE (1996) Stage-specific accumulation of the terminal differentiation factor LIN-29 during *Caenorhabditis elegans* development. Development 122(8):2517–2527.

47. Wagmaister JA, Gleason JE & Eisenmann DM (2006) Transcriptional upregulation of the *C. elegans* Hox gene *lin-39* during vulval cell fate specification. Mech Dev 123(2):135–150.

48. Davis MW, Birnie AJ, Chan AC, Page AP & Jorgensen EM (2004) A conserved metalloprotease mediates ecdysis in *Caenorhabditis elegans*. Development 131(23):6001–6008.

49. Stringham E, Pujol N, Vandekerckhove J & Bogaert T (2002) *unc-53* controls longitudinal migration in *C. elegans*. Development 129(14):3367–3379.

50. Wang M & Sternberg PW (2000) Patterning of the *C. elegans* 1° vulval lineage by RAS and Wnt pathways. Development 127(23):5047–5058.

Figure Legends

Figure 1. Signal intensity correlates with degree of sequence conservation. Lateral images of the developing vulva during the L4 stage. Fluorescence (left), Nomarski (center), and overlaid (right) images are shown. Expression patterns of three enhancer transgenes (RGN7, RGN21, and RGN 30) are shown along with their scores (= number of identical bases out of 150). The fragment with the highest score, RGN7, exhibits the brightest expression. The signal intensity and spatial resolution decreases in fragments with lower scores, RGN21 and RGN30, respectively.

Figure 2. Expression of *pax-2* and *sqv-4* are altered in a *ztf-16* mutant background. (A) *pax-2::GFP* expression in vulD cells during L4.2-L4.3 is enhanced in a *ztf-16* background. (B) Each image pair, Nomarski (top) and fluorescence (bottom), represents a single animal. For comparison of *pax-2::GFP* expression in wild type and *ztf-16* mutant backgrounds, five randomly-selected animals at each L4 sub-stage were imaged, and all images were collected with a 25 millisecond exposure time. (C) The percent of animals exhibiting *sqv-4::GFP* expression in vulE is increased in a *ztf-16* mutant background ($P = 0.0033$).

Figure 3. *ztf-16(ok1916)* exhibits a vulval phenotype and *ztf-16::GFP* is expressed in the vulva. The *ztf-16(ok1916)* allele exhibits an under-induced vulval phenotype in ~ 25% of animals. Only two VPCs are induced in animals shown. Expression of *ztf-16::GFP* [*sIs12144*] in the 2˚ cells (vulA-vulD) is present from the onset of the L4 stage,

and at mid-L4 the expression intensity appears to decrease. Weak expression is also present in the 1° cells; it is difficult to quantify, but it appears most strongly from mid- to late L4.

Figure 4. Gene models of F48B9.5, *sqv-4*, and *pax-2*. F48B9.5, *sqv-4*, and *pax-2* are shown with exons represented by turquoise rectangles. The relative positions of the nearest 5' upstream genes (gray gene models) are indicated for *sqv-4* and *pax-2*. F48B9.5's nearest 5' upstream gene, *nlp-7*, is located ~ 14-kb away and is not shown. Below each gene model the WABA *C. briggsae* alignments are shown (http://wormbase.org/). The light blue to dark blue regions code for low-to-high similarity regions, respectively. The green rectangles indicate the regions of sequence used to generate the reporter constructs referenced in this work. The numbered red rectangles show the positions of the predicted RGN enhancer elements. The expression patterns of each reporter are indicated. For reporters with vulval expression, cartoons of the L4 stage vulva with vulval cell nuclei represented by circles are used. Filled circles indicate the presense of expression and empty circles indicate absence of expression. (*) The *pax-2::GFP* expression in the vulE cells is present only during very early L4.

Supplemental Figure 1. Set of characteristics for scoring L4 sub-stages. Times indicated are approximate hours (hr) after L3/L4 molt. (L4.**0**) 0 hr; vulA, vulB, and vulE have divided, but vulC and vulF have not. (L4.**1**) 1.5 hr; vulC and vulF have now divided, and a narrow lumen has formed. (L4.**2**) Lumen has widened, and a prominent

kink (arrow) has formed between vulC and vulD. (L4.**3**) 4 hr; vulFs have separated and the apex of the lumen is flat and capped by the anchor cell (arrow). (L4.**4**) 5.5 hr.; utse is visible as a thin layer separating the vulval and uterine lumens (arrow). (L4.**5**) The side of the vulval lumen between vulC and vulD (arrow) is more curved. At sub-stage L4.4, the ventral side of vulD is convex, and the concave part against vulC is relatively small. At stage L4.5, the lumen wall against vulC and the ventral side of vulD forms a smooth curve. (L4.**6**) 8 hr; the "fingers" between vulB2 and vulC (arrow) are pointed ventrally. (L4.**7**) vulFs have migrated closer such that the lumen is narrowed in the dorsal section (arrow). The approximate cutoff is when the width of the channel is less than the width of the vulD nucleus. (L4.**8**) Lumen is partially collapsed. (L4.**9**) 10 hr; lumen is completely collapsed

Supplemental Figure 2. The scheme for construction of enhancer-*Δ-pes-10::GFP* constructs by fusion-PCR. The *Δ-pes-10::GFP* fragment was amplified from a vector containing the fusion (pP97.78) using primers PGdo and PGup3 (see Supplemental Materials for primer sequences). The candidate enhancer region was amplified from *C. elegans* genomic DNA using region-specific primers RGNxf and RGNxr. The 5' end of RGNxr primers contained a region complementary to the PGup3 primer, allowing for fusion PCR. The second round of amplification was done with primers RGNxf and PGdi3, using a mixture of the *Δ-pes-10::GFP* fragment and the candidate enhancer region fragment as the template. The primer PGdi3 is inset by 40 bases from the PGdo primer. This scheme minimizes the number of region-specific primers needed (two per

candidate region) while reducing the possibility of misprimed product by the use of partially nested reaction (by the use of PGdo/PGdi3 pair).

Figure 1

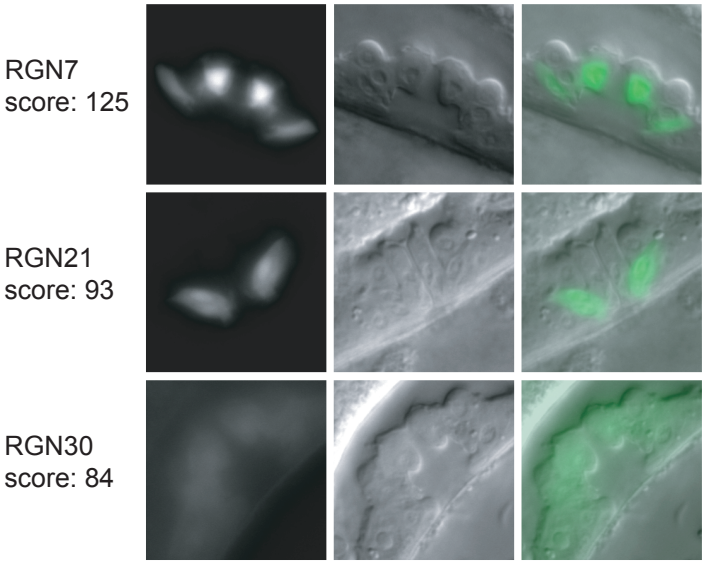


Figure 2

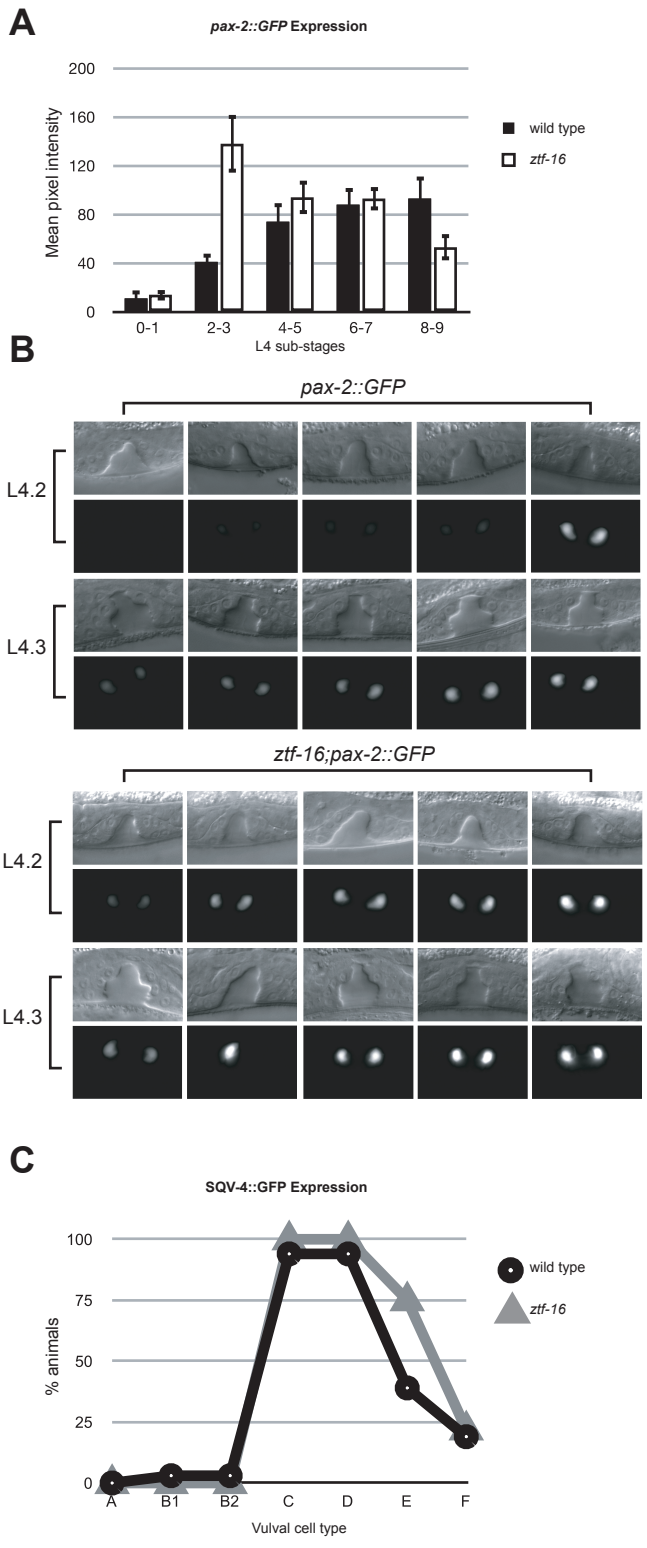

A
pax-2::GFP Expression
Mean pixel intensity
200
160
120
80
40
0
0-1
2-3
4-5
6-7
8-9
L4 sub-stages
wild type
ztf-16
B
pax-2::GFP
L4.2
L4.3
ztf-16;pax-2::GFP
L4.2
L4.3
C
SQV-4::GFP Expression
% animals
100
75
50
25
0
A
B1
B2
C
D
E
F
Vulval cell type
wild type
ztf-16

Figure 3

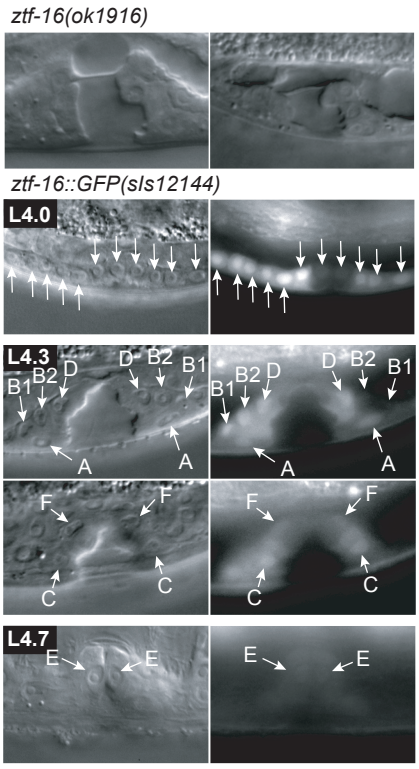
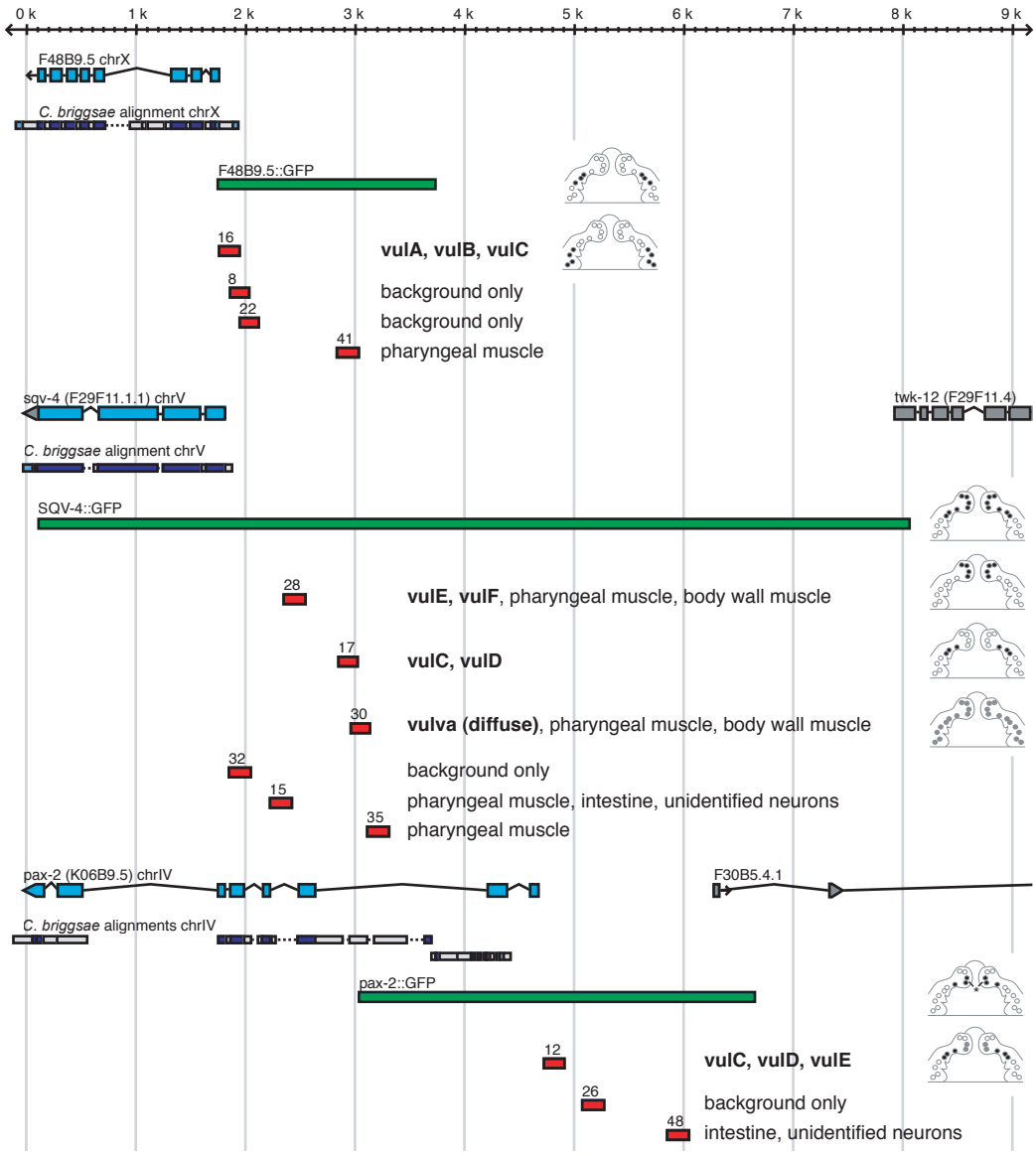
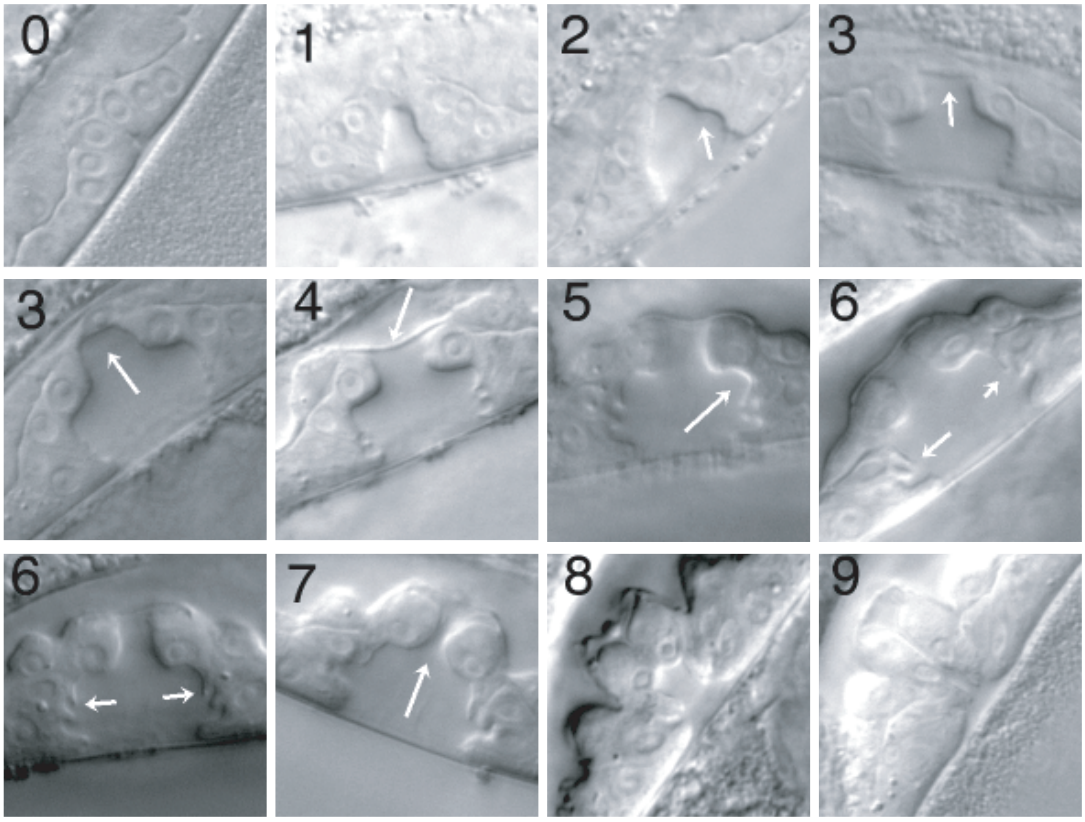


Figure 4


0 k
1 k
2 k
3 k
4 k
5 k
6 k
7 k
8 k
9 k
F48B9.5 chrX
C. briggsae alignment chrX
F48B9.5::GFP
16
vulA, vulB, vulC
8
background only
22
background only
41
pharyngeal muscle
sqv-4 (F29F11.1.1) chrV
twk-12 (F29F11.4)
C. briggsae alignment chrV
SQV-4::GFP
28
vulE, vulF, pharyngeal muscle, body wall muscle
17
vulC, vulD
30
vulva (diffuse), pharyngeal muscle, body wall muscle
32
background only
15
pharyngeal muscle, intestine, unidentified neurons
35
pharyngeal muscle
pax-2 (K06B9.5) chrIV
F30B5.4.1
C. briggsae alignments chrIV
pax-2::GFP
12
vulC, vulD, vulE
26
background only
48
intestine, unidentified neurons

Supplemental Figure 1



Supplemental Figure 2

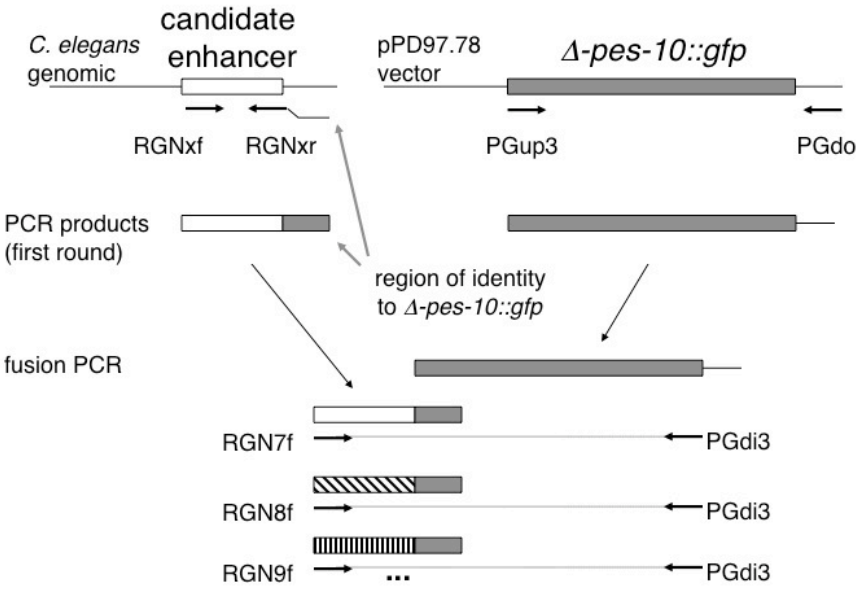


Table 1: Genes expressed in the *C. elegans* vulva

Gene	Expression (L4 or adult)	*C. briggsae* ortholog	Reference
B0034.1	E, F	CBG02550	Inoue, 2002
bam-2	F	CBG04211	Colavita, 2003
C55C3.5	F	CBG19878	Inoue, 2005
col-48	B1, B2, C, D*	CBG04191	see Materials and Methods
daf-6	E, F	CBG07607	Perens, 2005
egl-26	B, D, E	CBG06994	Hanna-Rose, 2002
F48B9.5	C, D	CBG14076	Gift of Hope Lab
grd-5	B1, B2, D	CBG01473	Hao, 2006
grd-12	C	CBG01466	Hao, 2006
grl-1†	none in vulva	CBG09376	
grl-4	A, B1, B2, D	CBG05832	Hao, 2006
grl-10	A, B1, B2	CBG20585	Hao, 2006
grl-15	B1, B2, C, D, E	CBG24285	Hao, 2006
lin-11	all vulval cells	CBG12236	Gupta, 2003
pax-2	D	CBG13498‡	Gift of Chamberlin Lab
sqv-4	C, D, E, F	CBG09668	Hwang, 2002
syg-2	E, F	CBG07568	Shen, 2004
cdh-3	C, D, E, F	CBG00038	Pettitt, 1996
ceh-2	B1, B2, C	CBG20241	Inoue, 2002
cog-1	all vulval cells	CBG20882	Palmer, 2002
col-7	A	none identified	see Materials and Methods
dhs-31	B1, B2, D	CBG03429	Inoue, 2002
egl-17	C, D (E, F in late L3)	CBG08160	Burdine, 1998
egl-38	F	none identified‡	Rajakumar, 2007
lin-29	all vulval cells	CBG02753	Bettinger, 1996; 1997
lin-3	F	none identified	Chang, 1999
lin-39	A	CBG09167	Wagmaister, 2006
nas-37	B	CBG01954	Davis, 2004
nhr-67	A, B, C	CBG06024	Fernandes, 2007
pepm-1	C, D, E, F	CBG04636	Inoue, 2002
unc-53	C	none identified	Stringham, 2002
zmp-1	A, D, E	CBG09053	Inoue, 2002; Wang, 2000

Genes in boldface were analyzed in this study. Orthology was based on WS140.
*L4 expression; adults express in all cells
†*grl-1* was included in this analysis because unpublished preliminary reports suggested *grl-1* expression in vulval cells. However, subsequent analyses found no evidence of *grl-1* expression in vulval cells (Hao et al., 2006).
‡PAX 2/5/8 genes *pax-2* and *egl-38* are apparently derived from a gene duplication event in the *C. elegans* lineage after the *C. elegans*/*C. briggsae* split.

Table 2. Candidate enhancer regions tested

Region	Gene	Score	Chrom. coordinates	Expression
RGN7	***grl-4***	**125**	**IV:7311259..7311452**	**vulA, vulB, vulD**
RGN8	F48B9.5	114	X:2149352..2149526	none
RGN9	*syg-2*	114	X:14654526..14654706	vulva/uterine muscle
RGN10	*grl-4*	111	IV:7311157..7311350	none
RGN11	*grl-1**	109	V:5533670..5533860	head neuron
RGN12	***pax-2***	**102†**	**IV:4204408..4204605**	**vulC, vulD**
RGN13	*bam-2*	100	I:2957958..2958142	none
RGN14	***col-48***	**99**	**I:3010421..3010614**	**vulB1, vulC, vulD,** hyp7, neuron, body wall muscle
RGN15	*sqv-4*	99	V:10661863..10662060	pharyngeal muscle, intestine, neurons
RGN16	**F48B9.5**	**97**	**X:2149252..2149439**	**vulA, vulB, vulC**
RGN17	***sqv-4***	**95**	**V:10662486..10662658**	**vulC, vulD**
RGN18	*syg-2*	95	X:14654708..14654874	uterus
RGN19	*grl-1**	94	V:5533789..5533958	none
RGN20	*syg-2*	94	X:14654369..14654563	pharyngeal muscle
RGN21	***grl-4***	**93**	**IV:7311360..7311544**	**vulva (diffuse),** head neurons
RGN22	F48B9.5	91	X:2149444..2149620	none
RGN23	*grl-4*	91	IV:7311501..7311695	vulva muscle, rectal epithelial
RGN24	*grd-5*	90	V:8618587..8618779	intestine, neurons
RGN25	*grl-4*	90	IV:7311056..7311253	vulva muscle, body wall muscle
RGN26	*pax-2*	90†	IV:4204763..4204956	none
RGN27	*syg-2*	89	X:14654626..14654769	vulva muscle, uterine muscle
RGN28	***sqv-4***	**86**	**V:10661980..10662177**	**vulE, vulF**, pharyngal muscle, body wall muscle
RGN29	*col-48*	84	I:3010289..3010487	NO LINES
RGN30	***sqv-4***	**84**	**V:10662590..10662767**	**vulva (diffuse)**, pharyngeal muscle, body wall muscle
RGN31	*sqv-4*	84	V:10662386..10662575	NO LINES
RGN32	*sqv-4*	84	V:10661486..10661684	none
RGN33	*grl-10*	83	V:7773889..7774085	none
RGN34	*bam-2*	82	I:2958056..2958247	NO LINES
RGN35	*sqv-4*	81	V:10662745..10662939	pharyngeal muscle
RGN36	*pax-2*	81†	IV:4205426..4205618	NO LINES
RGN37	*grl-10*	78	V:7773781..7773979	none
RGN38	*pax-2*	78†	IV:4205981..4206169	NO LINES
RGN39	*bam-2*	77	I:2958370..2958560	none
RGN40	*sqv-4*	77	V:10661750..10661948	NO LINES
RGN41	F48B9.5	76	X:2150334..2150528	pharyngeal muscle
RGN42	*grl-15*	76	III:12236244..12236442	intestine, coelomocytes
RGN43	*sqv-4*	76	V:10661295..10661492	NO LINES
RGN44	***daf-6***	**75**	**X:14887363..14887561**	**vulva (diffuse)**, lateral hypodermis
RGN45	*grl-10*	75	V:7773135..7773333	NO LINES
RGN46	*pax-2*	75†	IV:4205882..4206068	NO LINES
RGN47	*pax-2*	75†	IV:4205708..4205906	NO LINES
RGN48	*pax-2*	75†	IV:4205539..4205734	intestine, neurons
RGN49	*pax-2*	75†	IV:4205240..4205432	NO LINES
RGN50	*pax-2*	75†	IV:4204972..4205168	NO LINES
RGN51	B0034.1	74	II:5969843..5970039	NO LINES
RGN52	*daf-6*	74	X:14887906..14888097	NO LINES
RGN53	*grd-5*	74	V:8618686..8618857	NO LINES
RGN54	*grl-4*	73	IV:7311606..7311797	NO LINES

Regions in bold exhibit vulval expression. This work focused on RGNs with scores above 85 (RGN7-RGN28). Chromosomal coordinates are based on WormBase WS167 data set.
**grl-1* was included in this analysis because unpublished preliminary reports suggested *grl-1* expression in vulval cells. However, subsequent analyses found no evidence of *grl-1* expression in vulval cells (Hao et al., 2006).
†normalized from 50-bp score.

Table 3. Candidate 80-bp regions

Region	Gene	Chrom. coordinates	Vulva expression
RGN7a	***grl-4***	**IV:7311373..7311452**	**bright, vulB, vulC, vulD**
RGN7b	*grl-4*	IV:7311317..7311396	none
RGN7c	***grl-4***	**IV:7311260..7311339**	**faint (diffuse)**
RGN12a	*pax-2*	IV:4204526..4204605	none
RGN12b	*pax-2*	IV:4204468..4204547	none
RGN12c	*pax-2*	IV:4204409..4204488	none
RGN14a	*col-48*	I:3010421..3010500	none
RGN14b	***col-48***	**I:3010477..3010556**	**bright, vulB, vulC, vulD**
RGN14c	*col-48*	I:3010534..3010613	none
RGN16a	F48B9.5	X:2149360..2149439	none
RGN16b	F48B9.5	X:2149307..2149386	none
RGN16c	F48B9.5	X:2149253..2149332	none
RGN28a	***sqv-4***	**V:10662098..10662177**	**faint, vulE, vulF**
RGN28b	*sqv-4*	V:10662040..10662119	none
RGN28c	*sqv-4*	V:10661981..10662060	none

Regions in bold exhibit vulval expression. Chromosomal coordinates are based on WormBase WS167 data set.

Table 4. Results of the yeast one-hybrid analysis

DNA bait	Prey	Comment
RGN7 (*grl-4)*	None	DNA "bait" highly self-active
RGN12 (*pax-2)*	NHR-102	HIS positive only, but rare hit
	ZTF-16	HIS positive only, but rare hit
RGN14 (*col-48)*	CEH-22	HIS positive only
	SEM-4	HIS positive only
RGN16 (F48B9.5)	CEY-1	Double positive
	NHR-43	Double positive
RGN17 (*sqv-4*)	NHR-74	HIS positive only, but rare hit
	ZTF-16	Double positive, and rare hit
RGN30 (*sqv-4)*	None	

Table 5. A subset of Y1H positives effect enhancer expression.

Enhancer	Mutant	vulA, %	vulB1, %	vulB2, %	vulC, %	vulD, %	vulE, %	vulF, %	*n*
RGN12	+	0	0	7	100	70	20	7	37
RGN12	*nhr-102*	0	**15***	21	94	87	23	0	31
RGN12	*ztf-16*	2	6	6	97	73	5†	3	43
RGN16	+	96	78	39	39	8	0	0	51
RGN16	*cey-1*	100	90	**75‡**	**70§**	**55¶**	0	0	20
RGN16	*nhr-43*	87	91	**60ll**	51	3	0	0	65
RGN17	+	0	2	4	98	85	1	4	46
RGN17	*ztf-16*	0	5	0	95	75	5	0	20

Worms were grown and scored at 20˚C. *P* values were calculated using Fisher's exact test, comparing the fraction of cells, of the indicated cell type, that express the transgene versus those that do not and by comparing the expression of the transgene in a mutant versus a wild type background. $P < 0.05$ considered significant and represented by bold type. *, $P = 0.0163$; †, $P = 0.0734$; ‡, $P = 0.0086$; §, $P = 0.0334$; ¶, $P < 0.0001$; ll, $P = 0.0392$.

Supplemental Table 1. Relative GFP intensity of F48B9.5 and *pax-2* expression in vulD during the L4 sub-stages

		Sub-stage				
Reporter	**Mutation**	**L4.0-L4.1 (SEM)**	**L4.2-L4.3 (SEM)**	**L4.4-L4.5 (SEM)**	**L4.6-L4.7 (SEM)**	**L4.8-L4.9 (SEM)**
F48B9.5	+	1.08 (0.05)	1.15 (0.07)	1.87 (0.32)	5.47 (0.31)	8.66 (0.57)
F48B9.5	*cey-1*	0.73 (0.04)	0.74 (0.04)	2.32 (0.27)	5.79 (0.44)	7.19 (0.50)
F48B9.5	*nhr-43*	0.84 (0.03)	1.43 (0.22)	1.96 (0.30)	5.43 (0.25)	7.51 (0.31)
pax-2	+	11 (6)	41 (7)	74 (15)	88 (13)	93 (18)
pax-2	*nhr-102*	4 (1)	32 (5)	65 (11)	54 (14)	69 (14)
pax-2	*ztf-16*	14 (4)	138 (23)	94 (13)	93 (9)	53 (10)

CHAPTER IV:

Conclusions

During the development of multicellular organisms individual cells assume vastly divergent structural and functional roles. These roles range from the transmission of chemical signals to the absorption of nutrients to the reinforcement of anatomical shape. Underlying this wide-ranging diversification are gene regulatory networks (GRNs); logic maps that describe the intricate connections between signaling molecules, transcription factors, *cis*-regulatory modules, and target genes. Mapping these networks facilitates a better understanding of cell-fate patterning and enables us to make predictions about developmental processes.

While GRNs are being investigated in other systems and developmental processes, the *C. elegans* vulva provides us with an excellent opportunity for studying a relatively simple example of organogenesis. As shown in this thesis, the worm is a highly tractable system for conducting genetic, molecular biological, and biochemical analysis; and while the majority of networks under scrutiny in other systems focus on processes occurring during embryogenesis, the vulva provides us with a post-embryonic model. Also, its level of complexity strikes an elegant balance between the almost overwhelming complexity of mammalian organogenesis and the relative simplicity of single cell-type differentiation and specification.

While previous work had identified several aspects of transcriptional regulation during late-stage vulval organogenesis, in Chapter 2 I highlight my work in the initial consolidation and mapping of this data into a cohesive network diagram. Classical forward genetic screens (i.e., by EMS mutagenesis) are biased towards isolating mutants with strong phenotypes that act through nonredundant mechanisms. Since previous mutagenesis screens likely identified the majority of genes in this category, I utilized a

reverse genetic screen to identify additional genes, including those with more subtle effects, involved in late stage vulval organogenesis. Utilizing RNA interference (RNAi) to systematically disrupt the function of 836 known and putative *C. elegans* transcription factors in a *zmp-1::GFP* (zinc metalloproteinase) background, we identified *nhr-67* (ortholog of *Drosophila tailless*) and *nhr-113* (orphan nuclear hormone receptor) as positive regulators of *zmp-1* expression in vulA cells. *nhr-113*'s role in vulval organogenesis is possibly limited to regulation of differentiation in the vulA cell type. This is an excellent example of a gene involved in a subtle aspect of the vulval GRN that was missed by classical screening methods (Chapter 2).

A key component of GRNs are the *cis*-regulatory modules that mediate the spatiotemporal expression patterns of genes. Using Cistematic, a software package that incorporates the phylogenetic-footprinting and motif-finding capabilities of MEME, AlignACE, Co-Bind, and Footprinter for identifying *cis*-regulatory elements (1), we identified four conserved motifs upstream of the ZMP-1 translational start site. By deleting these motifs in the *zmp-1::GFP* transgene, three of the four motifs were revealed to possess *cis*-regulatory activity. One of these motifs includes the first example, in the vulva, of an enhancer that drives expression in unexpected cells, thus indicating the possible existence of a vulva-affiliated repressor element (Chapter 2).

Furthering our analysis of *cis*-regulation of transcription during vulval development, we utilized bioinformatic techniques to identify nine ~ 200-bp vulva-specific enhancer elements associated with six genes (Chapter 3). Yeast one-hybrid analysis of six of these elements revealed five transcription factors that bind these regions in yeast. Additional analysis of these transcription factors identified ZTF-16 as a

regulator of cell-fate specification in the vulva. At present, it is not known if any of the linkages in our current vulval GRN model are direct. The ZTF-16/*pax-2* interaction, however, may represent a direct regulatory connection.

Further characterization of the interactors identified by the yeast one-hybrid screen, in particular *ztf-16*, is necessary for defining their roles, if any, in vulval differentiation. Using our knowledge of the provisional vulval GRN it would be informative to conduct a series of pairwise tests with the *ztf-16* mutant and *ztf-16::GFP* reporter strains. *cog-1*, *egl-38*, *lin-11*, *lin-29*, and *nhr-67* are the major regulators of vulval differentiation. By assaying *ztf-16* expression patterns in backgrounds mutant for these genes, and conversely, by assaying their expression patterns in a *ztf-16* mutant background, we could elucidate *ztf-16*'s potential position and role in the regulatory network. Determing whether the protein-DNA interactions identified in yeast occur in the worm is also intriguing as we do not know whether any of the linkages in the current GRN are direct. To determine whether ZTF-16 directly regulates *pax-2* and *sqv-4* expression, additional experiments could be performed. Electrophoretic mobility shift assays (EMSA) with a variety of DNA probes (full-length, subdivided, and mutated RGN sequences) may verify protein-DNA interactions *in vitro* and fine-tune our knowledge of the actual binding sites. For *in vivo* analysis of binding, a series of transgenic reporters could be constructed. For example, the ~ 200-bp RGN sequences could be deleted from the more-complete reporter constructs and expression patterns between the two could be compared. The spatial and temporal expression patterns of these mutated transgenes would then need to be compared in wild-type and mutant backgrounds.

Considering that there are 934 predicted transcription factors in the *C. elegans* genome (2) and that phenotypes have only been identified for a fraction of these, it is likely that there are other transcription factors that are involved in the vulval GRN. These factors may act redundantly, or in parallel with each other to direct more subtle aspects of vulval differentiation, and consequently may have been missed in previous screens. Our strategy described in Chapter 3 for identifying vulval enhancer elements was limited to the region 2 kb upstream of the ATG start site. Enhancer elements, however, have been shown to exist outside of this region (3). It would be relatively simple to adapt our procedures to search for elements further upstream, in 3' regions, and in introns. Yeast one-hybrid analysis could again be utilized for identifying additional protein-DNA interactions.

Microarray technology has been successfully used for expanding the knowledge of several gene regulatory networks as they effectively determine the gene expression profiles of specific cells and tissues (4–6). Additionally, an adaptation of this technology, the combination of chromatin immunoprecipitation and microarrays (ChIP-chip), has been useful for identifying protein-DNA interactions (3, 6, 7). There are obstacles to modifying this technology for use in the vulval model. Mainly, due to the worm's small size, the vulva cannot be easily dissected from the animal for the harvesting of the material needed for these assays. Application of micromanipulator technology coupled with mRNA amplification protocols, however, holds promise as an effective way of overcoming this barrier. The combination of these approaches will likely extend our understanding of the vulval gene regulatory network. Cross-comparison of a more-thoroughly defined vulval network with genomic networks in other systems can elucidate

evolutionary diversification of the mechanisms controlling cellular differentiation, and enhance our knowledge of developmental processes in general.

References

1. Mortazavi A, Leeper Thompson EC, Garcia ST, Myers RM & Wold B (2006) Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. Genome Res 16(10):1208–1221.
2. Reece-Hoyes JS et al. (2005) A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol 6(13):R110.
3. Mockler TC et al. (2005) Applications of DNA tiling arrays for whole-genome analysis. Genomics 85(1):1–15.
4. Brunskill EW et al. (2008) Atlas of gene expression in the developing kidney at microanatomic resolution. Dev Cell 15(5):781–791.
5. Biemar F et al. (2006) Comprehensive identification of Drosophila dorsal-ventral patterning genes using a whole-genome tiling array. Proc Natl Acad Sci U S A 103(34):12763–12768.
6. Zeitlinger J et al. (2007) Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes Dev 21(4):385–390.
7. Liu XS (2007) Getting started in tiling microarray analysis. PLoS Comput Biol 3(10):1842–1844.