GENETIC REGULATION OF CAENORHABDITIS ELEGANS POST-EMBRYONIC DEVELOPMENT INVOLVING THE TRANSCRIPTION FACTORS EGL-38, VAB-3, AND LIN-14

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Ryan W. Johnson

*****

The Ohio State University
2008

Dissertation Committee:
Professor Helen M. Chamberlin, Advisor
Professor Heithem M. El-Hodiri
Professor Stephen A. Osmani
Professor Amanda S. Simcox

Approved by

______________________________
Advisor
Graduate Program in Molecular Genetics
ABSTRACT

Sequence-specific transcription factors are crucial to generating the gene expression patterns that drive the specification, morphogenesis, and physiology of organs and tissues. In order to better understand how organ form and function are orchestrated by transcription factors, we must better understand the genetic inputs and outputs of these critical regulators. In contrast to the complexity of higher organisms, the nematode Caenorhabditis elegans provides a powerful model system in which to address important questions in organ development and function. In this work, I have utilized C. elegans to characterize the genetic networks and organ functions of three post-embryonically functioning transcription factors.

The Pax family of transcription factors is highly conserved across animal species, and controls the development of multiple tissues and organs during development. The Pax-6 subclass of Pax genes is essential for the development of eyes and other sensory organs in species ranging from planaria to mice. In C. elegans males, two sensory mating structures, the copulatory spicules and the post-cloacal sensilla, are formed from stereotyped divisions of the two post-embryonic blast cells, B.a and Y.p, respectively. A C. elegans pax-6 transcript, vab-3, is necessary for the development of these sensory structures. Using a green fluorescent protein (GFP)-based vab-3 transcriptional reporter,
I found that expression is restricted to the sensory organ lineages of B.a and Y.p. Transcription of vab-3 in the tail region of the worm requires the Abdominal-B homeobox gene, egl-5. Opposing this activation, a transcription factor cascade and a Wnt signaling pathway each act to restrict vab-3 expression to the appropriate cell lineages.

Another C. elegans Pax gene, egl-38, is required for the development of the egg-laying system and rectum. However, few EGL-38 target genes are known. Using gene expression microarrays, we cross-referenced microarray data from an inducible EGL-38 strain and two egl-38 mutants that disrupt protein function in a tissue-preferential manner to identify potential tissue-specific EGL-38 target genes. One set of genes from this analysis was validated using GFP reporter transgenes. Most of these genes are expressed in egl-38-dependent tissues, and many display egl-38 dependence. In addition to the identification of target genes, this work revealed enrichments in gene classes that play a role in innate immunity. Consistent with this, we discovered a novel immune function for egl-38. We found that the gene activities of egl-38 and three egl-38-responsive genes from our validation set are associated with increased infection by the pathogenic bacterium M. nematophilum. However, we also show that egl-38 does not impact infection by a different pathogen, S. marcescens.

While Pax genes regulate spatial tissue/organ identity, some transcription factors regulate temporal identity. In C. elegans, heterochronic genes function to ensure the precise timing of stage-specific developmental events. I positionally cloned a novel missense allele of the heterochronic transcription factor LIN-14, and revealed a previously undiscovered ability of this protein to solely affect late larval development. lin-14(sa485) hermaphrodites exhibit asynchrony between vulval and gonadal
morphogenesis and maturation. Further, \textit{lin-14(sa485)} preferentially disrupts the timing of vulval cell morphogenesis, but not cell division. I also show that terminal differentiation of a uterine cell type is delayed in \textit{lin-14(sa485)} mutants.

Together, this work has identified the regulatory logic defining expression pattern of one \textit{Pax} gene, uncovered many candidate transcriptional targets of another Pax factor, and identified a novel mode of action for a temporal transcriptional regulator. These projects illustrate genetic principles applicable to not only \textit{C. elegans}, but to higher organisms by virtue of the conservation of \textit{Pax} genes and biological functions such as developmental timing control.
DEDICATION

To my wife Marisa and my mother Maryanna.
For their love, support, encouragement, and inspiration, I am eternally grateful.
ACKNOWLEDGMENTS

First and foremost, I would like to recognize my advisor Dr. Helen Chamberlin. Her guidance, patience, and encouragement have been a strong and positive influence throughout my graduate career. Her scientific insights and disciplined nature have helped me to learn from successes and failures alike, and to think independently. I am indebted to Helen for the knowledge she has shared with me over the years, and I am very grateful to have had such an excellent mentor.

I would also like to thank members of my Ph.D. committee: Dr. Heithem El-Hodiri, Dr. Stephen Osmani, and Dr. Amanda Simcox. Over the years, they have been generous to share their time, and have provided insightful suggestions. I feel better prepared as a young scientist due to the wisdom and experiences they have shared.

This research was supported through grants from the National Science Foundation and the National Institutes of Health.

I extend my thanks to other faculty members in the department of Molecular Genetics for their support and guidance. I would especially like to acknowledge Dr. Russell Hill, whose technical support and informative discussions have helped me grow professionally.
I am grateful to all current and former members of the Chamberlin lab for their professional fellowship and personal friendship, including Kristin Armstrong, Jennifer Greenberg, Dr. Hongtao Jia, Xin Li, Dr. Donha Park, Dr. Vandana Rajakumar, Dr. Sama Sleiman, Stephen Sewell, Jesse Sloan, Rong-jeng Tseng, Ashwin Uttam, Dr. Xiaodong Wang, Guojuan Zhang. Further, I would like to my colleagues in the Hill, Simcox, and Chang labs for their discussions and comraderie, including Dr. Rashmi Deshpande, Dr. John Butchar, Dr. Tom Jacobsen, Litty Paul, Steven Justiniano, and Dr. Rosemary Hage.

I thank Dr. Hongtao Jia for his collaborative efforts on the egl-38 project. His technical expertise and pragmatism were invaluable to my experimental contributions, and I would also like to recognize his aid in topical discussions and writing.

I would like to acknowledge my gratitude to Debbie Dotter, Marsha Hronek, and Jessie Siegman, for their tireless and cheerful administrative support.

I would like to thank all my family members (mom, dad, Ron, Mike, and all the extended siblings) for their support, patience, and perspective through the years. I am particularly grateful to my mother, who always put my needs ahead of her own, and without whom I would not and could not have become who I am as a person.

Finally, I thank my wife Marisa and my son Dylan. They are the inspiration for everything I set out to accomplish, and their love has helped me through these last difficult years.
VITA

July 12, 1978…………………………..Born – Akron, Ohio.

1997-2000……………………………..B.S. Biology, cum laude
University of West Florida

2001-2008……………………………..Teaching and Research Associate
The Ohio State University.

PUBLICATIONS


FIELDS OF STUDY

Major Field: Molecular Genetics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>VITA</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xv</td>
</tr>
<tr>
<td><strong>CHAPTER 1 INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 The role of specification and morphogenesis in animal development</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Transcriptional regulation of development</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Pax transcription factors are critical regulators of development in many species</td>
<td>5</td>
</tr>
<tr>
<td>1.4 The nematode C. elegans as a model to study Pax genes, specification, and morphogenesis</td>
<td>8</td>
</tr>
<tr>
<td>1.5 Defining genetic networks of two Pax genes and an additional transcriptional regulator of C. elegans post-embryonic development</td>
<td>10</td>
</tr>
</tbody>
</table>
CHAPTER 2 POSITIVE AND NEGATIVE REGULATORY INPUTS RESTRICT PAX-6/VAB-3 TRANSCRIPTION TO SENSORY ORGAN PRECURSORS IN C. ELEGANS

2.1 INTRODUCTION

2.2 MATERIALS AND METHODS

2.2.1 Genetic strains and strain construction

2.2.2 Microscopy

2.2.3 vab-3::gfp construction and expression

2.3 RESULTS

2.3.1 A transcriptional reporter recapitulates the VAB-3 expression pattern in male sensory structure precursors

2.3.2 A transcriptional cascade restricts vab-3 expression within the male hindgut

2.3.3 Wnt signaling restricts vab-3 transcription within the B lineage

2.3.4 vab-3 expression in the male tail requires the HOM-C gene egl-5

2.3.5 Ectopic vab-3 expression in non-sensory organ precursors requires egl-5

2.4 DISCUSSION

2.4.1 A domain refinement model for establishment of pax-6/vab-3

2.4.2 pax-6/vab-3 transcription is observed in sensory organ precursor cells

2.4.3 Regional induction of pax-6/vab-3 expression is mediated by the HOM-C gene egl-5

2.4.4 Negative regulatory inputs restrict pax-6/vab-3 transcription to sensory organ precursor cells: a transcriptional network

2.4.5 Negative regulatory inputs restrict pax-6/vab-3 transcription to sensory organ precursor cells: Wnt signaling

2.4.6 A domain refinement model for establishment of pax-6/vab-3

CHAPTER 3 EGL-38 TARGET GENE IDENTIFICATION VIA GENE EXPRESSION MICROARRAYS AND CHARACTERIZATION OF EGL-38 IMMUNE FUNCTION

3.1 INTRODUCTION
3.2 MATERIALS AND METHODS

3.2.1 Genetic strains and strain construction

3.2.2 RNA preparation

3.2.3 Microarray hybridization and data analysis

3.2.4 GFP reporter transgene construction

3.2.5 cDNA cloning and RNA interference

3.2.6 S. marcescens killing assays

3.3 RESULTS

3.3.1 Identification of potential EGL-38 target genes using gene expression arrays

3.3.2 Mining for prospective EGL-38-responsive genes by a microarray cross-referencing approach

3.3.3 Cell-specific expression of candidate genes is responsive to egl-38

3.3.4 egl-38 influences the susceptibility of C. elegans to infection by the gram-positive M. nematophilum, but not the gram-negative S. marcescens

3.3.5 Candidate EGL-38 target genes display egl-38-dependent rectal expression

3.3.6 A subset of egl-38-dependent genes mediate susceptibility to M. nematophilum infection

3.4 DISCUSSION

3.4.1 Identification of EGL-38-responsive genes using expression microarrays

3.4.2 A novel immune function for egl-38 does not influence S. marcescens infection

3.4.3 At least three downstream genes mediate the function of EGL-38 in M. nematophilum infection

3.4.4 Implications for innate immunity in vertebrates

CHAPTER 4 ISOLATION OF A NOVEL ALLELE OF LIN-14 AND CHARACTERIZATION OF THE RESULTING HETEROCHRONIC PHENOTypes IN THE C. ELEGANS EGG-LAYING SYSTEM

4.1 INTRODUCTION

4.2 MATERIALS AND METHODS

4.2.1 Genetic strains and strain construction

4.2.2 lin-14(sa485) mapping and cosmid rescue

4.2.3 lin-14 complementation and sequencing

4.2.4 lin-14(sa485) hermaphrodite anatomy and microscopy

4.2.5 ida-1::gfp studies

4.2.6 lin-14(n179) temperature shift experiments
4.3 RESULTS

4.3.1 Identification of a novel allele of the heterochronic gene lin-14

4.3.2 lin-14(sa485) mutants exhibit asynchrony between gonadal and vulval development

4.3.3 Specification of uterine uv1 cells is delayed in lin-14(sa485) mutants

4.3.4 The Evl-D phenotype results from an early defect in lin-14 function

4.4 DISCUSSION

4.4.1 Identification of a new allele of the heterochronic gene lin-14

4.4.2 An L4 function for the L1-specific transcription factor LIN-14

4.4.3 lin-14(sa485) mutants exhibit asynchrony between gonadal and vulval development

4.4.4 The cellular role of lin-14 in regulating temporal synchrony of the developing egg-laying system

CHAPTER 5 DISCUSSION

5.1 Project relationships

5.2 Transcriptional regulation of tissue development and function: lessons from egl-38

5.3 The role of temporal interactions in Pax-dependent tissues: lessons from lin-14

5.4 Genetic controls of transcription factor expression: lessons from vab-3

5.5 Significance of transcription factor specificity and temporal regulation models

5.6 Transcriptional control over C. elegans immunity: lessons from everyone

APPENDICES

A. Oligonucleotide primers

B. C. elegans genetic strains

C. lin-14(sa485) mutants exhibit abnormal B cell division and spicule formation

D. egl-38 mutants show reduced infection by M. nematophilum

LIST OF REFERENCES
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.</td>
<td><em>egl-5</em> is required for <em>vab-3</em> expression in hindgut blasts.</td>
</tr>
<tr>
<td>3.1.</td>
<td>Genes showing reciprocal response to <em>egl-38(n578)</em> and <em>hsp::egl-38</em> activity.</td>
</tr>
<tr>
<td>3.2.</td>
<td>Expression pattern summary for candidate EGL-38 target genes.</td>
</tr>
<tr>
<td>3.3.</td>
<td>Enrichment of <em>egl-38</em>-dependent expression among candidate genes.</td>
</tr>
<tr>
<td>3.4.</td>
<td>Overrepresentation of immune gene classes in altered EGL-38 activity array data.</td>
</tr>
<tr>
<td>4.1.</td>
<td><em>sa485</em> linkage analysis.</td>
</tr>
<tr>
<td>4.2.</td>
<td>Three-point <em>sa485</em> mapping.</td>
</tr>
<tr>
<td>4.3.</td>
<td>Two-point mapping using three snip-SNP’s across the X chromosome.</td>
</tr>
<tr>
<td>4.4.</td>
<td><em>lin-14(n179)</em> fails to complement <em>lin-14(sa485)</em>.</td>
</tr>
<tr>
<td>A.1.</td>
<td>Oligonucleotide primers.</td>
</tr>
<tr>
<td>B.1.</td>
<td><em>C. elegans</em> genetic strains.</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.</td>
<td>Expression of a vab-3::gfp reporter transgene</td>
<td>36</td>
</tr>
<tr>
<td>2.2.</td>
<td>Ectopic expression of vab-3 in the U blast cell in egl-38 and lin-48 mutants</td>
<td>38</td>
</tr>
<tr>
<td>2.3.</td>
<td>Correlation between vab-3 expression and the B.a cell fate</td>
<td>40</td>
</tr>
<tr>
<td>2.4.</td>
<td>egl-5 is required for proper B cell division and vab-3 expression</td>
<td>41</td>
</tr>
<tr>
<td>2.5.</td>
<td>Domain refinement model for regulation of pax-6/vab-3 expression in sensory organ precursors of the C. elegans male</td>
<td>42</td>
</tr>
<tr>
<td>3.1.</td>
<td>Microarray design and differentially expressed transcript groupings</td>
<td>73</td>
</tr>
<tr>
<td>3.2.</td>
<td>Multiple candidate genes are expressed in the vulva and/or rectum</td>
<td>75</td>
</tr>
<tr>
<td>3.3.</td>
<td>C. elegans death from reduced egl-38 activity precedes killing by S. marcescens</td>
<td>79</td>
</tr>
<tr>
<td>3.4.</td>
<td>Candidate gene expression patterns in the rectum</td>
<td>80</td>
</tr>
<tr>
<td>3.5.</td>
<td>Rectal expression of many candidate genes is dependent on egl-38</td>
<td>81</td>
</tr>
<tr>
<td>3.6.</td>
<td>RNAi knockdown reveals genes that function in M. nematophilum infection</td>
<td>82</td>
</tr>
<tr>
<td>4.1.</td>
<td>Genetic map of X chromosome markers used for sa485 mapping</td>
<td>113</td>
</tr>
<tr>
<td>4.2.</td>
<td>The sa485 strain JT8174 carries a G → A substitution in LIN-14</td>
<td>114</td>
</tr>
<tr>
<td>4.3.</td>
<td>The uterus of lin-14(sa485) mutants is asymmetric and abnormal</td>
<td>115</td>
</tr>
<tr>
<td>4.4.</td>
<td>lin-14(sa485) mutants exhibit a delay in vulval eversion relative to germline maturation</td>
<td>116</td>
</tr>
<tr>
<td>4.5.</td>
<td>Precocious completion of vulval inversion precedes Evl-D phenotype</td>
<td>118</td>
</tr>
<tr>
<td>4.6.</td>
<td>Intermediate levels of lin-14 activity in lin-14(n179) mutants result in the Evl-D phenotype</td>
<td>119</td>
</tr>
<tr>
<td>4.7.</td>
<td>Abnormal vulval eversion in lin-14 mutants is due to LIN-14 activity during the L1 larval stage</td>
<td>120</td>
</tr>
<tr>
<td>4.8.</td>
<td>The Pvl phenotype in lin-14(lf) mutants is a result of vulval collapse during the L4 larval stage</td>
<td>120</td>
</tr>
<tr>
<td>4.9.</td>
<td>uv1 specification is delayed in lin-14(sa485) animals, and delay is coincident with Evl-D phenotype</td>
<td>121</td>
</tr>
<tr>
<td>C.1</td>
<td>lin-14(sa485) mutants exhibit abnormal B cell division and spicule formation</td>
<td>144</td>
</tr>
<tr>
<td>D.1</td>
<td>egl-38 mutants show reduced infection by M. nematophilum</td>
<td>146</td>
</tr>
</tbody>
</table>
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>Asp/D</td>
<td>aspartic acid</td>
</tr>
<tr>
<td>AC</td>
<td>anchor cell</td>
</tr>
<tr>
<td>Bus</td>
<td>bacterially unswollen</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>Cog</td>
<td>connection of gonad defective</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Dpy</td>
<td>dumpy</td>
</tr>
<tr>
<td>Dar</td>
<td>deformed anal region</td>
</tr>
<tr>
<td>DTC</td>
<td>distal tip cell</td>
</tr>
<tr>
<td>δ</td>
<td>delta</td>
</tr>
<tr>
<td>ε</td>
<td>epsilon</td>
</tr>
<tr>
<td>Egl</td>
<td>egg-laying defective</td>
</tr>
<tr>
<td>Evl-D</td>
<td>eversion of vulval delayed</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>Symbol</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>gf</td>
<td>gain-of-function</td>
</tr>
<tr>
<td>G/Gly</td>
<td>glycine</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione S-transferase</td>
</tr>
<tr>
<td>Hi</td>
<td>hawaiian</td>
</tr>
<tr>
<td>HSN</td>
<td>hermaphrodite specific neuron</td>
</tr>
<tr>
<td>Hox</td>
<td>homeobox</td>
</tr>
<tr>
<td>hrs</td>
<td>hour(s)</td>
</tr>
<tr>
<td>ida</td>
<td>related to islet cell diabetes autoantigen</td>
</tr>
<tr>
<td>L</td>
<td>liter(s)</td>
</tr>
<tr>
<td>L1</td>
<td>first larval stage</td>
</tr>
<tr>
<td>L2</td>
<td>second larval stage</td>
</tr>
<tr>
<td>L3</td>
<td>third larval stage</td>
</tr>
<tr>
<td>L4</td>
<td>fourth larval stage</td>
</tr>
<tr>
<td>Lin</td>
<td>cell lineage abnormal</td>
</tr>
<tr>
<td>Lon</td>
<td>long</td>
</tr>
<tr>
<td>Let</td>
<td>lethal</td>
</tr>
<tr>
<td>If</td>
<td>loss-of-function</td>
</tr>
<tr>
<td>lact</td>
<td>β-lactamase domain containing</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
</tr>
<tr>
<td>µ</td>
<td>micro</td>
</tr>
<tr>
<td>M</td>
<td>moles per liter</td>
</tr>
</tbody>
</table>

**Figure:** xvi
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mab</td>
<td>male abnormal</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>mol</td>
<td>mole(s)</td>
</tr>
<tr>
<td>Mig</td>
<td>gonad migration defective</td>
</tr>
<tr>
<td>π</td>
<td>pi</td>
</tr>
<tr>
<td>Pax</td>
<td>paired homoebox</td>
</tr>
<tr>
<td>phat</td>
<td>pharyngeal toxin-related</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Pvl</td>
<td>protruding vulva</td>
</tr>
<tr>
<td>PD</td>
<td>paired DNA-binding domain</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>rt</td>
<td>room temperature</td>
</tr>
<tr>
<td>RMA</td>
<td>robust multi-array average</td>
</tr>
<tr>
<td>Ste</td>
<td>sterile</td>
</tr>
<tr>
<td>Sck</td>
<td>sick</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>Sma</td>
<td>small</td>
</tr>
<tr>
<td>Tris</td>
<td>trishydroxymethylaminomethane</td>
</tr>
<tr>
<td>Unc</td>
<td>uncoordinated</td>
</tr>
<tr>
<td>Uta</td>
<td>uterine asymmetry</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>utse</td>
<td>uterine seam</td>
</tr>
</tbody>
</table>

xvii
Vab  variably abnormal
VU  ventral uterine
ζ  zeta
CHAPTER 1

INTRODUCTION

1.1. The role of specification and morphogenesis in animal development

Developmental biology encompasses the precise coordination of events that transform a fertilized embryo into a complex multicellular organism. In general, development occurs via four processes: growth, fate specification, differentiation and morphogenesis. Each of these processes can be broken down further into categories, and can occur multiple times across the life of the organism. Ultimately, these different developmental processes must all be under tight control and coordination at the molecular level in order to faithfully orchestrate the combination and sequence of events required for normal development. One of the most important questions in developmental biology is identifying and characterizing the regulatory process responsible for the execution and coordination of these processes.

Proper control over developmental events is essential to all organisms, and it plays a significant part in human health. Approximately 3% of newborns harbor a physical defect resulting from abnormal development, and these account for roughly 20% of all infant deaths (http://www.cdc.gov/ncbddd/bd/default.htm). Furthermore, the
physical and behavior characteristics of actively dividing cells during development are notably similar to the characteristics of cancerous cells (rev. in Powers and Mu, 2008; Chaffer et al., 2007). In fact, stem cells that contribute to tissue growth and maintenance are currently thought to be a primary origin of cancer (rev. in Kucia and Ratajczak, 2006). Therefore, understanding of the fundamental principles underlying developmental processes and how they are regulated is of critical importance to medicine.

The processes of tissue specification and morphogenesis are central to developmental biology. Specification, or cell fate specification, involves the commitment of a cell to a specific developmental fate. The specification of myoblasts from a population of multi-lineage stem cells, for instance, commits these cells to the eventual production of muscle tissue (rev. in Shih et al., 2008). Morphogenesis is defined by physical changes of cell and tissue structure, and thus the combined action of all morphogenetic events during development is responsible for overall body shape and tissue organization of all multicellular organisms. After specification, for example, craniofacial muscles are formed from diverse migration and cellular restructuring of myoblasts (rev. in Noden and Francis-West, 2006). In a broader sense, therefore, individual cells must acquire the proper functional identity and physical orientation to establish populations of cells that will cooperatively perform physiological functions in a mature organism.

While the basic components and events for many developmental processes are known, many of the molecular genetic mechanisms responsible for their regulation and coordination remain unidentified. To cite an example, mammalian kidney formation involves an epithelial-to-mesenchymal transition, followed by specification of the uterine
bud, and subsequent crosstalk between these tissues (rev. in Bard, 1992). However, the identity and interaction of signaling pathways that control each of these steps (for example rev., see Gill and Rosenblum, 2006) is still under investigation. We can ask similar questions of a broad range of developmental events. For example, how are specification and morphogenetic events individually regulated? How are the factors that distinguish one tissue from another established in the proper spatiotemporal organization? In addition, these processes do not occur independent of one another, and they must be coordinated. How are individual specification and morphogenetic events correlated, and are those correlations controlled? Studies that begin to answer these questions will aid in understanding not only the “what”, “where”, and “when” of developmental processes, but the “how” of underlying gene expression and its regulation.

1.2. Transcriptional regulation of development

Cell identity and behavior are governed by the gene expression profile of the cell, as well as its developmental history, which is itself dependent upon the gene expression profile of precursor cells. In this way, gene expression is a core regulator of all developmental processes, and the diversity in cell identity and behavior results from dynamic changes in gene expression. There are multiple mechanisms that can initiate changes in gene expression at the DNA level, or transcription. Ultimately, all of these mechanisms alter transcription within the cell by modulating the activity of a class of proteins called transcription factors.
Transcriptional regulation of gene expression is a complex process involving several proteins and protein complexes. Some of these proteins act in a general manner, while some play gene-specific roles. For instance, gene products involved in direct recruitment and activation of RNA polymerase are required for transcription of all genes (rev. in Thomas and Chiang, 2006). This is also the case for chromatin remodeling complexes, such as ISWI and SWI/SNF, that function in DNA compaction and accessibility (rev. in Johnson et al., 2005). Alternatively, sequence-specific DNA-binding proteins recognize and regulate only subsets of genes. These factors alter transcription by modulating the recruitment and activity of general components such as the ones described above. Transcription factor structure and function is frequently modular in nature. Typically, these proteins contain a DNA-binding domain and a trans-activation domain. Further, other motifs are found in some transcription factors, including receptor domains that sense small molecule stimuli and repressor sequences that act in co-repressor recruitment.

As mentioned above, developmental processes are regulated by gene expression, and gene expression is regulated by the action of sequence-specific transcription factors. Single transcription factors have the capability to control the expression of tens or even hundreds of genes, so only a small number are necessary to establish distinct expression patterns within different cells. Moreover, cooperative function of transcription factors can induce transcription of genes that neither can induce alone. This is one way that the hierarchical nature of transcription factor biology emerges. In some cases, one (primary) transcription factor can initiate a tissue-specific developmental genetic program, including the expression of other (secondary) transcription factors. In turn, combinatorial
control between the primary and secondary transcription factors are able to induce further changes in gene expression. The hierarchical cascade is then reiterated by combinatorial action of tertiary transcriptional regulators with existing proteins, etc. When the function of any one of these transcription factors is lost, the resulting developmental process is abnormal, and the severity of abnormality is often reflective of how high in the hierarchy the affected transcription factor acts.

Transcription factors are typically necessary for normal development due to their ability to regulate expression of large sets of genes. In some instances, transcription factors acting near the top of a regulatory cascade are also sufficient for induction of a specific developmental program. One of the most famous examples of sufficiency among transcription factors is the *Drosophila melanogaster* gene *eyeless*. As implied by its name, mutations in *eyeless* produce fruit-flies with little/no eye tissue, demonstrating its necessity in eye development (Hoge, 1915). More dramatically, however, misexpression of EYELESS in sensory tissues across the fly body is sufficient to promote the formation of additional malpositioned—or ectopic—eyes (Halder et al., 1995). For this reason, precise control over the spatiotemporal expression of transcription factors such as EYELESS is absolutely critical to achieve normal animal development.

1.3. *Pax transcription factors are critical regulators of development in many species*

*eyeless* encodes one member of a family of transcription factors called Pax factors. Pax—or *paired box*—transcription factors are defined according to the presence of a unique type of DNA-binding domain. This domain is termed the *paired* domain (PD), after the founding gene member of the family. The PD is highly conserved from...
fruit flies and other invertebrates to humans, reflecting its importance in animal biology (rev. in Gruss and Walther, 1992). Like most transcription factors, Pax proteins also contain different combinations of other functional domains: homeobox DNA-binding domains, octapeptide repressor sequences, and C-terminal trans-activation domains. According to PD sequence and domain composition, Pax genes fall into distinct subclasses, and the different biological functions of Pax proteins generally reflect this classification (rev. in Underhill, 2000).

As evidenced by *eyeless* function, *Pax* genes play critical roles in animal development. *Pax* genes have been shown to function in a large number of different tissues, and this phenomenon is the result of two levels of tissue-specificity. First, each subclass of *Pax* genes has been found to regulate development of different sets of tissues, presumably due to differences in the PD and other protein motifs, as well as their expression pattern. Perhaps more interesting is the fact that individual *Pax* genes retain the ability to regulate the development of distinct tissues. *Pax*3, for instance, is necessary for proper maturation of muscle cells and for the commitment of progenitor cells to a melanocyte-producing cell type (Hornyak et al., 2001; Goulding et al., 1994). The importance of *Pax* genes during development is not only based upon their pleiotropic activities, but also their place high in the hierarchical cascade of transcription factor development. Each of these properties helps to explain the combinatorial nature of transcriptional regulation by Pax factors. Unlike some transcription factors, Pax proteins have a fairly degenerate DNA recognition sequence (Czerny et al., 1993). These proteins are often found to control gene expression in combination with other factors, whether they be gene products of a Pax transcriptional cascade or proteins required to distinguish
Pax gene regulation in one tissue versus another. Defining the logic of these context-dependent mechanisms is essential to understanding global and specific function of Pax genes, and there is much that is not yet understood.

Disruptions in Pax gene activity are associated with many developmental disorders, as well as progression of certain cancers (rev. in Lang et al., 2007). This makes understanding of both global and specific facets of Pax gene function during normal development all the more essential. Mutations have been identified in every subclass of human Pax gene, and are associated with diseases such as hearing loss (Waardenburg’s syndrome; Pax3; Tassabehji et al., 1992), vision impairment (Aniridia; Pax6; Jordan et al., 1992), and metabolic disorders (thyroid hypoplasia; Pax8; Macchia et al., 1998). Furthermore, many Pax proteins have been found to be aberrantly expressed in tumors, where they promote the survival of cancerous cells (Muratovska et al., 2003). In fact, efforts are underway to target Pax activity as a prospective cancer treatment (Lang et al., 2008; Jiang et al., 2007).

Murine models for examination of Pax-dependent disease pathology have been extremely useful, but are limited by the complexity and inaccessibility of in utero mouse development. In addition, many of the underlying features of Pax protein structure, expression, and function were first elucidated in invertebrate systems such as Drosophila. As the study of Pax genes has advanced, these features have repeatedly been found to be conserved in comparison to vertebrates (Onuma et al., 2002; rev. in Kammermeier and Reichart, 2001; Walther et al., 1991). Thus, investigation of Pax gene function in invertebrate model systems can provide useful information regarding the mechanisms of developmental and transcriptional regulation by vertebrate Pax genes.
1.4. The nematode C. elegans as a model to study Pax genes, specification, and morphogenesis

The invertebrate C. elegans has been demonstrated to be a powerful model system for addressing many biological questions, including those directed at developmental genetics (Brenner, 1974; Sulston and Horvitz, 1977). C. elegans is a free-living bacterivorous nematode (roundworm) found in the soil. Beyond the basic cellular processes shared among all multicellular organisms, most of the physiological functions necessary for vertebrate survival are also exhibited by C. elegans. Particular examples include a nervous system that directs behavior, muscles that allow movement, and an intestine for nutrient uptake. Therefore, proper specification and coordination of cells that act in these capacities is required in C. elegans development, just as it is in humans.

Several features of this organism make it amenable to developmental genetics research. Fertilized embryos develop briefly in utero before being laid into the environment as eggs, where they complete embryogenesis. Subsequently, embryos hatch and undergo four stages of larval development before reaching adulthood and producing mature gametes available for fertilization. This entire life cycle is approximately 3 days, allowing for rapid genetic analyses and convenient generation of progeny for developmental studies. Two things in C. elegans lend themselves well to examining tissue specification and morphogenesis: the somatic cell division pattern is reproducible from one animal to the next, and has been traced from the single-celled embryo to the adult, and transparency of the animals allows for in vivo observation at the single cell resolution (Sulston et al., 1983; Sulston and Horvitz, 1977). Finally, C. elegans are
tractable to both forward and reverse genetic approaches, and they are able to produce progeny through either mating or self-fertilization.

Not only do *C. elegans* develop analogous tissues to higher animals, but many of the molecular underpinnings of the development and function of these tissues are also conserved amongst phylogenetically divergent species. Signaling pathways are repeatedly found to function in similar processes, such as the regulation of cell polarity by Wnt pathway components (rev. in Jenny and Mlodzik, 2006), and the lateral inhibition of cell fates by Notch signaling (Chitnis, 1995). Moreover, many transcription factors and their role in development are conserved from *C. elegans* to vertebrates. An example of this is the transcription factor MyoD, which directs muscle development in both species (Harfe et al., 1998; rev. in Berkes and Tapscott, 2005). The conservation of molecular mechanisms among species further underscores the utility of *C. elegans* as a model for the examination of developmental gene expression and its regulation.

Like MyoD, the organogenic functions of Pax factors are highly conserved. All four subclasses of Pax factors are represented in the *C. elegans* genome. Two of these classes are comprised of a single gene, whereas the *Pax2/5/8* and *Pax4/6* subfamilies are each represented by two genes. *pax-2* and *egl-38/Pax2/5/8* are reported to be partially redundant, although only *egl-38* mutants exhibit strong developmental defects (Chamberlin et al., 1997; Park et al., 2006). Similarly, mutations in *vab-3/Pax6* cause embryonic and postembryonic defects, but the related gene *Y53C12C.1/pax6* has no obvious function (Chisholm and Horvitz, 1995; Zhang and Emmons, 1995). *egl-38* and *vab-3* act in several tissues during development, as do their vertebrate counterparts (Chamberlin et al., 1997; Park et al., 2006; Cinar and Chisholm, 2004).
1.5. Defining genetic networks of two Pax genes and an additional transcriptional regulator of C. elegans post-embryonic development

Although there is considerable knowledge about C. elegans Pax function at the cellular and tissue levels, the genetic networks surrounding these important genes are just beginning to be characterized. One area of our interest is how transcription factor expression patterns are established, and what genetic inputs regulate those expression events. Once a Pax factor is expressed in a cell or tissue, it helps to promote a cell fate, initiate signaling, or regulate a morphogenetic event. This is necessarily achieved through the transcriptional activation and/or repression of specific target genes, yet most of these targets are as yet unidentified. Therefore, a further goal of our work is aimed at learning what the tissue-specific targets of Pax proteins are, as well as how these downstream genes mediate Pax-dependent specification and morphogenetic functions.

Pax6 genes, including the previously mentioned Drosophila eyeless, encode proteins critical for specifying eye formation (rev. in Ashery-Padan and Gruss, 2001), in addition to other sensory tissues (Nomura et al., 2007). C. elegans do not have eyes, but the Pax6 ortholog is critical for the development of sensory organs in the male of the species (Zhang and Emmons, 1995). In one particular case, a product from this complex gene (termed vab-3) becomes expressed in two cells within the post-embryonically dividing rectum (Zhang et al., 1998). These cells generate two sensory mating structures, and vab-3 activity is required for their development (Chamberlin and Sternberg, 1995). We have created a transcriptional reporter for vab-3 that expresses green fluorescent protein (GFP) in these cells and their progeny. We use this reporter to uncover several
genetic inputs that activate and restrict its expression. Together, these regulatory controls illustrate the underlying genetic logic that an animal uses to define the expression of a critical regulator of development.

Another important aspect of Pax transcription factor biology that is not well understood is how tissue-specific gene expression is established, and how that gene expression leads to cellular and anatomical development. In vertebrates, Pax2/5/8 proteins govern ongoing specification and/or organogenesis of kidneys, the thyroid gland, the midbrain, and B lymphocytes (rev. in Dressler, 1996; Christophe, 2004; Reichert, 2005; Hagman et al., 2000). Similarly, the C. elegans Pax2/5/8 gene egl-38 plays a role in the development of multiple organs (Chamberlin et al., 1997, Park et al., 2006), providing an opportunity to address tissue-specific functions for an individual Pax gene in a relatively simple animal. To begin to answer these questions, we have taken the approach of identifying potentially new target genes for the Pax2/5/8 factor EGL-38 via gene expression microarrays. We report the use of a cross-referencing approach to increase the robustness of candidates recovered by these microarrays, and the validation of this approach through GFP reporter gene expression and RNA-mediated interference (RNAi) studies. Further, we outline a novel immune function for this transcription factor, and provide evidence that some candidate genes uncovered in the microarray analysis function in this biological process.

In our pursuit to better understand the regulation and function of Pax gene activity, we characterized a novel mutation that causes developmental defects in the same tissues as vab-3 and egl-38. It followed from our vab-3 regulation studies and from the well-documented combinatorial nature of Pax factor function (for example, Gong et al.,
2007), that other transcription factors represent a significant part of the genetic networks surrounding Pax genes. Concordantly, we found that the mutation of interest was an allele of another transcription factor gene. This gene, called lin-14, encodes a novel type of transcription factor that functions in controlling the timing of post-embryonic developmental events. lin-14 is part of a genetic pathway that acts in a concerted fashion to specify a developmental “clock”, ushering the worm from newly-hatched larva through successive stages of patterned cell division, specification, and morphogenesis to adulthood. Of specific interest regarding the allele we identified is the fact that this mutation caused emerging defects in late stages of development, despite the fact that the gene is limited in expression and function to the first larval stage. We found that specification and morphogenetic events in the egg-laying system are temporally disrupted in the lin-14 mutants, resulting in asynchronous maturation of the two tissue types comprising the egg-laying system.

In total, this work illustrates the genetic mechanisms by which developmental processes of specification and morphogenesis can occur. We have identified specific genetic inputs and outputs of Pax gene expression, and show how an additional transcription factor can play a distinct, but necessary role in Pax-dependent tissues. By addressing both spatial and temporal aspects of fate specification and morphogenesis, our work yields further insights and raises important questions concerning these developmental processes in higher animals such as humans.
CHAPTER 2

POSITIVE AND NEGATIVE REGULATORY INPUTS RESTRICT *PAX-6/VAB-3* TRANSCRIPTION TO SENSORY ORGAN PRECURSORS IN *C. ELEGANS*

2.1. INTRODUCTION

Organogenesis is a fundamental process in animal development requiring the strict coordination of cell fate specification, intercellular signaling events, and morphogenetic changes among multiple cell types. Given the complexity of such a biological process, it is extraordinary that the genetic program necessary to achieve proper organogenesis can be under control of a small number of critical genes. For example, members of the *Pax* gene family play a central role in the development of organs as disparate as the skeleton and kidney, and are also abnormally active in some cancer cells (reviewed in Lang et al., 2007; Robson et al., 2006). *Pax* gene family members encode proteins containing DNA-binding domains—a full or partial homeodomain, and a *paired* domain (Czerny et al., 1993)—and thus they influence development by regulating the transcription of other genes. Understanding how these key regulatory genes influence organ development remains an important question.
Pax-6 has been well characterized for its role in eye organogenesis (reviewed in Gehring, 1999; Ashery-Padan and Gruss, 2001). In Drosophila, loss of Pax-6/eyeless activity results in a total absence of eyes (Quiring et al., 1994). Similarly, Pax-6/Small eye mutant mice have small non-functional eyes (Hill et al., 1991), and the human conditions aniridia, Peter’s anomaly, and cataracts can result from various Pax-6 deficits (reviewed in Chi and Epstein, 2002). Additionally, misexpression of Eyeless in certain Drosophila tissues converts them into eye tissue (Halder et al., 1995), indicating that Pax-6 can be sufficient to define eye organogenesis. These findings identified Pax-6 as a potential master control gene due to its robust ability to be both necessary and sufficient for eye development. Although more recent work has demonstrated that eye development is dependent upon a network of genes (Silver and Rebay, 2005), Pax-6 remains a critical determinant for eye development.

Apart from eye development, Pax-6 has a broader role in the organogenesis of sensory and neural tissues. In vertebrates, it is critical for forebrain patterning (Stoykova et al., 1996) and proper generation of the nasal placode (Grindley et al., 1995). In the eyeless nematode C. elegans, the Pax-6 gene functions in head neurons and in precursors to sensory mating organs of the male tail (Chisholm and Horvitz, 1995; Zhang and Emmons, 1995; Chamberlin and Sternberg, 1995). These functions suggest that Pax-6 may have a conserved ancestral role in the development of sensory tissues, and examination of these alternate contexts can elucidate features of Pax-6 applicable to its broader role in sensory organogenesis versus those specifically adapted to eye specification and patterning.
Although there is a large body of work characterizing Pax-6 function, and recent genome-wide studies have identified downstream target genes (Jemc and Rebay, 2006; Ostrin et al., 2006; Nemeth et al., 2005), much less is known about the regulatory pathways that generate the Pax-6 expression pattern. However, the powerful nature of Pax-6 activity as a sensory organ determinant indicates that it is critical for animals to precisely control the expression of this gene. Therefore, to fully understand how Pax-6 functions in different sensory organ contexts, it is necessary to describe the mechanisms establishing the gene’s expression pattern. To this end, studies have begun to identify some cis-regulatory elements and signaling pathways regulating Pax-6 expression. Transgenic mouse experiments have revealed cis-regulatory modules that direct Small eye expression in a variety of tissues (Kammandel et al., 1999), and genetic studies have implicated Fgf and Bmp signaling in Small eye lens expression (Faber et al., 2001). Furthermore, facets of Pax-6 regulation have been evolutionarily conserved along with Pax-6 protein function. Regulatory sequences directing expression in different cell types of the eye are conserved between mice and flies (Xu et al., 1999), as is the ability of Notch signaling to promote Pax-6 expression in these species (Onuma et al., 2002). Despite these findings, the underlying regulatory logic employed by an organism to generate the Pax-6 expression pattern remains unclear.

The C. elegans male tail provides a relatively simple and powerful system in which to examine general features of sensory organogenesis. In the C. elegans male, four cells comprising the rectum, or hindgut, are specialized blast cells (Sulston et al., 1980). They generate sex-specific post-embryonic cell lineages. Differentiation and morphogenesis of these tissues lead to the formation of sensory mating structures.
Specifically, two cells, termed B and Y, contribute all the cells of the copulatory spicules and post cloacal sensilla, respectively. These sensory organs function in the location of the hermaphrodite vulva, as well as in sperm transfer (Sulston et al., 1980; Liu and Sternberg, 1995). Although less complex than the eye, development of these organs requires the coordination of multiple cell types for proper development and function. For example, the spicules are composed of neurons, glia-like sheath cells, and socket cells that secrete a chitin-like cuticle material that surrounds the nervous tissue. Examination of the critical genes that direct development of these sensory tissues may aid in a broader understanding of the underlying genetic logic employed by animals during organogenesis.

*pax-6* is a complex locus in *C. elegans*, producing one transcript coding for a protein with only a homeodomain (*mab-18*), and one transcript coding for a protein including both the homeodomain and a *paired* domain (*vab-3*; Chisholm and Horvitz, 1995; Zhang and Emmons, 1995; Cinar and Chisholm, 2004, Fig. 1B). This work focuses on regulation of the *vab-3* transcript. Previous work has demonstrated that *pax-6/vab-3* mutant males exhibit abnormal B and Y cell lineages, and that the spicules are severely reduced or absent in adults (Chamberlin and Sternberg, 1995, and unpublished). Further, the defect causing this loss-of-spicule phenotype was found to result from a specification defect early in the precursor cell lineage. The B cell normally divides asymmetrically late in the first larval stage to produce a large anterior daughter (B.a), which goes on to produce all cells of the spicules, and a smaller posterior daughter (B.p), whose derivatives have no sensory function. A similar asymmetric division and adoption of distinct fates occurs in Y, with Y.p giving rise to the post-cloacal sensilla. In *vab-
3(e648) males, the B.a cell adopts the fate of its sister cell, resulting in B.a or its daughters developing like B.p (Chamberlin and Sternberg, 1995). This result demonstrates that, as in the case of eye development in other organisms, Pax-6 can function as a critical regulator of sensory organ identity in C. elegans.

In this study, we characterize the C. elegans vab-3 expression pattern within the male B and Y lineages. We demonstrate that the expression pattern is established by both positive and negative regulatory inputs. We show that the Abdominal B ortholog, egl-5, is necessary for vab-3 expression in the sensory organ precursors B.a and Y.p, as well as in other cells when genetic manipulation causes vab-3 to be ectopically expressed. We find that the vab-3 expression domain is limited by a Wnt signaling pathway, as well as by a transcriptional cascade involving another Pax transcription factor gene, egl-38. These results outline the regulatory logic that establishes pax-6/vab-3 expression in a set of sensory organ precursors.

2.2 MATERIALS AND METHODS

2.2.1. Genetic strains and strain construction

C. elegans strains were cultured under standard conditions (Brenner, 1974; Sulston and Hodgkin, 1988). All experiments were done at 20°C. The strain DR466 him-5(e1490) was used as wild-type control in all experiments. The following mutations were used: Linkage Group I (LGI): lin-17(n671), lin-44(n1792); LGIII: lin-48(sa469), lin-48(gv4), egl-5(n486), unc-119(e2498); LGIV: egl-38(sy294); LGV: him-5(e1490); LGX: vab-3(e648). Additional genetic information is available at http://www.wormbase.org.
To transfer the \textit{vab-3::gfp} reporter transgene, \textit{guEx356}, into mutant backgrounds, we began with animals carrying \textit{unc-119(e2498)} and \textit{him-5(e1490)} in addition to the mutation of interest. These Uncoordinated (Unc) hermaphrodites were mated to CM549 males, non-Unc F1 hermaphrodites were selected and allowed to self-cross. F2 hermaphrodite offspring exhibiting the relevant mutant phenotype were selected and used to establish the transgene-bearing mutant strains.

All double mutants were created using the following basic strategy. Construction of the \textit{lin-17(n671)}; \textit{egl-5(n486)} strain is used as an example. CM549 males were mated to \textit{lin-17(n671)}; \textit{unc-119(e2498)}; \textit{him-5(e1490)} hermaphrodites. Non-Unc cross-progeny males were then mated to \textit{egl-5(n486)}; \textit{unc-119(e2498)}; \textit{him-5(e1490)} hermaphrodites. The resulting non-Unc progeny were picked individually and allowed to self-cross. Plates were assessed for appearance of animals with the \textit{lin-17} Bivulva (Biv) phenotype, and those hermaphrodites were again individually plated. Animals from the following generation were allowed to reach adulthood, and individuals displaying the \textit{egl-5} Egg-laying defective (Egl) phenotype were then chosen as founders for the double mutant strain with the full genotype: \textit{lin-17(n671)}; \textit{unc-119(e2498)}; \textit{egl-5(n486)}; \textit{him-5(e1490)}; \textit{guEx356}.

\subsection*{2.2.2. Microscopy}

Nomarski DIC microscopy was performed, as described previously (Sulston and Horvitz, 1977). Briefly, larvae were mounted on an agar pad containing sodium azide, and their hindgut blast cells were viewed under incident and ultraviolet light to analyze anatomy and GFP expression, respectively. Qualitative expression differences between
cells and between genotypes were observed. However, any cell displaying visible GFP fluorescence was scored as positive. Relative nuclear size (visualized as spherical clearings in the image of a cell) of the two B cell daughters was utilized to determine the symmetry/asymmetry of cell division. This simple measure was used, as we find cell size and fate differences to be highly correlated (data not shown). All anatomy was observed and photographed using a Zeiss Axioskope (Carl Zeiss, Inc.) equipped with Spot-RT CCD camera and software (Diagnostic Instruments, Inc.).

2.2.3. vab-3::gfp construction and expression

The vab-3 transcriptional GFP reporter plasmid pAJ51 was created in two steps. A 1.0-kb PstI-BamHI PCR fragment spanning portions of intron 2 and exon 3 was cloned into the GFP reporter pPD95.70 (kind gift from A. Fire). Next, a 13.2-kb PstI fragment of the cosmid F14F3 containing 4.7-kb of upstream sequence through intron 2 was subcloned into the PstI site used in the previous step. Restriction enzyme digests were used to confirm the presence and orientation of the insert.

Transgenic strains were created by standard germline transformation methods (Mello et al., 1991), with a modification to produce more stable expression. Low-copy, high-complexity extrachromosomal arrays were generated by injection of excess genomic carrier DNA with the transgene (Kelly et al., 1997). Specifically, linearized pAJ51 (15ng/µl) was co-injected with linearized unc-119(+) transformation marker pDP#MM016 (Maduro and Pilgrim, 1995; 1 ng/µl) and PvuII-digested yeast genomic DNA (300ng/µl) into unc-119(e2498); him-5(e1490). We isolated transgenic strains CM549 and CM550 containing the arrays guEx356 and guEx357, respectively. GFP
expression patterns were identical for both transgenic lines, and guEx356 was chosen for use in expression analyses. The stability of guEx356 transmission was assessed by counting the number of transgene-bearing offspring in the broods of several hermaphrodites. We find that 82.4% of the offspring from 8 unc-119(e2498); him-5(e1490); guEx356 hermaphrodites, and 78.9% of the offspring from 9 egl-5(n486) unc-119(e2498); him-5(e1490); guEx356 hermaphrodites, retain the transgene.

2.3. RESULTS

2.3.1. A transcriptional reporter recapitulates the VAB-3 expression pattern in male sensory structure precursors

*C. elegans* male mating is a complex behavior coordinated by several male-specific sensory structures (Liu and Sternberg, 1995). This work focuses on the development of two of these sensory structures, the copulatory spicules and the post-cloacal sensilla. The copulatory spicules and post-cloacal sensilla are generated from precursor cells in the hindgut, termed B and Y, respectively (Figure 2.1. A). These cells, along with two non-sensory precursor blast cells, F and U, do not divide in hermaphrodites. However, they all divide to produce a large number of offspring during male postembryonic development. Previous work has shown that *pax-6/vab-3* is expressed in, and required for normal development of, the B and Y cell lineages in males (Zhang et al., 1998; Chamberlin and Sternberg, 1995, and unpublished). Due to the central role of *Pax-6* in the development of sensory organs in animals, we set out to
define the regulatory inputs for vab-3 in order to understand how its pattern of gene expression is established.

We created a transcriptional reporter transgene that includes 4.7 kb upstream of the first exon, as well as the first two large introns of vab-3 (Figure 2.1. B; Materials and Methods). Using this reporter, we created transgenic strains and observed expression patterns consistent with previous observations for VAB-3 protein (Zhang et al., 1998). Both sexes bearing a vab-3::gfp transgene exhibit GFP expression in neuronal, and occasionally epidermal cell types in the head (Figure 2.1. C-D and data not shown). Neuronal transcription begins after gastrulation and persists throughout the life of the animals. In addition, vab-3::gfp displays sex-specific expression patterns, such as in the distal tip cells (DTC’s) of L4 hermaphrodites and the B.a and Y.p sensory organ precursors in males (data not shown and Figure 2.1. E-G). Epidermal cell and DTC expression are consistent with the Vab and Mig morphological defects displayed by vab-3 mutants (Chisholm and Horvitz, 1995; Nishiwaki, 1999). Neuronal function of vab-3 has not been well characterized, but VAB-3 protein expression has been observed in a variety of head neurons (Zhang et al, 1998). Notably, our vab-3::gfp transgene is not expressed in any ray lineages (data not shown). This is consistent with the prediction that it does not report expression of the mab-18 transcript. mab-18 encodes a paired domain-less PAX-6 isoform, and utilizes a transcriptional start site and enhancers internal to the pax-6 locus and downstream (3’) from the sequences contained within our reporter (Figure 2.1. B; Chisholm and Horvitz, 1995; Zhang and Emmons, 1995; Cinar and Chisholm, 2004).
To understand how *C. elegans pax-6/vab-3* expression is regulated in sensory precursors, we focused on the sex-specific expression of *vab-3::gfp* in the male hindgut. Consistent with antibody studies (Zhang et al., 1998), we found *vab-3::gfp* expression is visible in the B.a and Y.p cells beginning in early L2 (Figure 2.1. E-G). Expression levels among individual cell types derived from these precursors are not appreciably different through mid L3. In the B.a lineage, GFP persists through larval development. Expression diminishes during late L4 tail morphogenesis, and is absent in adults. Expression in the Y.p lineage shows a distinct temporal pattern. GFP in the Y.p lineage diminishes from mid L2 through mid L3, at which time only the eight B.a descendants exhibit expression (Figure 2.1. H-J). Interestingly, the timing of GFP loss in both lineages is roughly coincident with the timing when the cells undergo initial differentiation, suggesting *vab-3* may be important for the fate specification process, but not maintenance of the differentiated state.

### 2.3.2. A transcriptional cascade restricts *vab-3* expression within the male hindgut

Mutations in several genes that affect hindgut and male tail development have been identified. Among them are mutations in *lin-48* and *egl-38* (Chamberlin et al., 1999). *lin-48* encodes a zinc finger transcription factor that is expressed in, and required for, the proper development of cells in the mid-hindgut (Chamberlin et al., 1999; Johnson et al., 2001). In *lin-48* males, cells from the presumptive U cell lineage generate spicule cuticle, indicative of an abnormal production of socket-like cells normally generated by the B.a lineage (Jiang and Sternberg, 1999; Chamberlin et al., 1999). Prior genetic analyses have utilized hypomorphic *lin-48* alleles. In this study, we used the previously
reported missense allele *lin-48(sa469)* (Chamberlin et al., 1999), as well as *lin-48(gv4)*, thought to be a genetic null. This allele bears a small deletion and is predicted to produce a protein with no LIN-48 activity (M.W. Krause, unpublished communication). *egl-38* codes for a Pax-2/5/8 protein (Chamberlin et al. 1997), and EGL-38 directly activates *lin-48* hindgut expression (Johnson et al., 2001). Like *lin-48, egl-38* is required for the appropriate cell fate of several hindgut cell types. Development of the presumptive U lineage is abnormal in *egl-38(sy294)* males. However, adults exhibit crumpled spicules rather than the ectopic spicule material phenotype seen in *lin-48* mutants (Chamberlin et al., 1997).

To assess the relationship between *lin-48, egl-38*, and *vab-3* during male tail development, we scored *vab-3::gfp* expression in mutants. In *lin-48* mutant males, we found that the *vab-3::gfp* transgene is ectopically expressed in the presumptive U cell and its offspring (Figure 2.2. A-B). Early in the L2 stage, GFP expression in the U cell is infrequent, but becomes more robust as development proceeds. By mid-L3, roughly half of males show this ectopic expression event, approaching the percentage of *lin-48* mutant adults that generate ectopic cuticle-producing socket cells (Chamberlin et al., 1999; Jiang and Sternberg, 1999). Thus one function of *lin-48* is to restrict *vab-3* expression from the U cell lineage.

Introduction of *vab-3::gfp* into an *egl-38* mutant background also resulted in inappropriate GFP expression in the presumptive U cell and its offspring (Figure 2.2. C-D). This was expected, as *egl-38* is required for *lin-48* expression in the U cell (Johnson et al., 2001). However, comparison of the *egl-38(sy294)* ectopic expression phenotype to that of *lin-48* mutants revealed that the proportion of males affected was distinct. Fewer
egl-38 mutants were observed with ectopic vab-3::gfp expression, and this expression appeared to be more frequent in younger animals, in contrast to the lin-48 mutant males. We constructed an egl-38; lin-48 double mutant strain to further investigate the relationship between lin-48 and egl-38 with respect to vab-3 transcriptional repression. We found that the double mutants display an intermediate phenotype, rather than a strict epistatic relationship. Figure 2.2. E summarizes the transgene expression differences observed between these genotypes.

Reassessing the vab-3 transcription data in terms of temporal expression pattern, the distinction between egl-38 and lin-48 mutants becomes more apparent (Figure 2.2. F). As development progresses from a time point where B.a is undivided (late L1) to the point there are eight B.a descendents (late L2 through mid L3), the proportion of lin-48 mutant males transcribing vab-3::gfp in the U cell lineage increases considerably. During the same period, the percentage of ectopically expressing cells in egl-38 mutant males remains relatively static. Double mutants exhibit a complicated vab-3::gfp expression phenotype. GFP expression levels reflect that seen in egl-38(sy294) males at the earliest time point, but rise to levels exceeding that of either single mutant during intermediate time points. At the time when B.a has produced eight progeny, the frequency of expression in U-derived cells has dropped back to a level more similar to that of egl-38(sy294) single mutants. We conclude that although lin-48 expression is dependent on egl-38 (Johnson et al., 2001), the double mutant phenotype does not suggest a strict epistasis with respect to vab-3 expression. Since we used a presumed null allele of lin-48 for the double mutant analysis, we interpret that the intermediate
phenotype does not result from residual activity associated with this downstream gene. Consequently, we conclude that egl-38 influences vab-3 repression in the U cell through two mechanisms: a lin-48-dependent mechanism, and a lin-48-independent mechanism.

2.3.3. Wnt signaling restricts vab-3 transcription within the B lineage

During C. elegans development, Wnt signaling orchestrates many cell polarities and asymmetric cell divisions, including the first division of the male B cell (reviewed in Herman, 2002). Normally, the first division of B in wild-type males generates daughter cells distinct in both size and cell fate. When animals are mutant for the frizzled receptor gene, lin-17, B cell polarity is lost, cytokinesis of B is symmetric, and both daughters of B adopt a B.a-like cell fate (Sternberg and Horvitz, 1988). Unlike lin-17(n671) males, the B cell of lin-44/Wnt mutants divides asymmetrically, but does so in a reversed orientation, resulting in a posterior B.a-like daughter and an anterior B.p-like daughter (Herman and Horvitz, 1994). Although why the receptor and the ligand mutants exhibit different phenotypes continues to be an area of investigation, the mutant analysis demonstrates that both genes are important for establishing proper orientation of B cell division and daughter cell fate.

As vab-3 is required for proper B.a fate, and Wnt signaling is necessary for suppressing a B.a-like fate in the presumptive B.p cell, we tested the possibility that vab-3 may be transcriptionally repressed in the posterior daughter of B as a result of Wnt signaling. Examination of vab-3::gfp expression in lin-17(n671) and lin-44(e1472) mutant animals revealed ectopic vab-3 expression in the posterior daughter of B. Specifically, GFP can be observed in both daughters of B in lin-17(n671) males (Figure
2.3. A-B), while expression in lin-44(e1472) males is frequently observed in the posterior
daughter and its offspring only (Figure 2.3. C-D). Transcription in these altered lineages
is generally maintained throughout larval development, similar to transcription in a wild-
type B.a lineage (data not shown). Figure 2.3. E illustrates the correlation between \textit{vab-3}
expression and cytokinesis in these mutants. When males of each genotype are grouped
according to B cell polarity and asymmetry of cytokinesis (see Materials and Methods),
the majority of animals exhibit a strong coincidence of \textit{vab-3::gfp} expression and B.a-
like cell fate. We conclude that \textit{vab-3} transcription is negatively impacted by Wnt
signaling, and that it is tightly associated with the B.a cell fate, consistent with the
requirement for \textit{vab-3} activity in spicule development.

\textbf{2.3.4. \textit{vab-3} expression in the male tail requires the HOM-C gene egl-5}

The experiments thus far have elucidated some of the genetic mechanisms acting
to restrict \textit{vab-3} transcription to two blast cell lineages of the male tail. However, in
order for these pathways to limit \textit{vab-3} expression to these sensory organ precursors, we
reasoned that there must exist some genetic activator of \textit{vab-3} transcription. A primary
candidate for such transactivation potential was the \textit{egl-5} gene, which encodes a HOM-C
protein orthologous to \textit{Drosophila} Abdominal B (Chisholm, 1991; Chow and Emmons,
1994). As with \textit{HOM-C} genes in other species, \textit{egl-5} is expressed in the posterior region
of the worm body, and is required for proper development of, among other cells, the
male-specific blasts of the hindgut. In \textit{egl-5} mutant males, B and Y both undergo their
first division at the appropriate time and in the proper axis orientation. However, the
asymmetric site of cytokinesis of B is frequently lost (Figure 2.4. A-B). We did not
assess the cytokinesis phenotype of Y in this mutant background. The resulting lineages of both B and Y are abnormal, and do not resemble the wild-type lineages of neighboring or sister cells (Chisholm, 1991; Ferreira et al., 1999).

\textit{egl-5(n486)} males display a fully-penetrant loss of spicule phenotype. We utilized the \textit{vab-3::gfp} transgene to assess the influence of \textit{egl-5} activity on \textit{vab-3} expression. In contrast to wild type, \textit{vab-3::gfp} expression does not initiate in the tail of \textit{egl-5(n486)} mutants, and is absent from the B and Y cell lineages (Figure 2.4. C-F). The transgene expression pattern among neuronal cells of the head is unaltered, however, confirming the presence of the transgene in the animal. This result indicates that the \textit{HOM-C} gene is necessary to promote the expression of \textit{vab-3} in the B.a and Y.p lineages.

\textbf{2.3.5. Ectopic \textit{vab-3} expression in non-sensory organ precursors requires \textit{egl-5}}

Although normal expression of \textit{vab-3} in B.a and Y.p is lost in \textit{egl-5(n486)} mutants, we have found additional genetic pathways act to repress \textit{vab-3} transcription outside of these sensory precursor cells. This led us to ask whether the ectopic transcription exhibited in the transcription factor and Wnt signaling mutants also requires \textit{egl-5} activity. Strains carrying either \textit{lin-48(gv4), egl-38(sy294), or lin-17(n671)} in combination with \textit{egl-5(n486)} were constructed. In all double mutants, no GFP expression was observed in the hindgut region (Table 2.1.). These experiments indicate that the ectopic \textit{vab-3} expression in the U and B.p lineages, just as in the B.a and Y.p lineages, requires \textit{egl-5}.
2.4. DISCUSSION

2.4.1. A domain refinement model for establishment of pax-6/vab-3 expression

The pax-6/vab-3 gene is expressed in a precise pattern of sensory organ precursor cells in the C. elegans male hindgut. We propose that this tight regulation is established through refinement of a broad expression domain. We have found both positive and negative regulatory inputs controlling the expression of vab-3 (Figure 2.5.). A HOM-C gene, egl-5, allows for the activation of vab-3 expression in multiple cells in the hindgut region. In cells that normally are not sensory organ precursors, this activation is overridden by repressive inputs. We have identified two of these inputs: a transcriptional cascade for one cell (U) and a Wnt signaling event for another (B.p). egl-5 expression is widespread among cells in the posterior body region of the worm (Ferreira et al., 1999), and it is plausible that egl-5 has the capacity to promote vab-3 throughout its expression domain. The implication of this is that additional genetic pathways that act to repress vab-3 expression remain to be identified. How widespread the mechanism of domain refinement in the transcriptional regulation of master regulators like Pax-6 is unknown. Given the apparent conservation of genetic mechanisms surrounding Pax-6 gene regulation, we speculate that the Pax-6 expression pattern in other sensory organ systems is established through layered restriction of a broad domain of potential activation.
2.4.2. *pax-6/vab-3 transcription is observed in sensory organ precursor cells*

We report here a transcriptional GFP reporter for the *C. elegans* *pax-6* gene, *vab-3*. Expression in head hypodermis, hermaphrodite DTC’s, and male tail sensory organ precursors is closely coincident with the tissues that require *vab-3* function. The reliable expression of *vab-3::gfp* in the developing B.a and Y.p lineages makes it a useful experimental tool for the study of *vab-3* regulation in these sensory organ precursor cells. Although the dynamics of GFP stability may not be the same as endogenous VAB-3, our results suggest that in both the B.a spicule lineage and the Y.p post-cloacal sensilla lineage, cessation of *vab-3* expression precedes terminal differentiation of the tissue. There is accumulating evidence that *Pax* gene expression is closely associated with a progenitor cell state in various organogenic contexts. For instance, PAX-3 expression diminishes prior to initiation of the MyoD-directed differentiation program in mouse myoblasts (Williams and Ordahl, 2000), and Pax-6 inactivation causes retinal progenitor cells to lose their pluripotency (Marquardt et al., 2001). Furthermore, the undifferentiated, stem cell-like characteristic of cancer cells is often accompanied by abnormal expression of Pax factors in these cells (reviewed in Lang et al., 2007; Robson et al., 2006). Our expression data suggests that *vab-3* may have a progenitor maintenance function in the *C. elegans* male tail. Taken together, these results open up the intriguing question of whether the role of *Pax* genes in progenitor cells is a case of convergent evolution, or whether this function has been conserved across a broad range of metazoan species.
2.4.3. **Regional induction of pax-6/vab-3 expression is mediated by the HOM-C gene egl-5**

We found the *HOM-C* gene *egl-5* is necessary for male hindgut expression of *vab-3*. *egl-5* mutant males not only lack the endogenous *vab-3* expression pattern, but also lose ectopic transgene expression in double mutants with the transcription factors and Wnt components we have examined. Our work demonstrates a second example of a HOM-C factor regulating expression of a *Pax* gene, as Yi and Sommer (2007) have shown that *pax-3* expression is directly activated by Hoxd4/LIN-39 in the developing vulva of the nematode *Pristonchus pacificus*. Various studies have identified other types of interaction between Hox and Pax factors, as well. For example, Pax-6 binds and activates a central nervous system enhancer of *Hoxd4* (Nolte et al., 2006), and Hox11 acts in a complex with Pax-2 and Eya1 to regulate gene expression during early kidney development (Gong et al., 2006). Together, these data suggest that animals have utilized *HOM-C* and *Pax* genes in a variety of cooperative ways during organ development.

2.4.4. **Negative regulatory inputs restrict pax-6/vab-3 transcription to sensory organ precursor cells: a transcriptional network**

We have identified two independent regulatory pathways that restrict *pax-6/vab-3* transcription to the sensory organ precursor cells. One pathway is a transcriptional cascade involving a Pax-2-related transcription factor, EGL-38, and a zinc finger protein, LIN-48. Ectopic expression of *vab-3* in the presumptive U cell lineage in *egl-38* and *lin-48* mutants indicated that each of these factors can genetically, if not directly, repress *vab-3* expression. The finding of ectopic *vab-3* expression in *lin-48* mutant males was
particularly interesting, as socket-like cell types normally produced from the B.a lineage are derived from the U cell in these males (Chamberlin et al., 1999; Jiang and Sternberg, 1999). This suggests that, in addition to the requirement for vab-3 activity in B.a to produce sensory structure cell types, vab-3 may be sufficient for generating sensory organ cell types in a non-sensory cell lineage. This would be analogous to the role of Pax-6 in the specification of eyes in other animals, where Pax-6 misexpression can cause ectopic eye formation (Halder et al., 1995). Additionally, the finding that egl-38 restricts the expression domain of vab-3 has evolutionary implications. In the developing mouse nervous system, Pax-6 and Pax-2 engage in reciprocal inhibition of gene expression, helping to define the boundary between the optic stalk and optic cup (Schwarz et al. 2000). This raises the possibility that Pax-2 repression of Pax-6 expression is an evolutionarily conserved mechanism of Pax-6 regulation.

We also observe a difference in temporal vab-3 misexpression in egl-38 and lin-48 mutants. In wild-type males, Y.p expresses vab-3 only through the second larval stage, while expression in the B.a lineage persists through larval development. Ectopic vab-3 expression can be observed in the presumptive U cell of young egl-38(sy294) males. However, expression is rarely observed prior to U cell division in lin-48 mutant males. Furthermore, expression of vab-3 in egl-38 mutants peaks in mid-L2 stage, and declines until GFP is no longer visible by early L3. In contrast, lin-48 mutants express vab-3 more robustly as the animals progress into the L3 stage and beyond. One attractive explanation for these differences is that the presumptive U cell of egl-38(sy294) males is transformed to a Y.p-like cell fate, while in lin-48 males it exhibits a B.a-like fate transformation. This interpretation is consistent with cell lineage and cell fate analyses of
lin-48 and egl-38 (Chamberlin et al., 1997; Chamberlin et al., 1999). We hypothesize that the expression phenotype of egl-38; lin-48 double mutants reflects a combination of gene expression changes reflecting both U-to-B.a and U-to-Y.p cell fate transformations. Since the egl-38; lin-48 double mutants do not recapitulate the lin-48 mutant phenotype, we interpret that egl-38 does not only influence vab-3 expression through its effect on lin-48. As lin-48 is the only known target for EGL-38 in the hindgut, it is unclear whether lin-48-independent regulation of vab-3 involves direct repression of vab-3 by EGL-38, or if there are other transcription factors that might mediate this egl-38 function.

### 2.4.5. Negative regulatory inputs restrict pax-6/vab-3 transcription to sensory organ precursor cells: Wnt signaling

We have found that a second regulatory pathway that restricts pax-6/vab-3 expression and the sensory organ precursor cell fate is a Wnt signaling pathway. Several components of a Wnt signaling pathway have been found to influence the asymmetric division of the male B cell (Wu and Herman, 2007). The mutants with the most severe disruption of normal B cell asymmetry are mutants in the Wnt ligand-encoding gene, lin-44, and the Frizzled-type receptor gene, lin-17. Our results show that in these mutants, vab-3 expression correlates with a B.a-like cell fate. Many additional proteins participate in the asymmetric B cell division, including non-canonical Wnt pathway components such as RHO-1/RhoA and LET-502/ROCK (Wu and Herman, 2006). The mechanism by which Wnt signaling impacts vab-3 expression remains unclear. It is possible that the TCF/LEF transcription factor POP-1 directly mediates the observed transcriptional repression. Alternatively, the vab-3 expression decision may be strictly dependent upon
B cell polarity, and therefore, alterations in the \textit{vab-3} expression pattern may reflect the altered B cell polarity observed in Wnt signaling mutants.

The B.a-to-B.p cell fate transformation in \textit{vab-3} mutant males suggests that B.p is the default epithelial state of a B cell daughter. More generally, we can say that \textit{pax-6/vab-3} modifies the B.a cell to become a sensory organ precursor. Mutant phenotype data indicate that Wnt signaling acts to inhibit the organ precursor state in favor of the default state. Consistent with this idea, the LIN-44 protein is produced in the posterior of the animal, nearer the daughter that will adopt the default cell fate (Herman et al., 1995). Interestingly, Wnt-mediated repression of \textit{Pax-6} expression also occurs in the fly eye (Pichaud and Casares, 2000; Bess and Casares, 2005), although there are clear mechanistic differences in comparison to \textit{C. elegans} B cell regulation. In each case, however, Pax-6 activity, which can induce a sensory organ-specific developmental program, is spatially limited through a transcriptional boundary established by directional Wnt signaling.

Our work also illustrates how \textit{egl-5} and the Wnt pathway can function together differently in different cells in \textit{C. elegans}. For example, in the nearby ventral blast cells P11.p and P12.p, Wnt and EGF signaling converge to activate \textit{egl-5} expression in the presumptive P12.p, thereby preventing it from adopting the P11.p fate. \textit{egl-5} and Wnt mutants display the same developmental defect, a P12-to-P11 fate transformation (Jiang and Sternberg, 1998; Maloof et al., 1999). In the B cell, however, \textit{egl-5} and \textit{lin-17} mutants display opposite phenotypes. Whereas \textit{lin-17(n671)} males generate two B.a-like daughters, \textit{egl-5(n486)} males generate two daughters more like B.p. We suggest that \textit{egl-5} may be necessary to specify the B cell fate, and that Wnt acts to pattern B cell division,
rather than controlling egl-5 activity itself. In the case of egl-5(n486); lin-17(n671) double mutants, for instance, the loss of vab-3 repression in B.p is masked, as the cells are not competent to transcribe vab-3. A similar relationship between Wnt and HOM-C genes is observed in another asymmetric cell division in the C. elegans tail, the T cell. In the T cell, an egl-5 paralog, nob-1, is necessary for expression of a fate determinant (psa-3) in the T lineage, and Wnt signaling controls asymmetric psa-3 expression between daughter cells (Arata et al., 2006). Altogether, it is clear that HOM-C genes and Wnt signaling can function together in different ways, depending on the specific developmental decision.

2.4.6. Implications for the regulatory logic controlling pax-6/vab-3 gene expression

We have characterized regulatory inputs into Pax-6 transcription in sensory organ precursors in C. elegans. We uncover a domain refinement mechanism that underlies the precise expression pattern for pax-6/vab-3. Our findings support the idea that facets of the regulatory network governing Pax-6 expression are shared amongst species. Whether this shared regulatory logic arises from conservation of cis-element—trans-acting factor interactions, or is a case of convergent evolution of genetic regulations, these pathways are clearly of widespread importance in the control of Pax-6 expression during sensory organogenesis. Is this a feature specific to Pax-6, or something more generally true of Pax gene regulation in other organogenic processes? Have cancer cells expressing Pax factors overcome multiple repressive inputs in order to express these factors and maintain
an undifferentiated, anti-apoptotic state? As evidence highlighting the key role for Pax-6 and other Pax genes in organogenesis and pathology accumulates, these will be important questions to address.
Figure 2.1. Expression of a vab-3::gfp reporter transgene
(A,C,F,I) Nomarski DIC micrographs of males. In all photos, anterior is left. (D,G,J) Epi-flourescent images of the same animals. (A) A diagram of the four post-embryonic blasts of the male hindgut and their progeny in early L2 and mid L3, and micrograph of the adult male tail. Cells of the B cell lineage (gray) generate the copulatory spicules (arrows). The Y cell progeny (black) also generate sensory structures (the post-cloacal sensilla) in adults. (B) A diagram of the genomic structure of the C. elegans pax-6 locus. Coding exons are black, with parentheses marking mab-18-specific exons. The exons coding for the DNA-binding paired domain (PD) and homeodomain (HD) are noted. Dashed arrows denote two functionally distinct pax-6 transcripts, vab-3 and mab-18. Sequences contained in the vab-3::gfp reporter transgene guEx356 are indicated as follows: solid line denotes the included vab-3 genomic sequence, gray box denotes GFP coding sequence, and dotted line denotes unc-54 3’ UTR. (C,D) L2 larval vab-3 expression in unidentified neuronal cells of the head. This expression begins in embryogenesis, persists throughout the animals’ lifespan, and was used as an internal control for hindgut expression experiments. (E,H) Diagrams of the early L2 (E) and mid L3 (H) male hindgut. (E,F,G) The B and Y cells have undergone one asymmetric cell division. Nuclei are labeled. vab-3 is expressed in the sensory organ precursors B.a and Y.p. (F,G) The dotted circle indicates the position of U, which lies out of the focal plane. (H,I,J) The B.a and Y.p precursors have generated eight progeny. Nuclei of B.a descendents are labeled with greek letters. vab-3 expression is robust in all B.a cell types. (I,J) The ε and ζ lateral pairs are out of the plane of focus (dotted circles), as are the Y.p descendents, which have differentiated and have diminished vab-3 expression.

Continued on next page…
Figure 2.1. continued.
Figure 2.2. Ectopic expression of vab-3 in the U blast cell in egl-38 and lin-48 mutants

(A,C) Nomarski photomicrographs and (B,D) corresponding epi-fluorescent images of early L2 mutant males. Anterior is left, and nuclei are labeled. ‘l’ designation in cell labels denotes left daughter of the corresponding cell. (A,B) lin-48(gv4) and (C,D) egl-38(sy294) males exhibit ectopic vab-3 expression in the presumptive U cell. The B.p cell in (A,B) is out of the plane of focus. (E,F) Population-based quantification of vab-3 expression in the U cell lineage. (*) lin-48 expression data includes data from lin-48(sa469) and lin-48(gv4) males. (**) egl-38; lin-48 denotes an egl-38(sy294); lin-48(gv4) or egl-38(sy294); lin-48(sa469) genotype. No expression differences were observed between animals bearing the different lin-48 alleles. (E) Ectopic U lineage expression is infrequent in comparison to B.a lineage expression. egl-38 and lin-48 mutants exhibit differences in ectopic U lineage expression. Double mutants exhibit an intermediate expression phenotype. Data correspond to animals from late L1 to mid L3 stage. All samples, n ≥ 73. (F) Temporal profile of ectopic U lineage expression from the same population of animals represented in (E). Larval stage ‘time points’ are grouped according to the number of B.a progeny (1, 2, 4, and 8), which are present in the indicated developmental stage. Expression increases over time in lin-48 mutant males, but is relatively stable in egl-38(sy294) animals. The double mutant profile does not suggest a clear epistatic relationship between egl-38 and lin-48 with respect to vab-3 expression.
Figure 2.2. continued.
Figure 2.3. Correlation between vab-3 expression and the B.a cell fate
(A,C) Nomarski photomicrographs and (B,D) corresponding epi-fluorescent images of early L2 mutant males. Anterior is left, and nuclei are labeled. (A,B) Following a symmetric cell division, both B cell daughters in lin-17(n671) males express vab-3. (C,D) The B.p lineage, rather than the B.a lineage, expressed vab-3 in lin-44(n1792) males. In this L2 male, B.p has divided, while B.a has not. This is analogous to a reversal of the wild-type male pattern in which B.a divides before B.p. (E) Correlation chart between vab-3 expression and B.a cell fate patterning. In animals showing normal asymmetric B cell division, vab-3 is expressed in the B.a and its descendents. Typically, after abnormal B cell divisions, expression of vab-3 is observed in B.a-like, but not B.p-like, cells and descendents.
Figure 2.4.  

(e)l-5 is required for proper B cell division and vab-3 expression

(A,B,C,E) Nomarski photomicrographs of early L2 males.  (A,B) Anterior is left, nuclei are labeled, and cell boundaries are highlighted (dotted lines).  (A) Wild-type male B cells undergo asymmetric cytokinesis, producing daughter cells of unequal size.  (B) The B cell daughters of egl-5(n486) males are frequently equal in size, and appear B.p-like.  

(D,F) Epi-fluorescent images of (C,E). vab-3 expression is observed in the B.a cell of wild type, but not egl-5(n486) males (arrows). Expression is observed in head neurons of both strains (brackets).  (F) Image overexposed to highlight lack of GFP in B.a.
Table 2.1. *egl-5* is required for *vab-3* expression in hindgut blasts

<table>
<thead>
<tr>
<th>Genotype*</th>
<th>n</th>
<th>B.a</th>
<th>B.p</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>139</td>
<td>86</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>egl-5(n486)</em></td>
<td>42</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>lin-17(n671)</em></td>
<td>42</td>
<td>86</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td><em>lin-17(n671); egl-5(n486)</em></td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td><em>lin-48</em>*</td>
<td>47</td>
<td>91</td>
<td>-</td>
<td>26</td>
</tr>
<tr>
<td><em>lin-48(gv4); egl-5(n486)</em></td>
<td>35</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

*Full genotype includes *unc-119(e2498); him-5(e1490); guEx356

**Data from both *lin-48(sa469)* and *lin-48(gv4)* males

Figure 2.5. Domain refinement model for regulation of *pax-6/vab-3* expression in sensory organ precursors of the *C. elegans* male

*vab-3* expression (black) in the region requires activity of the HOM-C gene, *egl-5* (gray). Restriction of the *vab-3* expression domain is accomplished in part by combined *egl-38* and *lin-48* function in the U cell lineage, and via *Wnt/lin-44—Frizzled/lin-17* signaling in the B.p cell.
CHAPTER 3

EGL-38 TARGET GENE IDENTIFICATION VIA GENE EXPRESSION MICROARRAYS AND CHARACTERIZATION OF EGL-38 IMMUNE FUNCTION

3.1. INTRODUCTION

Understanding how a single Pax protein can regulate the formation of very different organs is an important question that has been a focus of interest for many years. For instance, the well-studied Pax6 gene is necessary for both eye development and specification of insulin-producing β cells in the pancreas (rev. in Dohrmann et al., 2000). One attractive explanation for how an individual Pax gene coordinates development in disparate tissues is that Pax factors act in a combinatorial fashion, and that interaction with different proteins results in transcriptional regulation of different sets of target genes. Target regulation by complexes such as Pax5—Ets1 in B lymphocyte specification (Fitzsimmons et al., 1996), Pax2—Hox11—Eya1 in early kidney development (Gong et al., 2007), Pax6—Sox2 in lens development (Kamachi et al., 2001), and Pax6—Cdx2 in pancreatic cell function (Hussain and Habener, 1999) provide strong evidence that combinatorial expression is a standard mechanism for generating
Pax factor functional specificity. Therefore, understanding the tissue-specific nature of 
*Pax* gene networks, including Pax protein cofactors and target genes, is critical to 
understanding and addressing specific developmental defects associated with *Pax* gene 
mutations.

Despite a large body of work describing the biological roles of different Pax 
factors, and some characterization of transcriptional mechanisms, much is still unknown 
about *Pax* gene networks. Is there a general set of target genes that function broadly in 
response to Pax activity that carry out general developmental functions, or are all targets 
tissue-specific? What are the intrinsic features of Pax proteins that provide for tissue-
specificity, and how are those features converted into target gene expression? To address 
such questions about the nature of *Pax* tissue-specificity and target gene profiles, *C.
elegans* provides a good experimental system for multiple reasons. First, the highly 
conserved nature of Pax proteins among animals indicates that principles of Pax function 
in lower organisms can be applicable to humans (rev. in Gruss and Walther, 1992). In 
addition, while most *Pax*-dependent biology in vertebrates is very complex, the invariant 
cell-lineages, transparent anatomy, and relatively simple organ structures of *C. elegans* 
allow for a more tractable system for genetic manipulation of organogenesis and 
functional dissection of *Pax* activity. The *C. elegans* genome has representatives from 
each *Pax* gene subclass, and these genes are known to regulate diverse developmental 
processes, as they do in vertebrates (Hobert and Ruvkun, 1999).

The *Pax2/5/8* subclass gene *egl-38* is the most well studied *Pax* gene in *C.
elegans*. Various non-null alleles of *egl-38* have been found to disrupt protein function in 
a tissue-preferential manner, thus providing unique tools for the study of EGL-38 activity
in vivo and in vitro (Zhang et al, 2005). Genetic studies have shown that egl-38 acts to promote cell survival in somatic and germline tissues and to coordinate the development of the egg-laying system and the rectum (Park et al., 2006; Chamberlin et al., 1997; Rajakumar and Chamberlin 2007). The egg-laying system provides a particularly interesting model for understanding egl-38 function, as this gene acts in both the vulva and the uterus during their development. Coordination between the two tissues is disrupted by the mutation egl-38(n578), while other egl-38-dependent processes remain relatively unaffected. Specifically, in egl-38(n578) mutants, one cell type of the ventral uterus (uv1) is mis-specified as another (uterine seam, or utse), and the dorsolateral vulF cells of the vulva fail to separate and form an open channel through which eggs can pass (Rajakumar and Chamberlin, 2007). The uv1 specification defect is due to the combination of an autonomous uv1 defect and a loss of LIN-3/epidermal growth factor (EGF) signaling from the vulF cells to the pre-uv1 cells. In contrast, the failure of vulF cells to undergo proper morphogenetic separation is due solely to a vulF-intrinsic egl-38 deficiency. Although these processes are well described at the cellular/tissue level, the precise molecular mechanisms that mediate EGL-38 activity remain unresolved.

The studies discussed above demonstrate that egl-38 can mediate diverse organogenic and cellular processes, and open up the possibility that additional, uncharacterized roles for egl-38 remain undiscovered. To fully identify and elucidate the tissue-specific roles of egl-38 in C. elegans, more must be known about the transcriptional targets of the EGL-38 protein. That is, through identification of target genes, the molecular and cellular mechanisms by which EGL-38 imparts its physiological effects can be better understood. Furthermore, identification of multiple target genes is a
requisite for understanding more broad mechanisms of transcriptional control, such as defining features of tissue-specific combinatorial regulation. To date, few target genes of EGL-38 have been discovered. EGL-38 is known to directly activate lin-48 transcription in the rectum, and ced-9 in the germline (Johnson et al., 2001; Park et al., 2006). Clearly, these genetically identified targets represent a small fraction of genes one would expect to be regulated by EGL-38. The lack of targets led us to ask whether we might identify more using a systemic experimental approach.

Gene expression profiling using microarrays is a powerful technique for exploring gene expression in C. elegans (Lee et al., 2004; Lamitina, 2006). Consequently, we took a genomic expression analysis approach to identify EGL-38 target genes that may function in established and/or undiscovered egl-38-dependent processes. We utilized a carefully chosen combination of strains exhibiting distinct alterations in EGL-38 activity to take full advantage of this high-throughput expression profiling strategy. Here, we report expression microarray data defining multiple sets of potential target genes. Further, we used a microarray cross-referencing approach to identify and sort a robust group of candidate EGL-38 targets. To complement the information obtained in silico, GFP reporter studies were used to validate our microarray design.

In addition to identifying new genes expressed in response to EGL-38, this work has uncovered a novel function for EGL-38 in mediating the innate immune response in C. elegans. A collaborator has shown that EGL-38 mediates the susceptibility of hermaphrodites to rectal infection by the gram-positive bacterium Microbacterium nematophilum (H. Jia, personal communication). In contrast, I show that egl-38 activity has no effect on the survival of C. elegans in the presence of a lethal gram-negative
bacterium *Serratia marcescens*, illustrating functional specificity for *egl-38* in the immune response. Several genes identified by our microarray cross-referencing approach display *egl-38*-dependent rectal expression, and reducing the activity of three of these genes using RNAi influences the extent of rectal infection by *M. nematophilum*. In total, this work has defined a previously undiscovered innate immune function for a *Pax* factor, and has provided a textured dataset that not only outlines potential EGL-38 target genes, but carries information related to the tissue- and process-specificity of EGL-38 action.

3.2. MATERIALS AND METHODS

3.2.1. Genetic strains and strain construction

*C. elegans* strains were cultured under standard conditions (Brenner, 1974; Sulston and Hodgkin, 1988). All strain maintenance and experiments were done at 20°C using the *E. coli* OP50 bacterial strain as a food source, unless otherwise noted. The following mutations were used: Linkage Group III (LGIII): *unc-119(e2498)*; LGIV: *egl-38(sy294), egl-38(n578), nT1(IV; V)*; LGV: *nT1(IV; V)*. The unmapped integrated transgene *guls21 [unc-119(+); hsp-41::egl-38]* was also used (Zhang et al., 2005). All strains used can be found in Appendix B. Additional genetic information is available at [http://www.wormbase.org](http://www.wormbase.org).

Transgene expression was assessed in the *egl-38(n578)* background by crossing wild-type transgenic lines with strain CM1501 [*unc-119(e2498); egl-38(n578)/nT1*]. Specifically, ten L4 hermaphrodites from each transgenic strain were heat-shocked two hours at 30°C. Male progeny arising from non-disjunction of X were picked with siblings
to seed a mating population. Males from these populations were crossed to CM1501, and the resulting non-Unc F1 progeny were selected for self-crossing. F2 animals which displayed the Egl phenotype were then chosen as founders for the transgenic egl-38(n578) strains.

3.2.2. RNA preparation

Total RNA was extracted from synchronous populations of 4 C. elegans strains: N2 wild type, CM1452 [egl-38(n578)], CM1421 [egl-38(sy294)], and CM748 [guls21(hsp-41::egl-38)]. These strains were synchronized according to standard alkaline hypochlorite methods (http://arojanlab.ucsd.edu/protocols/hatch-off.htm), with the exception that the bleach solution contained 400 µl NaOH, 1ml NaOCl, and 6 ml H2O. For most genotypes/conditions, 2 100mm pre-starved plates were synchronized, but more plates were required to yield an equivalent number of egl-38(sy294) embryos. Upon synchronization, worms were grown for 42 hours post-feeding, harvested as L4 animals in saline buffer M9, and frozen at -80°C until all samples had been collected. For heat shock conditions, animals were grown to 40.5 hours, subjected to heat shock at 35°C for 30 minutes, and allowed to recover at 20°C for 1 hour before harvesting at 42 hours. L4 populations from each genotype/condition were isolated in triplicate from independently synchronized populations. To recover total RNA from lysates, Trizol extraction was performed according to the manufacturer’s protocol (Invitrogen). The recovered RNA was subsequently purified using RNeasy kits (Qiagen), and sample quality was tested. RNA concentration and degradation were analyzed with the RNA 6000 Nano Assay using an Agilent bioanalyzer 2100 (Agilent).
By employing induced EGL-38 expression as a gain-of-function condition, we hoped to enrich for genes that respond to EGL-38 - a category that includes direct target genes – and select against genes that are altered as a secondary effect of the altered development in egl-38 mutants. For the induced expression experiments, the time from initial heat shock to worm collection and subsequent RNA isolation was in the range of 2-3 hours. We chose the 2-hour recovery timepoint based upon data for the direct EGL-38 target lin-48, whose transcript abundance peaks at that time (D.H. Park, personal communication). However, increased lin-48 transcript levels were detectable by semi-quantitative RT-PCR as early as 0 hours post-heat shock (30 minutes after exposure to 30°C), indicating that by 2 hours post-heat shock, target genes of LIN-48, which is a Zn-finger transcription factor, could potentially be responding. Therefore, although we are confident that many of the EGL-38-responsive genes represent direct targets, our method does not rule out the potential for indirect downstream target genes in the candidate list.

3.2.3. Microarray hybridization and data analysis

Five μg RNA of each sample was utilized by the Ohio State University Microarray Shared Resource (MASR) to perform hybridization on Affymetrix C. elegans genome Genechips (Affymetrix). This facility uses the Affymetrix GeneChip System including GeneChip Hyb-Station Oven 320/640, Fluidics Station 450, and GeneChip Scanner 3000. Raw microarray data were analyzed by the robust multi-chip average algorithm (RMA; Irizarry et al., 2003), provided by Biometric Research Branch (BRB) software (http://linus.nci.nih.gov/BRB-ArrayTools.html; developed by Dr. Richard Simon and Amy Peng Lam). Genes for which transcript levels were increased or decreased ≥1.5
fold relative to wild type based on a 95% confidence interval were selected for further analysis. Initial lists were based on Affymetrix probe set identifications, and we arrived at final gene lists by elimination of probe sets corresponding to predicted duplicate genes, pseudogenes, transposons, and sequences found at multiple locations in the genome (http://wormbase.org). A short list of genes to be validated was assembled by applying two biological filters. Selection of a gene for the list required that transcript abundance was significantly altered in both egl-38(n578) and hsp-41::egl-38 experiments. Further, the direction of transcript abundance change had to be opposite in mutant compared to the heat-induced condition.

### 3.2.4. GFP reporter transgene construction

GFP reporter genes were generated via a PCR fusion approach (Hobert, 2002). GFP coding sequence was amplified from the vector pPD95.69 (kind gift of A. Fire) using primers PR1945 and PR1946 (see Appendix A) to generate a 1.8 kb product. This fragment included sequences upstream of the GFP coding region in order to incorporate the synthetic intron, which aids in reporter gene expression (Fire et al., 1990). In parallel, regulatory sequences upstream and including the annotated start codons of each candidate were amplified. The length of the amplicon varied among the candidates, and was chosen by two criteria. When there were no genes within 5kb of a candidate’s 5’ end, primers with a predicted Tm ≥ 55°C were chosen as near to the 5kb mark as possible (2 of 13 genes). In most cases, other genes were present within 5kb upstream of the candidate, so the entire region between coding sequences was chosen (11 of 13), again using the constraint of primers with a Tm ≥ 55°C. To build reporter genes for the three
candidates that are members of operons (H19N07.3, C41H7.3, C41H7.5), the same regulatory sequence selection was done, except that the most 5’ member of the operon was used as the reference point, rather than the operon member in our gene list.

To create the GFP fusion products from these PCR fragments, the GFP and regulatory sequence templates were treated separately. The GFP product was purified by QIAquick PCR purification kit (Qiagen) and used for all second round reactions. For the gene-specific PCR reactions, if the correct bands were discernable following agarose gel electrophoresis, and present at a reasonable concentration, 1 µl of these PCR reactions was used directly as template for the fusion PCR step. For this step, primers nested within the 3’ end of the GFP amplicon and the 5’ end of the candidate amplicon were used to increase the specificity and efficiency of amplification. The DNA concentration of successfully-generated GFP reporter fusion genes was roughly estimated in agarose gels by comparison to the manufacturer-specified concentration of the 3.0kb fragment in a GeneRuler 1kb DNA Ladder (Fermentas) as a reference. All primer information can be found in Appendix A.

DNA transformation of RH10 [unc-119(e2498)] worms was performed according to standard germline microinjection methods (Mello et al., 1991). GFP reporter fusion gene PCR reactions were directly used to start injection mixtures according to the visual concentration estimates. These concentrations were also applied to determine the amount of pTJ1043 [unc-119(+)] marker DNA to be used, which in all cases was roughly 10-fold less than PCR product concentration. For most reporter gene constructs, 2-3 transgenic lines were established to confirm the reliability and reproducibility of transgene expression. However, only one strain per reporter was used for wild-type data collection,
as well as for establishing and scoring *egl-38(n578)* transgenic strains. All strain information can be found in Appendix B.

### 3.2.5. cDNA cloning and RNA interference

To generate candidate gene knockdown via RNAi, we amplified cDNA for each genes, using total RNA from mixed stage worms as a template. Primers were designed to include all coding sequence except for the stop codon, and were tagged with restriction enzyme sites (see Appendix A). Addition of an NheI or SpeI cutting site at the 5’ end, and KpnI, SacI, or EcoRV at the 3’ end, facilitated cDNA cloning into the vector pBlueScript. Single-stranded RNAs were generated from these cDNA vectors via Ampliscribe T7/T4 in vitro transcription reactions, according to manufacturer’s protocol (Epicentre Biotechnologies). These RNAs were then purified using RNeasy columns (Qiagen), and annealed to form dsRNA in annealing buffer (1mM Tris pH 7.5, 1mM EDTA). We used a microinjection strategy for our RNAi experiments (Fire et al., 1998). Briefly, young adults were injected in each gonad arm with dsRNA at a concentration of 1μg/ul, or with annealing buffer alone. After initial controls were tested using wild-type hermaphrodites, a strain carrying *eri-1(mg366)* was used to enhance the RNAi response. One day post-injection, the injected animals were transferred to new plates to ensure all of the progeny to be analyzed would be exposed to the RNAi treatment.
3.2.6. *S. marcescens* killing assays

*C. elegans* survival analysis was performed similar to previous studies (Pujol et al., 2001), with one exception. Plates were maintained at 20°C to maintain consistency with the egl-38(RNAi) *M. nematophilum* infection experiments. For each trial, 50 L4 hermaphrodites were placed on plates seeded with either *E. coli* OP50 or *S. marcescens* DB11. The animals were scored each successive day as alive, dead, or not scorable. The latter category included all animals/carcasses that could not be found, were found dead on the plastic, or were damaged by picking. These animals were not counted towards the total animals scored. In one wild type—DB11 trial, 33 of 50 were scorable; in all other cases, 40 or more of 50 were scorable. Animals were scored as alive if they showed normal locomotion and/or responded to touch. Animals were counted and transferred to new plates each day until only 5-10 worms remained alive, at which point the same plate was used to complete the assay.

3.3. RESULTS

3.3.1. Identification of potential EGL-38 target genes using gene expression arrays

The *C. elegans* Pax-2/5/8 protein EGL-38 regulates many developmental processes, but to date, few target genes have been identified (Chamberlin et al. 1997, Johnson et al., 2001; Park et al., 2006). In order to better understand the mechanisms by which *egl-38* mediates developmental patterning in different cells and tissue types, we set out to identify *egl-38*-responsive genes using a genomic expression array strategy. To generate a comprehensive preliminary list of potential target genes, we employed both
gain-of-function (gf) and loss-of-function (lf) conditions. Previous studies have shown that different hypomorphic alleles of egl-38 preferentially disrupt functions in different tissues (Chamberlin et al., 1997; Zhang et al., 2005). Since null alleles of egl-38 are lethal, we independently analyzed two strains carrying different hypomorphic egl-38 alleles, n578 and sy294, which are known to disrupt distinct sets of egl-38 functions. Specifically, egl-38(n578) mutants exhibit several defects in the development of the egg-laying system, while egl-38(sy294) animals display stronger defects in rectal cell development. As no hyperactive alleles of egl-38 are known, we utilized induced expression of EGL-38 under control of the heat shock promoter, hsp-41. Induced expression confers an egl-38 gain-of-function phenotype in germline cell death assays (Park et al., 2006).

We compared the gene expression profiles between heat-shock treated wild type (N2) vs. heat-shock treated hsp::egl-38 transgenic worms, N2 vs. egl-38(n578), and N2 vs. egl-38(sy294). Using the robust multi-chip average (RMA) algorithm (Irizarry et al., 2003) and Biometric Research Branch software (http://linus.nci.nih.gov/BRB-ArrayTools.html), differentially expressed genes were defined according to liberal statistical cutoffs (|fold change| ≥1.5, p ≤ .05). These experiments identified 550 genes altered in response to induced expression of EGL-38, 727 genes altered in egl-38(n578) mutant animals and 1019 genes altered in egl-38(sy294) animals (Figure 3.1.).

To analyze how successful these microarray experiments were in retrieving appropriate sets of candidate genes, we first searched the gene lists for known EGL-38 target genes. Transcript levels of the well-studied EGL-38 target gene lin-48 were significantly reduced in the wild type vs. egl-38(sy294) experiment, but not in the others.
This result confirmed predictions about the allele specificity aspect of our microarray design. EGL-38 activates transcription of *lin-48* in a set of rectal cells, and *lin-48::gfp* expression is absent in *egl-38(sy294)* mutants, but not in *egl-38(n578)* animals (Johnson et al., 2001). While *lin-48* levels were not found to be upregulated in the induced EGL-38 experiment, there is evidence that *lin-48* is regulated by EGL-38 in combination with at least one other factor. Thus, a limited pool of the other factor could prohibit increased activation of *lin-48* transcription (Johnson et al., 2001; Zhang et al., 2005). Although EGL-38 can also directly regulate *Bcl-2/ced-9* expression, this gene was not differentially represented in any of our experiments. This may be because our RNA samples were derived from L4 populations, prior to the onset of germline apoptosis (Gumienny et al., 1999).

Although no other EGL-38 targets have been confirmed, we also reviewed our gene list for genes acting in *egl-38*-dependent tissues, and found many of them to be altered in response to *egl-38* activity. *ida-1* is expressed in uterine uv1 cells, that are missing in *egl-38* mutants (Zahn et al., 2001, Rajakumar and Chamberlin, 2007), and *ida-1* transcripts are altered in both *egl-38* mutant backgrounds. Similarly, *peb-1* and *ceh-6* have been shown to express and to function in the hindgut, and the abundance of these transcripts is altered in *hsp-41::egl-38* and *egl-38(n578)* strains, respectively (Thatcher et al., 2001; Burglin and Ruvkun, 2001). Thus our microarray approach was successful in selecting known genes that function in *egl-38*-dependent tissues.
3.3.2. Mining for prospective EGL-38-responsive genes by a microarray cross-referencing approach

Compiling statistically filtered lists of differentially expressed genes from microarray data can be informative regarding transcriptional regulation patterns, as evidenced by the discovery of a novel immunological role for EGL-38 in *C. elegans* (H. Jia, personal communication; this work). Often, stringent statistical criteria can also be applied to identify the best candidate genes for validation. However, selection of individual EGL-38 target genes through statistical analysis alone was not possible in this case. Inherent challenges with the genetic conditions and utilization of whole worms as a transcript source limited the resolution with which we might distinguish true and false positives. To overcome these statistical limitations, we employed an array cross-referencing approach. In particular, we required that candidate transcripts be altered reciprocally in *gf* and *lf* experiments. As with any microarray experiment, one must be careful to balance the proportion of false positives and negatives according to the experimental goals. The primary objective of our study was identification of new EGL-38 targets genes, and so we wished to minimize the number of false positives. We concede that we may have eliminated a significant number of true positives through our selection process. However, we favored this strategy for evaluating a first set of candidate genes, as it provides for greatest enrichment of true positives. By requiring reciprocal responses from the EGL-38 *lf* and *gf* experiments, we have effectively created a biological filter with which to refine the statistically determined candidate gene lists.
We chose to focus our array validation efforts on transcripts affected in the *egl-38(n578)* background, rather than those impacted by *egl-38(sy294)*. RNA samples were collected from L4 populations, the developmental timeframe in which *egl-38* is functioning in the egg-laying system. Therefore, the allele known to preferentially disrupt development of the egg-laying system stood out as the most appropriate dataset in which to validate our cross-referencing approach. Of the 55 transcripts that respond to both induced EGL-38 and *egl-38(n578)*, 14 were altered in a reciprocal fashion (Table 3.1.). Three individual Affymetrix probesets represent transcripts predicted to be members of the same operon (*C41H7.5, C41H7.4, and C41H7.3*), so these were considered as only one candidate. In addition, one other transcript is annotated as part of an operon. In the microarray experiments, *H19N07.3* was down-regulated in *egl-38(n578)* mutants and up-regulated in response to induced EGL-38 expression. We looked at the array expression values of other transcripts within the same operon, and although these did not pass our statistical filter, each of these transcripts responds in a similar way to *H19N07.3* (data not shown). Similarly, the abundance of members of the *C41H7* operon all exhibited a positive correlation with EGL-38 activity, being increased in the *hsp-41::EGL-38* experiment and decreased in the hypomorphic condition. Taking the distinct genes and operon members together, the array cross-referencing method provided 13 primary transcripts for further investigation.
3.3.3. **Cell-specific expression of candidate genes is responsive to egl-38**

Initial validation of the 13 candidate genes was addressed by reporter transgene analysis. Green fluorescent protein (GFP) reporter transgenes were synthesized via a PCR-fusion method (Hobert, 2002), using roughly 1-5 kb of regulatory sequence upstream of these genes, or in two cases, upstream to the operon containing the gene of interest. The expression patterns were observed in wild-type L4 hermaphrodites, and are summarized in Table 3.2. Strikingly, GFP expression was repeatedly observed in four tissues known to require egl-38 function. Of the ten candidate reporters examined, six (60%) displayed significant expression in the vulva, three (30%) in the excretory system and the uterus, and eight (80%) in the rectum (Table 3.3.). In contrast, roughly 5, 2, 1, and 2% of the 9,260 curated *C. elegans* gene expression patterns include cells of these four tissues, respectively ([www.wormbase.org](http://www.wormbase.org), WormBase Release WS180). Together, nine of ten transgenes express in the vulva and/or rectum (Figure 3.2.). Although the sample size is small, we conclude that our array cross-referencing approach enriched for genes that express in egl-38-dependent tissues.

The expression patterns we observed revealed interesting similarities, as well as differences. Although several genes had unique patterns of expression among egl-38-responsive tissues, most also exhibited expression in shared subsets of epithelial tissue not known to require egl-38 activity. In the most extreme case, reporters for *F18E3.11*, *F08F3.4*, and *H13N06.3*, exhibit a nearly identical epithelial expression pattern outside of the vulva and rectum (Table 3.2.). These genes appear to all be expressed in the hypodermal syncytium hyp7, the Pn.p cells, as well as some additional unidentified epithelia. All of these are absent or minimally expressed in the vulval lineages, while
F18E3.11 and F08F3.4 are noticeably missing in most rectal cell types, as well (Figure 3.2. E,H).

To further evaluate the microarray results, we examined the vulval expression of these reporter transgenes in the egl-38(n578) mutant background. Of the six transgenes showing vulval expression, five were responsive to lowered egl-38 activity (Table 3.3.). Furthermore, each of these transgenes behaved in the manner predicted by the microarray results. That is, reporters for the transcripts with reduced abundance in the egl-38(n578) arrays displayed reduced expression, while those representing transcripts with increased abundance in egl-38(n578) arrays displayed increased vulval expression. These results provide strong confirmation that the cross-referencing approach to choosing potential EGL-38 target genes was robust, and capable of identifying egl-38-dependent genes that are expressed in (or excluded from) egl-38-dependent tissues.

3.3.4. egl-38 influences the susceptibility of C. elegans to infection by the gram-positive M. nematophilum, but not the gram-negative S. marcescens

Apart from confirming the success of our microarray approach to the discovery of prospective EGL-38 target genes important for development, brief evaluation of the differentially expressed gene lists revealed enrichments of gene families previously linked to innate immunity and C. elegans-pathogen interactions. For example, C. elegans contains six pharyngeal gland toxin-related (phat) genes (www.wormbase.org, WormBase Release WS185), which contain ion transport inhibitor domains. This corresponds to 0.03% of the roughly 20,000 protein coding genes in C. elegans. Four, two, and one phat genes are found on the list of transcripts altered in egl-38(n578), egl-
38(sy294), and hsp-41::EGL-38 populations, respectively. These correspond to 16.7-, 6.0-, and 5.3-fold enrichments with respect to their representation in the genome. The enrichment of other gene classes known to contribute to innate immunity in *C. elegans* showed fold enrichments ranging from 1.5 to 17.9 in our initial array lists (Table 3.4.).

*C. elegans* provides a good model for the study of innate immunity and host-pathogen interactions (rev. in Mallo et al., 2002; Gravato-Nobre and Hodgkin, 2005). Human opportunistic pathogens such as *Salmonella* species and *Pseudomonas aeruginosa* can cause infections and induce innate immune responses in *C. elegans* (Labrousse et al., 2000; Aballay et al., 2000; Tan et al., 1999; Mahajan-Miklos et al., 1999), allowing for direct study of pathogen-specific virulence factors. Many genes that respond to bacterial infection have been identified through traditional genetic screens and expression microarray analysis (Gravato-Nobre et al., 2005; Yook and Hodgkin, 2006; O’Rourke et al., 2006). However, no sequence-specific transcription factors have yet been identified that mediate the transcriptional response to infection. Coupled with these facts, the enrichment of immune response gene families in our microarray experiments indicated that EGL-38 may control gene expression in response to pathogen exposure.

The gram-positive bacterium *Microbacterium nematophilum* is known to specifically infect the *C. elegans* rectum (Hodgkin et al., 2000), an organ in which egl-38 plays a critical developmental role. Additionally, egl-38 mutants exhibit a protruding rectum phenotype strongly resembling the *M. nematophilum*-induced Deformed anal region (Dar) phenotype. Together, these findings prompted us to hypothesize that EGL-38 could play a role in *M. nematophilum* infection by regulating the expression of a subset of immune response genes.
In a collaborative effort, we tested our hypothesis that EGL-38 influences the immune response. The infectivity of *M. nematophilum* in *egl-38* mutants was assayed previously (H. Jia, personal communication). Normally, culturing *C. elegans* on mixed bacterial lawns of *E. coli* and *M. nematophilum* results in significant colonization of the rectum, the Dar phenotype, and intestinal constipation. We chose to focus primarily on rectal colonization, as it is the most direct measure of bacterial infection. A series of rectal colonization experiments found that *egl-38(n578), egl-38(RNAi)*, and to a lesser extent *egl-38(sy294)* mutants showed significantly reduced bacterial infection compared to wild type (Appendix D). These results indicated that EGL-38 facilitates rectal infection by *M. nematophilum*.

Previous work examining the transcriptional response of *C. elegans* to different pathogenic species illustrated that each bacterium elicits a unique expression profile of immune response genes (Alper et al., 2007; Wong et al., 2007). Due to these findings, we hypothesized that the infection susceptibility conferred by EGL-38 may be pathogen-specific. Furthermore, the ability of EGL-38 to regulate expression of immune response genes, either directly or indirectly, suggested that this factor may serve a protective role against other pathogens. To address this question, we challenged *egl-38* mutants with another bacterium infectious to *C. elegans*. Although no other pathogens have been reported to infect the rectum specifically, many infect the digestive tract more generally (Gravato-Nobre and Hodgkin, 2005). The gram-negative species *Serratia marcescens* was chosen for three reasons: (1) it is a well characterized *C. elegans* pathogen, (2) gram-negative and gram-positive bacteria can elicit distinct host responses (for examples, see Yu et al., 2007; Scott et al., 2005), and (3) unlike *M. nematophilum, S. marcescens*
exposure is lethal (Pujols et al., 2001), and therefore is a more dangerous pathogen to \textit{C. elegans}.

Growth of worms on plates of the \textit{S. marcescens} strain DB11 has been demonstrated to cause a dramatic shortening of normal lifespan (Kurz et al., 2003). We find near 100% lethality by 10 days post-DDB11 exposure, compared to around 18 days for worms grown on \textit{E. coli} OP50 (Figure 3.3. A). We could not assess the survival curve of \textit{egl-38(n578)} mutants, which die within 2 days after reaching adulthood due to the Egl “bag of worms” phenotype. Therefore, we followed the survival of \textit{egl-38(RNAi)} and \textit{egl-38(sy294)} mutants in this study. Neither of these loss-of-function conditions displays a strong Egl phenotype, yet both were observed to have decreased \textit{M. nematophilum} infections. We found that \textit{egl-38(sy294)} mutants exhibit a decreased lifespan compared to wild-type animals (Figure 3.3. B). However, this lethality was observed in populations grown on OP50, and no difference was observed on the DB11 plates. In \textit{egl-38(RNAi)} mutants, the survival curves were roughly similar to wild-type curves on DB11 plates, but this lethality was unchanged in uninfected conditions (Figure 3.3. C). In both \textit{sy294} and RNAi mutant populations, many animals were observed to have strong rectal defects at the time of death (Figure 3.3. D). Together, these results indicate that reductions in \textit{egl-38} activity cause premature lethality, but that these animals are not notably susceptible to \textit{S. marcescens} killing. More generally, \textit{egl-38} does not appear to serve a protective role for at least one gram-negative pathogen, suggesting that EGL-38 does not confer broad sensitization/protection to gram-negative/positive pathogens, respectively.
3.3.5. **Candidate EGL-38 target genes display egl-38-dependent rectal expression**

The rectal expression of most candidate EGL-38 target genes we tested was initially curious, as *egl-38(n578)* had not been previously observed to have defects in rectal cell development. However, the discovery that these mutants have altered rectal infection suggested a possible reason for the recovery of rectally-expressing genes from the microarray cross-referencing strategy focused on *egl-38(n578)*. Considering this, we characterized the rectal expression of these genes in more detail. Panels A-H of Figure 3.4. illustrate the relative frequency of gene expression within each of the rectal cell types. For the candidate GFP reporters, each pattern is unique, with most exhibiting expression in multiple cell types. Taking all reporters in total, the intestinal-rectal valve (IRV) cells and the K/K’ left-right pair were the cells most commonly showing GFP fluorescence, followed by F and U. Other cells exhibiting transgene expression were P12.pa, B, and the three rectal epithelial cells, RepD, RepVL, and RepVR. Expression patterns in this tissue are diagramed in Figure 4I. From these patterns, we conclude that the genes whose transcript levels change in response to EGL-38 activity are expressed to various degrees in all rectal cell types.

*egl-38* functions in multiple cells of the rectum. In *egl-38(sy294)* mutants, the postembryonic blasts K’, F, and U develop abnormally, and an *egl-38::GFP* reporter is expressed in these and other rectal cells (Chamberlin et al., 1997; Johnson et al, 2001; X. Wang and H.M. Chamberlin, unpublished communication). Together with our GFP reporter expression results, these data strongly indicate that rectal expression of some of the candidate genes is regulated by *egl-38*. To address this hypothesis, we crossed each reporter transgene into the *egl-38(n578)* mutant background and scored L4 animals for
rectal expression. Of the eight genes with hindgut expression, four were predicted by the array results to be positively regulated by EGL-38. Three of the four display a statistically significant reduction in the proportion of animals with detectable rectal expression and/or the number of GFP(+) rectal cells per animal in egl-38(n578) animals. In contrast, expression did not significantly increase for the four genes predicted to be repressed in response to EGL-38. In fact, expression for each of them decreased in the rectal cells (Figure 3.5.). The full expression pattern of these reporter transgenes reveals the reason; while expression decreases in the rectal cells, GFP levels rise in many other tissues (e.g. intestine, seam, uterus; Table 3.2.). Altogether, these results suggest that the rectal expression of many genes identified by our microarray experiments is egl-38 dependent.

3.3.6. A subset of egl-38-dependent genes mediate susceptibility to M. nematophilum infection

Reporter expression studies confirmed that some genes identified by our microarray experiments are expressed in hindgut cells in response to egl-38. To address whether any of the candidates in our study are functionally involved in mediating the immune susceptibility function of EGL-38, we undertook RNAi knockdown studies for a subset of genes. phat-3, gst-21, F11E6.8, and C05C9.1 were chosen based upon predicted immune function of the encoded gene product.

Of the four egl-38-dependent genes tested, only phat-3(RNAi) animals failed to show a reduced infection phenotype (Figure 3.6.). gst-21(RNAi) and F11E6.8(RNAi) animals displayed the strongest resistance, with nearly half showing reduced infection...
levels. This indicates that these genes, which are up-regulated in response to EGL-38, confer a knockdown phenotype similar to that seen in egl-38 mutants. Neither set of RNAi-treated animals, however, was resistant to the same extent as egl-38(RNAi) worms, which had >75% of the population exhibiting reduced infection. These data suggest that the egl-38-dependent susceptibility to infection is mediated by the EGL-38 responsive genes gst-21, F11E6.8, and C05C9.1, but that the full egl-38 phenotype results from altered expression of more than one downstream gene.

3.4. DISCUSSION

3.4.1. Identification of EGL-38-responsive genes using expression microarrays

We report here a biological filtering approach for gene expression array data to identify target genes for the C. elegans Pax protein EGL-38. Using RNA collected from L4 larvae, we independently identified groups of transcripts altered in two hypomorphic egl-38 mutants and an inducible EGL-38 over-expressing strain compared to wild type. Overlapping transcripts from these experimental outputs defined distinct sets of candidate genes for further investigation. A robust group of genes reciprocally altered in egl-38(n578) and hsp-41::egl-38 array lists was chosen for validation studies. GFP reporter analysis revealed that almost all transcripts investigated are expressed in one or more tissues dependent upon egl-38 function. These findings confirmed the efficacy of using a genome-wide approach to discover genes responsive to EGL-38.
Although microarrays have become a standard gene discovery tool, they yield data that are inherently noisy. A major difficulty associated with microarrays and other genome-wide techniques is the relatively high recovery frequency of false positives (rev. in Hatfield et al., 2003; Armstrong and van de Wiel, 2004). Many statistical methods have been developed to reduce the inclusion of false positives during data analysis (rev. in Lai, 2006). However, as statistical filtering becomes more stringent, true positives tend to be filtered out along with the false positives. For this reason, microarray data is best applied within a biological framework.

In our experiments, we used very liberal statistical criteria with which to filter the normalized expression data. To alleviate concerns over false positives, we chose to filter one set of microarray data through another set of microarray data; one that had been produced under very different biological conditions. This approach allows us to minimize the recovery of false positives inherent to each individual data set. For instance, some differentially expressed transcripts in *egl-38(n578)* mutants (vs. wild type) are almost assuredly the result of the cumulative developmental defects present in the mutants. Ultimately, a measurable change in gene expression during the L4 stage may result from cellular or anatomical defects that reflect inappropriate signaling or morphogenesis during an earlier stage. The use of an inducible EGL-38 expression condition for comparison to the *egl-38(n578)* mutant condition mitigates the problem of confounding RNA changes resulting from indirect developmental defects with more direct effects. By inducing EGL-38 early in L4 and collecting the RNA soon thereafter, differential expression will not reflect any developmentally accrued abnormalities. Similar arguments can be made for the elimination of biologically spurious transcript
changes in the \textit{hsp-41::egl-38} strain by filtering them through a gene list obtained from analysis of RNA from \textit{egl-38} mutants. Therefore, we believe this approach has successfully addressed a fundamental problem of false positive recovery, and that the logic is supported by the robustness of our validation results.

Given the existence of multiple hypomorphic \textit{egl-38} mutations, we used two alleles with distinct tissue-preferential phenotypes to create a more informative set of microarray data. While \textit{egl-38}(n578) is known to preferentially disrupt development of the egg-laying system, \textit{egl-38}(sy294) mutants display strong defects in rectal development along with lesser egg-laying system phenotypes (Chamberlin et al., 1997). Therefore, we expected that creating cross-referenced gene lists between each of these and the \textit{hsp-41::egl-38} dataset would yield two functionally distinct sets of target genes. The unexpected finding that both \textit{egl-38} alleles affect the expression of immune response genes confounded this expectation, and these genes were present in both lists. Examination of rectal development in \textit{egl-38(sy294)/hsp-41::egl-38} shared gene knockdown experiments would help to illustrate the level of tissue-specificity achieved.

While we have found cross-referencing of array data from independent experiments to effectively filter out false positives, we recognize that expression profiling will always miss real targets (i.e. yield false negatives). Moreover, there is reason to believe that we have focused on an incomplete set of candidates within our data set. For example, the enrichment of genes annotated for immune function was not limited to the reciprocal category of genes present in the shared \textit{egl-38(n578)/hsp-41::egl-38} list. One can imagine scenarios in which real EGL-38 targets might be present in the loss- and gain-of-function gene lists, but not show reciprocal mRNA levels from one to the other.
In point of fact, validation of previous microarray studies performed in an analogous manner was not limited to reciprocally responsive transcripts (Fisher and Lithgow, 2006). In our case, EGL-38 likely acts as an activator in some tissues and a repressor in others for several target genes, as we observed with C05C9.1. Depending on the wild-type levels of a target gene transcript within each tissue, and the extent to which the altered EGL-38 activities in our study affect each tissue, such a target may show increased or decreased total mRNA levels in both conditions. We hypothesize that such complexities are occurring in vivo, and future studies of potential EGL-38 target genes need not rule out the non-reciprocally responding genes shared in our array output lists.

3.4.2. A novel immune function for egl-38 does not influence S. marcescens infection

Genome-wide expression analysis from C. elegans strains with various levels of EGL-38 activity allowed us to uncover a novel immune function for this Pax transcription factor. Specifically, these microarray data were enriched for immune response genes. egl-38-dependent expression of GFP reporters for most candidate genes tested supports the enrichment data, and EGL-38 was subsequently found to mediate infection by the gram-positive bacterial pathogen Microbacterium nematophilum (H. Jia, personal communication). In contrast, I found that EGL-38 does not appear to act in the resistance of C. elegans to the gram-negative Serratia marcescens. One potential problem with these data, however, is that the egl-38 mutants examined died prior to the expected timeframe of S. marcescens killing, even when cultured on the normal food source E. coli OP50. While we can firmly conclude that loss of egl-38 activity does not accelerate death caused by infection, these data do not fully assess whether egl-38
mutants are less susceptible to infection, as they are in the case of *M. nematophilum*. We hypothesize that EGL-38 does not serve a protective role for at least one gram-negative pathogen. However, further experiments probing EGL-38 function in other known *C. elegans* infections and/or by assays other than survival would help to determine whether this speculation is valid.

3.4.3. **At least three downstream genes mediate the function of EGL-38 in *M. nematophilum* infection**

Our processing of microarray data allowed us to identify three genes that play a role in the EGL-38-mediated susceptibility to *M. nematophilum*: *gst-21, C05C9.1*, and *F11E6.8*. We show that these genes are expressed in the rectum, the site of *M. nematophilum* colonization. Furthermore, their rectal expression is diminished in egl-38 mutants. *gst-21* and *C05C9.1* encode molecules predicted to participate in an immune response, and thus reducing gene function through RNAi might be expected to cause an enhanced sensitivity to bacterial infection. In contrast, we show that these genes, along with *F11E6.8*, facilitate infection in a fashion similar to *egl-38*, and RNAi depletion reduces sensitivity to bacterial infection. Mechanistic functions for each gene are considered below.

*gst-21* encodes a member of the glutathione-S-transferase (GST) protein family, which have been shown to function in metabolic detoxification. GST-21 presumably is not functioning in the detoxification of bacterial molecules, as in this case we would expect *gst-21(RNAi)* animals to show an enhanced infection phenotype, rather than reduced infection. However, in addition to metabolic detoxification, studies have shown
that GST proteins attenuate intracellular signaling, and can directly inhibit the MAPK Jun-N-terminal kinase (JNK; Villafania et al., 2000). Since previous results show that MPK-1-mediated signaling confers protection against *M. nematophilum* infection, it is possible that a GST-21 inhibition of MPK-1 is responsible for the *gst-21(RNAi)* reduced infection phenotype we observe.

*C05C9.1* is a putative LBP/BPI/CETP protein; a category of protein involved in lipid binding, and in some cases, bacterial killing. We hypothesize that the reason *C05C9.1(RNAi)* animals did not exhibit an enhanced infection phenotype is that this type of protein acts as a pattern recognition molecule for the LPS of gram-negative bacteria (rev. in Beamer et al., 1999). Gram-positive bacteria, including *M. nematophilum*, have no LPS layer surrounding the peptidoglycan cell wall, so *C05C9.1* is not likely to act as a defense protein for this bacterial pathogen. Why *C05C9.1(RNAi)* animals exhibit reduced infection is not clear. Although these proteins are generally found in soluble form, membrane forms of BPI and LBP proteins have been described (Weersink et al, 1993; Muller et al., 2003), suggesting that a membrane-bound form of C05C9.1 could serve as an attachment point in the rectum for invading bacteria.

*F11E6.8* encodes a predicted membrane-associated tyrosine kinase of unknown biological function. Interestingly, the F11E6.8 protein is 48% identical (64% similar) to the human macrophage stimulating 1 receptor (MST1R). Macrophages are phagocytic white blood cells that act in both innate and adaptive immunity. Activation of this receptor stimulates macrophage migration and phagocytosis, as well as inhibiting LPS-induced inflammation (Wang et al., 2002). Moreover, human MST1R is closely related to the *met* proto-oncogene, and it mediates jaagsiekte sheep retrovirus (JSRV) induction.
of lung cancer (Miller et al., 2004). Neither of these activities is suggestive of F11E6.8
function as a facilitator of *M. nematophilum* infection in *C. elegans*. However, the role of
a similar protein in the human innate immune response does indicate that this protein may
be involved in pathogen-induced signaling. This is a particularly exciting prospect, as no
receptors identified in genetic screens or expression arrays have been shown to act in *M.
nematophilum* infection (Gravato-Nobre et al., 2005; O’Rourke et al., 2006).

Overall, our data suggest a model wherein EGL-38 regulates at least three genes
whose functions allow for increased bacterial infection. More work must be done to
clarify the precise mechanism(s) by which rectal infection is enhanced via the function of
these host genes. However, one attractive explanation is that *M. nematophilum* has
evolved the ability subvert the anti-microbial nature of these genes, and even exploit
them. Such adaptations are common among pathogens as they battle to evade and
overcome host immune responses (rev. in Bhavsar et al., 2007).

3.4.4. Implications for innate immunity in vertebrates

Do Pax2/5/8 proteins play a conserved role in the innate immune response? This
is an important unanswered question. It has long been known that *Pax5* is required for B
lymphocyte commitment and maintenance of B cell identity (rev. in Singh et al., 2007).
However, this reflects a developmental function rather than an immune-specific activity.
Pax2/5/8 proteins are also known to act as anti-apoptotic factors in both *C. elegans* and
vertebrates (Park et al., 2006; rev. in Lang et al., 2007), and programmed cell death is
important tool of the innate immune response for many host species responding to
infectious pathogens (rev. Hoole and Williams, 2004). Although *M. nematophilum* does
not induce apoptosis of rectal cells in *C. elegans*, apoptosis of germline nuclei is increased in response to *Salmonella typhimurium* infection (Aballay and Ausubel, 2001), and *egl-38* plays a role in developmentally regulated germline apoptosis (Park et al., 2006). This suggests a mechanism that might confer an evolutionary advantage for conserving an immunological role for *Pax2/5/8* genes. Further experiments can determine whether a role in the innate response to infection is a conserved function for *Pax2/5/8* proteins.

Irrespective of whether *Pax* genes themselves act in the innate immune response in humans, the role of *EGL-38* in *C. elegans* infection can be instructive regarding transcriptional regulation during the human innate immune response. For example, *EGL-38* plays a role in the development of rectal cells, as well as their mature immune function. Likewise, STAT proteins are transcriptional regulators involved in the development and differentiated immune response of dendritic cells (rev. in Onai and Manz, 2008), which are critical effectors of the innate immune system. How are cell fate/maintenance target genes and immune response target genes regulated by these transcription factors? Are they controlled simultaneously or are they subject to different regulation? Given the specificity of *egl-38* alleles discussed here, as well as other described alleles (Zhang et al., 2005), one could begin to tackle these questions in a relatively simple system. In addition, further examination of the *EGL-38* response to *C. elegans* rectal infection may help to address other questions, such as the role of transcription factor cross-talk and interaction/recruitment of cofactors during the immune response.
Figure 3.1. Microarray design and differentially expressed transcript groupings
RNA was collected in triplicate from synchronized L4 populations. Three genetic backgrounds with altered EGL-38 activity (\textit{hsp-41:egl-38}; \textit{egl-38(n578)}; \textit{egl-38(sy294)}) were individually compared to wild type (N2) populations. Normalization (RMA) and statistical filtering (|fold-change| ≥1.5, \( p = 0.05 \)) generated individual lists of differentially expressed genes. Overlap of these lists was used to classify potential tissue-specific \textit{egl-38}-responsive candidate genes (\(^*\) = egg-laying system, \(^*\) = rectum, ? = tissues affected by both hypomorphic mutations). The numbers listed in overlapping sections indicate all genes from the datasets, ignoring the directionality in which transcript abundance is altered, which was used as a subsequent filter.
<table>
<thead>
<tr>
<th>Transcript</th>
<th>Gene Description</th>
<th>Fold Change</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H19N07.3</td>
<td>-</td>
<td>-3.4</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>C41H7.5</td>
<td>-</td>
<td>-1.7</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>C49G7.4</td>
<td>phat-3</td>
<td>-1.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>C41H7.3</td>
<td>-</td>
<td>-1.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>ZK697.6</td>
<td>gst-21</td>
<td>-1.8</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>F41D3.3</td>
<td>nhr-265</td>
<td>-1.7</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>F11E6.8</td>
<td>PTK</td>
<td>1.7</td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>F08F3.4</td>
<td>NAD-dependent epimerase</td>
<td>1.6</td>
<td>-1.6</td>
<td></td>
</tr>
<tr>
<td>Y42A5A.2</td>
<td>lact-8</td>
<td>2.2</td>
<td>-1.7</td>
<td></td>
</tr>
<tr>
<td>T21H3.2</td>
<td>ptr-16 family</td>
<td>3.2</td>
<td>-1.6</td>
<td></td>
</tr>
<tr>
<td>C05C9.1</td>
<td>BPI family</td>
<td>1.7</td>
<td>-1.7</td>
<td></td>
</tr>
<tr>
<td>H13N06.2</td>
<td>von Willebrand factor</td>
<td>1.9</td>
<td>-2.6</td>
<td></td>
</tr>
<tr>
<td>F32D8.7</td>
<td>-</td>
<td>2.1</td>
<td>-1.7</td>
<td></td>
</tr>
<tr>
<td>F18E3.11</td>
<td>-</td>
<td>4.1</td>
<td>-3.5</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1. Genes showing reciprocal response to egl-38(n578) and hsp-41::egl-38 activity

* N2 strain was used as the wild-type control.

Down-regulated conditions are highlighted (gray).
Figure 3.2. Multiple candidate genes are expressed in the vulva and/or rectum
Nomarski and fluorescent images of the L4 vulva and rectum of transgenic animals. Gene names indicate regulatory sequence driving GFP expression. Left- and right-handed columns for each panel are vulval and rectal images, respectively. Top images within panels are light micrographs, bottom images are GFP fluorescence photographs. Vulval cells are bracketed. The rectal lumen is marked by a dotted line.
<table>
<thead>
<tr>
<th>Gene</th>
<th>vulva</th>
<th>uterus</th>
<th>system</th>
<th>rectum</th>
<th>tissues</th>
</tr>
</thead>
<tbody>
<tr>
<td>H19N07op</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>seam, pharynx, Pn.p's</td>
</tr>
<tr>
<td>phat-3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>head neurons</td>
</tr>
<tr>
<td>C05C9.1</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>pharynx, spermatheca</td>
</tr>
<tr>
<td>gst-21</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>head neurons</td>
</tr>
<tr>
<td>F18E3.11</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>hypodermis, Pn.p's</td>
</tr>
<tr>
<td>H13N06.2</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>pharynx, hypodermis, Pn.p's</td>
</tr>
</tbody>
</table>

* lact-8 - - + + b.w.m.*

<table>
<thead>
<tr>
<th>Gene</th>
<th>vulva</th>
<th>uterus</th>
<th>system</th>
<th>rectum</th>
<th>tissues</th>
</tr>
</thead>
<tbody>
<tr>
<td>F11E6.8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>hypodermis</td>
</tr>
<tr>
<td>F08F3.4</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>hypodermis, Pn.p's</td>
</tr>
<tr>
<td>F41D3.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>seam</td>
</tr>
</tbody>
</table>

**Table 3.2. Expression pattern summary for candidate EGL-38 target genes**

(+) presence of one or more cells in the tissue displaying significant expression.

(*) b.w.m. = body wall muscle.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Genotype</th>
<th>Predicted regulation*</th>
<th>n</th>
<th>% Vulva#</th>
<th>% Rectum#</th>
</tr>
</thead>
<tbody>
<tr>
<td>H19N070p</td>
<td>wt</td>
<td>+</td>
<td>37</td>
<td>86.5</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>38</td>
<td>52.6</td>
<td>94.7</td>
</tr>
<tr>
<td>phat-3</td>
<td>wt</td>
<td>+</td>
<td>38</td>
<td>0.0</td>
<td>36.8</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>39</td>
<td>0.0</td>
<td>17.9</td>
</tr>
<tr>
<td>F41D3.3</td>
<td>wt</td>
<td>+</td>
<td>35</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>+</td>
<td>35</td>
<td>0</td>
<td>5.7</td>
</tr>
<tr>
<td>gst-21</td>
<td>wt</td>
<td>+</td>
<td>34</td>
<td>76.5</td>
<td>73.5</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>36</td>
<td>55.6</td>
<td>47.2</td>
</tr>
<tr>
<td>F18E3.11</td>
<td>wt</td>
<td>-</td>
<td>37</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>38</td>
<td>2.6</td>
<td>0.0</td>
</tr>
<tr>
<td>H13N06.2</td>
<td>wt</td>
<td>-</td>
<td>36</td>
<td>41.7</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>36</td>
<td>47.2</td>
<td>97.2</td>
</tr>
<tr>
<td>lact-8</td>
<td>wt</td>
<td>-</td>
<td>36</td>
<td>5.6</td>
<td>86.1</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>38</td>
<td>5.3</td>
<td>71.1</td>
</tr>
<tr>
<td>F11E6.8</td>
<td>wt</td>
<td>-</td>
<td>28</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>29</td>
<td>100.0</td>
<td>93.1</td>
</tr>
<tr>
<td>F08F3.4</td>
<td>wt</td>
<td>-</td>
<td>15</td>
<td>40.0</td>
<td>46.7</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>9</td>
<td>66.7</td>
<td>33.3</td>
</tr>
<tr>
<td>C05C9.1</td>
<td>wt</td>
<td>-</td>
<td>35</td>
<td>11.4</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>egl-38(n578,</td>
<td>-</td>
<td>36</td>
<td>36.1</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Table 3.3. Enrichment of egl-38-dependent expression among candidate genes (wt) wild type. (*) Expected relationship between EGL-38 and candidate gene, as predicted by microarray results (+ = activation; - = repression). (#) % animals with expression in 1 or more cell of the tissue.
<table>
<thead>
<tr>
<th>Gene Family</th>
<th>Array comparisons and associated fold enrichment</th>
<th>Family Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>egl-38(n578) vs. N2</td>
<td>F.E.</td>
</tr>
<tr>
<td>gaLECtin</td>
<td>lec-7 vs. N2</td>
<td>2.2</td>
</tr>
<tr>
<td>LYSOzyme</td>
<td>lys-2, lys-7 vs. N2</td>
<td>5</td>
</tr>
<tr>
<td>Glutathione S-Transferase</td>
<td>gst-21, gst-38 vs. N2</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saposin-like Protein</td>
<td>spp-1, spp-7, spp-15</td>
<td>3.8</td>
</tr>
<tr>
<td>Downstream Of DAF-16</td>
<td>dod-19, dod-22, dod-24</td>
<td>7.6</td>
</tr>
<tr>
<td>PHaryngeal gland</td>
<td>phat-1, phat-2,</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>phat-3, phat-4</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.4. Overrepresentation of immune gene classes in altered EGL-38 activity array data

(*) Both strains treated with heat-shock.
F.E. = fold enrichment compared to genome.
Figure 3.3. *C. elegans* death from reduced *egl-38* activity precedes killing by *S. marcescens*

(A-C) Survival curves for various populations grown on non-pathogenic *E. coli* strain OP50 and pathogenic *S. marcescens* strain DB11. In all conditions, approximately 50 L4 animals were plated on bacterial lawns and scored daily for survival. Curves represent average of two independent trials. (A) wild type (N2); (B) *egl-38*(sy294); (C) RNAi experiments, including mock dsRNA injection control. (D) Light micrograph of the rectal defects associated specifically with dying *egl-38* mutant animals (*egl-38*(RNAi) pictured).
Figure 3.4. Candidate gene expression patterns in the rectum
Schematics of rectal cells. (A-H) Degree of shading reflects the percentage of animals that exhibit transgene expression in each cell type (black = 100%, white = 0%). Transgenes represented are (A) H19N07op; (B) F11E6.8; (C) H13N06.2; (D) gst-21; (E) lact-8; (F) phat-3; (G) F08F3.4; (H) C05C9.1. (I) Summary of all transgene expression patterns. Degree of shading reflects the proportion of candidate gene reporters that display significant expression in that cell type.
Figure 3.5. Rectal expression of many candidate genes is dependent on egl-38
Rectal expression of seven candidate genes is reduced in an egl-38(n578) background. There are a total of eleven rectal cells (9 cell types), however no gene was expressed in every cell type. All expression was scored in L4 animals.
Figure 3.6. RNAi knockdown reveals genes that function in *M. nematophilum* infection
Altered infection graph illustrating the percent of animals in a given treatment condition that deviate from a moderate/baseline level of rectal infection by *M. nematophilum*. Increased SYTO13 fluorescence reflects enhanced bacterial colonization, while decreased fluorescence suggests reduced bacterial colonization. Fluorescence was scored manually in four qualitative categories (“-“ and “+” were grouped as reduced infection, “++” was considered moderate/baseline infection, and “+++” was considered enhanced infection). All RNAi populations are L4 progeny from dsRNA-injected mothers. n ≥ 21.
CHAPTER 4

ISOLATION OF A NOVEL ALLELE OF LIN-14 AND CHARACTERIZATION OF THE RESULTING HETEROCHRONIC PHENOTYPES IN THE C. ELEGANS EGG-LAYING SYSTEM

4.1. INTRODUCTION

During development, the temporal progression of pattern formation is of equal importance to spatial orientation of cells and tissues. Premature formation of adult tissues can lead to disease, as in the skin conditions porokeratosis and Darier’s disease, which are characterized by dysregulated differentiation of keratinocytes (Shen et al., 2002; Kassar et al., 2008). Similarly, failure of tissues to complete their temporal developmental program can be deleterious. For instance, cancer stem cells arising from differentiated tissue exhibit hallmarks of immature cells, suggesting that they have reversed the normal temporal differentiation process (rev. in Gao, 2008). While there is a considerable amount of diversity in the mechanisms organisms use to temporally regulate gene expression, many features are conserved. On the smallest scale of conserved temporal control is the cell cycle, which employs the same families of molecules to direct cell growth and division in virtually all eukaryotes (rev. in Murray and Kirschner, 1989).
At the other extreme, gene ontology (GO) category groupings of global gene expression profiles throughout development have been shown to be highly similar between mammals and *Drosophila* (Wagner et al., 2005).

Organogenesis requires that cells and tissues be organized properly for continuous and coordinated interaction. In many cases, tissues are found to progress through stepwise developmental specification events, with each step requiring that the previous step(s) occurred normally. For instance, the generation of hematopoietic progenitors from embryonic stem cells requires sequential exposure to Bmp4, activin A, bFGF, and VEGF. Simultaneous exposure to these signals, however, is less effective in promoting cell commitment to a hematopoietic lineage (Pearson et al., 2008). Crosstalk between tissues in the kidney (rev. in Davies, 2002) and cooperative phases between retinal and lens tissue during eye development (Jean et al., 1998) provide further examples of sequential temporal regulation. In total, these studies illustrate that sequential specification and morphogenetic events require coordination between two synchronously developing tissues.

In contrast to sequential temporal regulation, other types of temporal regulation are more modular. Reiterative formation of somites in vertebrate embryos relies on a temporal segmentation clock, and disruptions in this clock lead to skeletal defects. However, experimental transplants of presomitic tissue fragments introduced in a reversed orientation maintain the original directionality and rate of somite formation, demonstrating that this temporal patterning occurs by the reiterative formation of autonomous units (rev. in Shifley and Cole, 2007). Additionally, some adult amphibian species can regenerate limbs, illustrating that tissue development can occur in a modular
fashion, independent from normal temporal development (rev. in Yokoyama, 2008). Another form of modular temporal control is the circadian clock. The *Drosophila period* gene encodes a core component of the circadian clock in flies and vertebrates (rev. in Hardin, 2005). In the invertebrate *C. elegans*, the *period* ortholog *lin-42* is not involved in circadian control, but does play a role in the heterochronic—or developmental timing—pathway (Jeon et al., 1999; Tennessen et al., 2006). The heterochronic pathway controls the modular execution and sequence of post-embryonic larval stages of *C. elegans* development (rev. in Moss, 2007). From these examples, one can see that temporal regulation can take on different apparent forms.

Although genes in many organisms exhibit both sequential and modular temporal developmental regulation, it is unclear whether or how the different processes are related. Are these processes under the control of shared temporal regulators, or do they use very different mechanisms to direct development? Many genes have been shown to regulate the temporal behavior of tissues across phylogenetically diverse species, such as the *lin-42/period* gene family mentioned above. Another gene in the *C. elegans* heterochronic pathway, *lin-28*, is also conserved among metazoans, and appears to be developmentally regulated (Moss and Tang, 2003). While *lin-28* expression delineates a stepwise progression of maturing epithelia in vertebrates (Yang and Moss, 2003), it plays a role in the modular larval stage development of *C. elegans* (Ambros and Horvitz, 1984). Thus, by studying the individual genes and gene interactions known to play a role in these processes (e.g. *lin-28*), we can better understand how cells receive, interpret, and execute temporal instructions.
In addition to *lin-42* and *lin-28*, genetic experimentation in *C. elegans* has identified other factors acting in the heterochronic pathway. These genes act to specify temporal, rather than spatial or lineal identity of cells. In *C. elegans*, post-embryonic development occurs in four larval stages, each punctuated by a molting of larval cuticle. Post-embryonically dividing cell types undergo stage-specific developmental patterns of cell division and/or differentiation, and these stage-specific patterns are under the control of the heterochronic pathway. Transcription factors such as LIN-14 and HBL-1 are important for the specification of larval-specific fates, as are members of the *lin-4* and *let-7* microRNA families, which act to downregulate the translation of these transcription factors and other heterchronic gene products (reviewed in Moss, 2007). Apart from the genes in the heterochronic pathway, perturbations in other genes have been found to display developmental timing defects. For example, *cog-3* alters the progression of gonadogenesis, and the resulting asynchrony between gonadal and vulval development leads to a defect in the mature egg-laying system (Huang and Hanna-Rose, 2006).

*lin-14* is a heterochronic gene that encodes a novel type of transcription factor. Loss of *lin-14* activity is associated with the elimination of L1-specific developmental patterns and premature, or precocious, execution of L2-specific patterns. Moreover, *lin-14* gain-of-function (*gf*) mutations result in reiteration of L1-specific developmental programs (Ambros and Horvitz, 1984, 1987). In combination, these data suggest that *lin-14* is necessary and sufficient for promoting L1-specific cell fates in *C. elegans*. *lin-14* was also recognized as the first target of a microRNA (Wightman et al., 1991; Lee et al., 1993). As with all microRNA’s the *lin-4* gene encodes a short RNA with complementary sequence to the 3’ untranslated region of its target mRNA(s) (Lee et al., 1993), and the
interaction between lin-4 and the lin-14 3’ UTR inhibits lin-14 translation. It is this regulation that is disrupted in lin-14(gf) mutations, while all known lin-14 loss-of-function (lfl) mutations are found within coding sequences (Reinhart and Ruvkun, 2001). lin-14 expression is detectable during late embryogenesis, and continues through the L1 stage of post-embryonic development. Subsequently, LIN-14 expression is repressed as lin-4 levels rise at the end of L1, allowing other heterochronic genes to promote the switch to L2-specific fates (Lee et al., 1993). LIN-14 levels remain at undetectable levels throughout the remainder of larval development (Ruvkun and Giusto, 1989).

The work in this chapter focuses on the temporal development of the C. elegans egg-laying system. The egg-laying system in C. elegans is a relatively simple organ composed of two main tissue types, the uterus and the vulva. During normal development, proper connection between the vulva and gonad requires cells of the ventral uterine (VU) lineages. Explicitly, six intermediate precursors called π cells are specified by the action of Notch signaling and multiple transcription factors, including one encoded by the heterochronic gene lin-29 (Newman et al., 1995, 2000; Cinar et al., 2003). These π cells then go on to produce two cell types, the uterine seam (utse) and the uv1 cells, both of which are important for the uterine-vulval connection. One defect in cog-3(ku212) mutants is that uv1 cells are not fully specified, and instead undergo necrotic cell death (Huang and Hanna-Rose, 2006). It is thought that the delay in gonad development indirectly results in disruption of epidermal growth factor (EGF) signaling from the vulva to the π cells. In the absence of EGF receptor activation, the pre-uv1 cells not only fail to differentiate normally, but they die due to the absence of an EGF-dependent survival pathway. The cog-3 data demonstrate how subtle asynchrony
between coordinating tissues during organogenesis can result in dramatic consequences, but the mechanisms by which synchrony is tightly controlled are still not well understood.

In this chapter, we describe the characterization of a new lin-14 allele, and uncover a role for this heterochronic gene in synchronizing the vulva and uterus during maturation of the egg-laying system. The mutation lin-14(sa485) was isolated on the basis of abnormal male tail morphology. We found that the male defect in lin-14(sa485) mutants results in part from inappropriate patterning of an asymmetric cell division occurring just before the L1-to-L2 transition (see Appendix C). Closer observation of this strain revealed that the hermaphrodite vulva fails to evert in a small proportion of animals. Through further investigation, we found that eversion does not fail, but exhibits a delay relative to the maturation of the germline and gonad. This delay is brief, and represents a subtle alteration of vulval-gonadal synchrony, similar to what is seen in cog-3(ku212) mutants. Unlike cog-3(ku212), however, we found lin-14(sa485) animals to exhibit precocious vulval development during L4, which alters coordination between the uterus and gonad, and subsequently causes a delay in vulval maturation. Finally, as with cog-3(ku212) worms, lin-14(sa485) mutants have abnormal uv1 fate specification. Rather than dying by necrosis, however, lin-14(sa485) uv1 cells show only a delayed differentiation phenotype. Together, this work defines a temporal control function of lin-14 that is carried over into late larval development. In addition, the work demonstrates that a single regulatory gene can contribute to both modular and sequential developmental timing processes.
4.2. MATERIALS AND METHODS

4.2.1. Genetic strains and strain construction

*C. elegans* strains were cultured under standard conditions (Brenner, 1974). All experiments were done at 20°C, unless otherwise noted. Either the Bristol isolate N2, or in most cases the strain DR466 [*him-5(e1490)*], were used as wild type in all crosses and experiments. The following mutations, rearrangements, and transgenes were used:

- Linkage Group I (LGI): *dpy-5(e61)*; LGII: *rol-6(e187)*, *inls179*, *tra-2(q276)*; LGIII: * unc-32(e189)*; LGIV: *unc-5(e53)*, *him-8(e1489)*; LGV: *dpy-11(e224)*, *him-5(e1490)*,
  - *mnDp1(V;X)*; LGX: *unc-78(e1217)*, *lin-18(e620)*, *lon-2(e678)*, *dpy-6(e14)*, *lin-14(sa485)*, *lin-14(n179)*, *lin-2(e1309)*, *unc-9(e101)*, *let-4(mn105)*, *unc-3(e151)*, *let-41(mn146)*. All strain information can be found in Appendix B. Additional genetic information is available at [http://www.wormbase.org](http://www.wormbase.org).

For most anatomical experiments involving *lin-14(sa485)*, a derivative of strain JT8174 [*sa485; him-5(e1490)*] was used—CM1500—which had been backcrossed four times. JT8174 was used directly for mapping experiments.

4.2.2. *lin-14(sa485) mapping and cosmid rescue*

Strain JT8174 was isolated in a male abnormal (Mab) mutagenesis screen (H.M. Chamberlin and J.H. Thomas, unpublished communication). Linkage analysis was performed using the marker strains MT464 [*unc-5(e53)*; *dpy-11(e224)*; *lon-2(e678)*] and MT3751 [*dpy-5(e61)*; *rol-6(e187)*; *unc-32(e189)*]. DR466 males were mated to the marker strain hermaphrodites, male F1 cross-progeny were crossed to JT8174
hermaphrodites, and resulting F2 hermaphrodite progeny were self-crossed. Resulting F3 hermaphrodites with a generally wild-type appearance were plated individually, and F4 progeny were used to confirm sa485 homozygosity. The offspring of verified lin-14(sa485)/lin-14(sa485) F3 mothers were scored for segregation of linkage group markers.

Three-point genetic mapping of sa485 was performed in a similar fashion for all crosses. By way of example, wild-type males were mated to the strain SP928 [dpy-6(e14) unc-9(e101)]. Resulting male progeny were then crossed with JT8174 to generate dpy-6(e14) unc-9(e101)/sa485 F2 hermaphrodites. These F2 animals were self-crossed, Dpy nonUnc and Unc nonDpy progeny were selected and plated individually, and these plates were assessed for segregation of sa485 males with defective spicule production. This method was followed using the following marker strains: SP928, MT1790 [unc-78(e1217) lin-18(e620) lon-2(e678)], SP224 [mnDp1 (X; V); let-4(mn105) unc-3(e151)], SP392 [mnDp1 (X; V); unc-3(e151) let-41(mn146)], CB1309 [him-5(e1467); lin-2(e1309)], CB4856 [wild type “Hawaiian (Hi)” strain; Tc1 pattern IX].

Snip-SNP mapping was carried out in two ways. Initially, broad positioning of sa485 on X was done using a two-factor approach. Here, males from the Hawaiian (Hi) polymorphic strain CB4856 were mated directly to JT8174, F1 cross-progeny were picked, and sa485 homozygotes were selected in the F2 on the basis of the Mab and/or Uta phenotypes. Three polymorphisms were tested in the recovered sa485 lines. The second snip-SNP mapping strategy was similar to the three-point genetic crosses. In these cases, however, the double mutant strain (CM1455) carried sa485 and the marker mutation dpy-6(e14), and males from the polymorphic strain CB4856 were crossed to this
strain. F1 cross-progeny were selected and allowed to self-cross, and single mutant F2 recombinants (Dpy nonMab/Uta, Mab/Uta nonDpy) were recovered. The recombinant chromosome was homozygosed prior to SNP testing. For the two-point snip-SNP mapping strategy, whole worm lysates of a population of animals were made as described (Barstead, 1999). A similar protocol was used for the three-point crosses, except that the entire lysate reaction was scaled up to provide a long-term template DNA source. Amplification of genomic DNA fragments containing the snip-SNP’s of interest were carried out using a nested PCR approach. Briefly, one round of PCR is run using the genomic DNA in worm lysates as a template source. A second round of PCR is then run, using primers nested between the primer sequences used in the first reaction and the first round reaction product as the DNA template. The final PCR products are then subjected to cutting by the restriction enzymes appropriate for the genotype at the polymorphism of interest. The following polymorphisms and their associated primers were used: pkP6010—PR 367/368; pkP6110—PR 369/370; pkP6139—PR 371/372; pkP6114—PR 1515/1516; pkP6126—PR 1829/1830. Sequences for these can be found in Appendix A.

Cosmids were provided by Audrey Fraser (Wellcome Trust Sanger Center Institute), and purified according to Midi prep protocol (Qiagen). We verified all cosmids by restriction enzyme digestion, and found only one (C03G4) to be incorrect. Four pools of three cosmids each were prepared, providing almost continuous coverage of the X chromosome genetic region between 2.86-3.86. From these pools, two types of injection mixtures were created, yielding final cosmid concentrations in the range of 24-28 ng/μl each, as well as either 150 ng/μl of the rol-6(su1006) plasmid pTJ43 or the myo-2::gfp vector pTJ871. These solutions were injected into wild-type animals and
subsequently crossed into the *sa485* strain JT8174 to assess rescue of the male spicule
defect.

**4.2.3. lin-14 complementation and sequencing**

The complementation assay was performed against the temperature sensitive *lin-14(n179)* allele. Crosses and scoring were performed at 20°C, rather than at a strict
permissive or restrictive temperature. This was done because the penetrance and severity
of the *sa485* male spicule phenotype had only been characterized at 20°C, and we were
able to observe spicule defects in *lin-14(n179)* mutants at this temperature. Given that
*lin-14* is located on the X chromosome and males are normally hemizygous for X, we
utilized the *tra-2(q276)* mutation to generate males carrying two copies of X. Animals
homozygous for this mutation are always phenotypically male. Specifically, *tra-2(q276)*
males were crossed with the *lin-14(n179)* strain DR441. Resulting hermaphrodites were
then self-crossed to generate *tra-2(q276)* males carrying one copy of *lin-14(n179)*. As
not all of these males were heterozygous for the *lin-14(n179)* mutation, we plated these
males individually with 4-6 JT8174 hermaphrodites for mating. Successful mating was
confirmed by the appearance of wild-type cross-progeny. F1 hermaphrodites were
chosen as L4’s, and allowed to self-cross. The appearance of Pvl in F2 self-cross
progeny from these hermaphrodites verified the presence of *lin-14(n179)*, and sibling F2
males were scored for their spicule phenotype.

To sequence the coding region of *lin-14* in JT8174, we amplified genomic
sequences (including introns) as 6 PCR fragments. Exons 4-6 were amplified using
primers PR1895-1898, exon 7 using PR1899-1902, exons 8-12 using PR1903-1906, exon
13 using PR1907-1910, and the 3'UTR using PR1911-1914. Sequences can be found in Appendix A. PCR products were column-purified (Qiagen) and sequenced using a 3730 DNA analyzer (Applied Biosystems) with the nested primers. Sequence outputs were verified by chromatogram examination, and alignment was done using web-based ClustalW 1.8 software (http://searchlauncher bcm tmc.edu/multi-align/multi-align.html). The sa485 gGt-to-gAt nucleotide substitution was confirmed in sequencing reads from both ends of the PCR product covering exon 10. The full sequence of lin-14 exon 10 in sa485 mutants is as follows:

cattccgaccagtgtcagaacaacaattatctaacagatcattcaaggcaaaagtatgaagagatggaagttggaatgtcgatg
attcagttttgtaagaagttgacagaaagctgtgctttggaccacgtctatgtctacaaaccacagtgtgcgtAtctcaaccactc
gaactatgcaaatctttcaatcaaggaatgtatatattcaac.

4.2.4. lin-14(sa485) hermaphrodite anatomy and microscopy

Categorization of vulval development stages was qualitative, based upon gross morphology. Early milestones were a count of the number of vulval precursor cell (VPC) progeny nuclei present, and these counts included cells derived from un-induced VPC’s (P3.p, P4.p, and P8.p). After the first two rounds of VPC cell division were completed, vulval development was partitioned into five morphological milestones: initial inverting, the “pyramid” stage (which we defined as the arrangement of vulval cells in the central plane of the worm in the shape of a pyramid, with the gonad anchor cell (AC) at the apex), the well-defined “Christmas tree” stage, evverting, and completed eversion. These subjectively labeled stages were not equally spaced in developmental time, so we further characterized animals as being between any two milestones to increase the temporal
resolution of the data. Similarly, we developed a milestone-labeling system for tracking
gonadal progression through the larval development and into adulthood. The first was
the clear distinction of the AC from other somatic gonad tissue, followed by dorsal
turning of the distal tip cell (DTC), DTC return (it had reflexed and migrated to a position
immediately dorsal to the vulva), sperm differentiation, oocyte differentiation, one
fertilized embryo per gonad arm (1-2 eggs), and two fertilized embryos per gonad arm (3-
4 eggs). Like vulval development, these milestones do not provide a temporally uniform
or complete series of events, so less strict categorizations of gonadal migration were
noted to aid in following progress through the larval stages. Not all milestones were
applicable to each experiment, but experiments with overlapping developmental
timeframes shared the same system of milestone labeling.

The Evl-D phenotype was initially observed using a dissecting stereoscope
(Leica). All subsequent analysis was performed using a Zeiss Axioskop 2 (Zeiss). For
studies of the Evl-D phenotype and vulva-gonad maturation, animals of the appropriate
general stage were selected from mixed-stage populations and scored for their precise
developmental progression. Of the 140+ animals scored for each genotype in the
experiment defining the Evl-D phenotype, 35-36 animals were scored in each of the four
gonadal maturation stages. For the vulva-gonad synchrony/time-course experiment,
synchronous worm populations were generated, and samples from these populations were
taken for scoring. The time-course data do not represent one synchronous population, but
rather reflect combined sets of data coming from several independent trials.
Synchronization was performed as described (http://aropianlab.ucsd.edu/protocols/hatch-
off.htm), so time-course hours indicate hours after transferring arrested L1 larvae to
plates. The number of animals scored in each stage was not a parameter of the experimental design, and therefore there was significant variability. However, we typically scored 8-10 animals within each hour-long period.

4.2.5. *ida-1::gfp* studies

The strain BL5717 [*inls179; him-8(e1489)*] was used as the control for uv1 specification experiments. We believe this to be a suitable control strain, because like *him-5(e1490),* *him-8(e1489)* causes a high incidence of males, and neither mutation is expected to influence the timing of uv1 cell specification. The uv1 cells were scored on a per animal side basis, so although there are four uv1 cells specified per animal, we only scored up to two GFP(+) cells/worm. While all four cells were not scored formally, we did take note of the other pair of cells, and found the expression from each animal side to match in most animals. Expression of *ida-1::gfp* in the HSN motor neuron acted as the internal control verifying transgene presence in all animals scored. In total, 47 wild type and 57 *lin-14(sa485)* animals were examined in this assay. These animals were not evenly spread throughout the temporal categories, potentially skewing the precision of the proportions shown, but not likely altering the delay pattern observed.

4.2.6. *lin-14(n179)* temperature shift experiments

Vulval eversion was assessed in the temperature-sensitive strain DR441 [*lin-14(n179)*]. Animals were synchronized as before, except embryos were hatched overnight at 15°C or 25°C. L1 populations were then shifted to the opposite temperature condition at the following hours post-L1 feeding: 0, 6, 12, 18, and 24. Due to the slow
growth rate of worms at 15°C, upshifted animals also included groups shifted at 30, 26,
and 42 hours. Vulval stage/phenotype data were collected for animals throughout the
four stages of germline maturation described above. In both representations of the data,
normal vulval eversion includes animals that had completed eversion and were not Pvl.
Over 50 animals were scored for each temperature shift condition. Within individual
temperature shift conditions, the number of animals scored for each germline maturation
stage varied, but most condition-stage combinations had n values between 14 and 18. For
Evl-D analysis, percentage data are reported directly. In the time-of-action graph, only
temperature shift conditions representative of shifts at rough developmental intervals
were used, and all adults were grouped together (germline maturation stages combined).
The chosen times were as follows: 15°C → 25°C upshift: none—15°C only, feeding—up
at 0 hrs., midL1—up at 12 hrs., L1/L2—up at 24 hrs., mid-L2—up at 30 hrs., L2-L3—up
at 42 hrs.; 25°C → 15°C downshift: none—25°C only, feeding—down at 0 hrs., midL1—
down at 6 hrs., L1/L2—down at 12 hrs., mid-L2—down at 18 hrs., L2-L3—down at 24
hrs.

4.3. RESULTS

4.3.1. Identification of a novel allele of the heterochronic gene lin-14

In the *C. elegans* male, four rectal cells divide post-embryonically, producing
sensory mating structures (Sulston et al., 1980). One of these cells, termed B, divides
asymmetrically to produce one sensory organ precursor and one epithelial cell. The
sensory organ precursor, B.a, goes on to generate all of the cells of the copulatory
spicules, which aid in vulval location and sperm transfer during mating (Liu and Sternberg, 1995). Many genetic screens have been performed to isolate genes involved in spicule development. For instance, genetic screens have identified components of Wnt, EGF, and TGF- signaling pathways important for the specification and development of B.a and its progeny (Wu and Herman, 2006; reviewed in Patterson and Padgett, 2000; Sternberg et al., 1995).

The sa485 mutant strain JT8174 was isolated from one such spicule-defective mutant screen. These animals exhibit variable defects in the asymmetric division of B, as well as abnormal patterning of B cell progeny, ultimately resulting in reduced or absent spicules. To identify the gene perturbed by the sa485 mutation, we performed genetic linkage analysis using the male spicule phenotype to track sa485 in crosses. We first attempted to construct double mutants between sa485 and marker mutations associated with each chromosome. The mutual exclusion of spicule-defective and long (Lon – caused by lon-2(e678)) phenotypes indicated that sa485 was likely to lie on the X chromosome (Table 4.1.).

Linkage experiments were followed by 2- and 3-point mapping crosses. The loci involved are illustrated in Figure 4.1. and the relevant data are presented in Tables 4.2. and 4.3. Using a chromosome marked with unc-78(e1217) and lon-2(e678) chromosome, we found that 3/3 uncoordinated (Unc) non-Lon recombinant animals, but 0/8 Lon non-Unc animals picked up the sa485 mutation, suggesting that sa485 lay near or to the right of lon-2 on the X chromosome (Table 4.2.). lon-2 is on the left arm of X, so we used single nucleotide polymorphism markers which alter restriction enzyme sites (snip-SNP’s) from each arm and the center of X. Unlike pkP6010 and pkP6110, we found 0/19
recovered *sa485* mutants carrying the Hawaiian (Hi) strain polymorph for pkP6139, suggesting this mutation lay on the right arm of X (Table 4.3.). We analyzed the far right arm of X using two *unc-3(e151) let* double mutants, however these two strains gave conflicting results, and the data were discounted (Table 4.2.). Next, we recovered recombinants from a mother with a *dpy-6(e14) unc-9(e101)* chromosome in trans to *sa485* to determine that *sa485* lay between these two loci on the near right arm of X. Subsequently, we used a *dpy-6(e14) sa485* double mutant strain recovered from this experiment for further mapping. First, we created *dpy-6(e14) sa485/lin-2(e1309)* animals, and recovered both *dpy-6* and *sa485* single recombinants. 4/4 Dpy mothers failing to segregate Mab (sa485) males picked up *lin-2(e1309)*, while only 1/20 Mab-segregating non-Dpy mothers was considered to have *lin-2(e1309)*. Taking the latter data (1/20) together with the observation of Egl animals in the *sa485* mutant population, however, suggested that it was equally possible we had recovered 0/20 *sa485 lin-2* double mutants. Given that discrepancy, we used the *dpy-6(e14) sa485* strain to once again cross into the polymorphic Hi strain. We recovered *dpy-6(e14)* and *sa485* single mutants, and examined the genotypes for two snip-SNP’s. 9/16 Dpy animals picked up the Hi polymorph for pkP6126, while 11/16 picked up Hi pkP6114, indicating the *sa485* genetic position to be ~3.75. Similarly, 11/31 and 2/28 *sa485* homozygotes carried these same polymorphisms, suggestive of a position near 3.00.

The relative consistency and narrow range of these mapping data lead us to consider rescue experiments. Examining the relationship between the physical and genetic maps in this region of the X chromosome revealed a non-linear correlation between sequence and genetic markers. Specifically, 38 genes are predicted to fall
between 2.86 and 3.0 on the genetic map, and these are represented in 4-5 cosmid sequences. Examination of a much larger genetic region, 3.0-4.0, reveals only 18 additional loci, represented by a roughly equivalent amount of sequence as is found from 2.86-3.0. For this reason, we chose to perform rescue experiments for the approximate genetic region 2.8-3.8. We obtained 21 cosmids and fosmids that provided 2-fold coverage of this region. After verification of cosmid identity, we tested for sa485 rescue by injecting mutant hermaphrodites with mixtures of 3-4 cosmids. Due to small brood size, stress/damage from injection, and abnormal egg-laying system development, we could not establish stable transgenic lines. Moreover, preliminary observations of transformed F1 males suggested that no cosmid pools were able to rescue the JT8174 Mab phenotype, so we turned to a candidate approach for cloning the gene affected by sa485.

Given preliminary failure to rescue lin-14(sa485) and the unusual genetic-physical map correlation, we suspected that our genetic mapping data could be “compressed” relative to true physical map position. Therefore, we began to look a further distance to the right from the pkP6114 snip-SNP. Of the 24 loci between genetic positions 4.0 and 5.0, one gene stood out as a candidate that might explain aspects of the sa485 mutant strain. In addition to the male spicule defect, we had observed several other phenotypes in the sa485 strain, including general sickness (Sck), slow movement (Unc), lethality (Let), sterility (Ste), egg-laying defects (Egl), small males (Sma), and protruding vulvas (Pvl). Most of these phenotypes had also been reported in RNAi analysis for the gene lin-14 (Kamath et al., 2003), which is located at 4.28 on the right arm of X. These shared phenotypes, as well as an unusual developmental phenotype we
had been tracking in *sa485* hermaphrodites, suggested that we should examine the *lin-14* gene in the JT8174 strain for lesions. We amplified the coding region of *lin-14* from this strain using PCR and sequenced the resulting fragments. We found a single nucleotide transition mutation in exon 10, and this is predicted to cause a glycine to arginine substitution in the LIN-14 protein (Figure 4.2.).

To confirm that *sa485* was an allele of *lin-14*, we carried out a complementation test. As *lin-14* is on the X chromosome, and males have only one copy of X, we used a genetic background that generates males, irrespective of X copy number. In this way, we were able to generate hermaphrodites that would produce male progeny of three genotypes: *sa485* homozygotes, heterozygotes bearing *sa485* and the canonical *lin-14* allele *n179*, and *lin-14(n179)* homozygotes. If *sa485* and *n179* were mutations in separate genes, we would expect the heterozygous males to have normal spicules. Therefore, in the whole male population, presence of ≥50% animals with wild-type spicules would indicate separate genes (complementation). *lin-14(n179)* males display incompletely penetrant spicule defects at the temperature tested, so a low (but non-zero) percentage of wild-type males would indicate that *sa485* was an allele of *lin-14*. Only 10% of male progeny were observed to have normal spicules, arguing that *lin-14(n179)* failed to complement the spicule defect associated with *sa485* (Table 4.4.).
4.3.2. *lin-14(sa485) mutants exhibit asynchrony between gonadal and vulval development*

In addition to the phenotypes mentioned above, we observed *lin-14(sa485)* mutants to have an unusual developmental phenotype within the hermaphrodite’s egg-laying system. A proportion of otherwise adult-looking animals were observed to have a vulval morphology indicative of the L4 larval stage. Normally, as hermaphrodites enter adulthood, the germ cells undergo a series of maturation events. Initially, germ cells most proximal to the vulva differentiate into sperm, followed more distally by differentiation of oocytes. Oocytes are then fertilized by sperm in the spermatheca, and move into the uterus. These embryos/eggs are generally produced in pairs, one from each gonad arm. Through the L4 larval stage, the vulva undergoes sequential inversion and eversion, along with other morphogenetic and differentiation processes. Simultaneous to the germline maturation events described, vulval eversion completes, yielding the mature vulval structure. In contrast to wild type, we frequently found *lin-14(sa485)* animals with mature sperm and an immature, inverted vulva, which was almost never observed in wild-type worms. We also noted the immature vulva in some animals with mature oocytes and fertilized eggs (Figure 4.3. A).

From the initial phenotypic characterization, it was not clear whether the immature vulva was a terminal and variable defect, or if it represented a transient feature, where vulval eversion simply occurred more slowly than in wild-type individuals. To address this question, we scored the proportion of hermaphrodites that had completed vulval eversion at each phase of germline maturation. The increasing percentage of *lin-14(sa485)* worms with normally everted vulvas over developmental time suggests that
this phenotype is the result of a temporal delay in vulval maturation. Therefore, we termed this the eversion of vulva-delayed phenotype (Evl-D; see note in Materials and Methods; Figure 4.3. B).

Given that \emph{lin-14} is known to regulate temporal development, it was not altogether surprising that \emph{lin-14(sa485)} mutants exhibited an asynchrony between vulval and gonadal development. However, \emph{lin-14} null mutants exhibit large stage-specific shifts in developmental events, while \emph{lin-14(sa485)} animals show a more modest temporal defect. The simplest explanation for this is that \emph{lin-14(sa485)} is a hypomorphic allele, and that reduced LIN-14 activity in these mutants causes an incompletely penetrant delayed, or retarded, phenotype in vulval maturation. If this were the case, we reasoned that reducing the level of functional LIN-14 protein by another mechanism could also elicit the Evl-D phenotype. \emph{lin-14(n179)} is a well-characterized temperature sensitive loss-of-function (\emph{lof}) mutation. We performed a series of temperature shift experiments with \emph{lin-14(n179)} mutants to effectively generate varying levels of functional LIN-14, and analyzed these animals for vulval eversion over the same gonad maturation timeframe as was used in the \emph{lin-14(sa485)} studies. \emph{lin-14(n179)} mutants grown at the permissive temperature are indistinguishable from wild type, and those grown solely at the restrictive temperature are almost totally defective in the normal vulval eversion process. Animals shifted to the restrictive temperature in the middle or end of the L1 stage, however, exhibit an eversion curve strikingly similar to that of \emph{lin-14(sa485)} animals (Figure 4.4.). Together, these data indicate that the Evl-D phenotype results from a reduction in LIN-14 activity.
It is clear that the Evl-D phenotype described above represents a loss of synchronous development between tissues. Although the Evl-D phenotype was characterized as retarded maturation of vulval tissue, this seemed to conflict with the established function of lin-14. Loss of lin-14 function has been reported to cause precocious developmental events, rather than retarded ones (Ambros and Horvitz, 1984, 1987). Due to this, we considered the alternative possibility that the Evl-D phenotype may be the result of precocious germline maturation. This was also an intriguing hypothesis, given that lin-14 has been shown to function in the vulva, but not the gonad. To address these possible interpretations, we modified an established method for analyzing developmental progression of vulval tissue (Huang and Hanna-Rose, 2006). We tracked developmental milestones within both the vulva and the gonad over real time (Figure 4.5. A-B). Larval molts were scored as additional landmarks of developmental progress. Figure 4.5. C illustrates the developmental time-course data, and raises several important points. First, these data confirm previous observations that there is no full stage-specific shift of developmental events in lin-14(sa485) mutants. Secondly, the time spent in each larval stage examined is longer for lin-14(sa485) mutants relative to wild type. Most importantly, examination of relationships between vulval and gonadal milestones reveals that synchrony of events are roughly maintained, but that there are some minor differences. The most dramatic distinction is the time in which vulval tissue arranges into the “Christmas tree” stage (milestone 5; Figure 4.5. A), which appears to happen precociously in the mutant strain, especially taking into consideration the extended L4 stage these animals exhibit. In wild-type hermaphrodites, this event is typically reached 6 hours into L4 development, almost exactly halfway through the stage.
In contrast, this event typically occurs 4 hours into the L4 stage of \textit{lin-14(sa485)} animals, just about one quarter of the way through L4. The other key difference between vulval-gonadal development in these animals is the retarded vulval eversion that we observe as the Evl-D phenotype (milestone 6; Figure 4.5. A). Together, these data suggest that \textit{lin-14(sa485)} hermaphrodites develop more slowly than wild type, that the Evl-D phenotype results from subtle changes in temporal development, and that these changes involve precocious and retarded developmental events within the vulval tissue.

\subsection{Specification of uterine uv1 cells is delayed in lin-14(sa485) mutants}

Our demonstration that \textit{lin-14(sa485)} mutants exhibit a precocious developmental event was consistent with the precocious defects observed in canonical \textit{lin-14(II)} mutant strains. Interestingly, we found the precocious achievement of vulval inversion (“Christmas tree” stage) was followed by retarded vulval eversion. This indicated that the precocious inversion event alters synchrony between the vulva and the gonad, and may temporarily disrupt signaling between these tissues. Within the gonad, we focused our analysis on the uterus, which contacts directly with the developing vulva. Several lines of evidence supported the hypothesis that vulva-uterus crosstalk is altered in \textit{lin-14(sa485)} mutants. Experiments assessing the morphology and function of uterine cells have shown that the ventral uterus is important as a structural anchor for the inverting and everting vulva (Newman et al., 1996). In addition, null \textit{lin-14} mutants display a strong protruding vulva phenotype (Pvl), and we found during our \textit{lin-14(n179)} gene dosage experiment that as LIN-14 activity decreases, the vulval structure becomes increasingly weak, and vulvas at the Christmas tree stage begin to collapse and show the Pvl
phenotype. This phenotype is clearly distinguishable from the cellular behavior during normal vulval eversion, as shown in Figure 4.6. Furthermore, lin-14 canonical loss-of-function mutants fail to specify the uv1 cells, which are ventral uterine cells adjacent to the vulva (Huang and Hanna-Rose, 2006). Finally, uv1 cells are known to require vulval cells for their proper specification, and many genes have been implicated in this specification signaling. Together, these data raised the possibility that the precocious-to-retarded shift in the lin-14(sa485) vulva could be due to altered interaction between the vulva and the uv1 cells of the ventral uterus.

To test our hypothesis, we examined uv1 cell specification in lin-14(sa485) hermaphrodites using an ida-1::GFP reporter transgene, which acts as a terminal differentiation marker for the uv1 cell fate. This transgene is initially weakly expressed in a uterine equivalence group called the π cells. Subsequently, expression becomes robust in the uv1 cells, and diminishes in other cell types. We scored animals for robust GFP expression, and found that, like vulval eversion, uv1 specification is delayed in these mutants relative to wild type. Not only did uv1 specification lag behind gonad development, we also observed a lag behind vulval development, although to a lesser degree (Figure 4.7. A). We also observed uv1 specification to be highly correlated with vulval eversion. That is, all animals with an Evl-D phenotype lacked ida-1::gfp-expressing uv1 cells, and all animals with normal vulval eversion had GFP(+) uv1 cells (Figure 4.7. B). This suggested that a delay in uv1 specification is likely to be one cellular defect underlying the gross anatomical defect described as the Evl-D phenotype.
4.3.4. The Evl-D phenotype results from an early defect in lin-14 function

Phenotypic analysis of the epithelial seam cells in lin-14 loss- and gain-of-function mutants has revealed that lin-14 directs L1-specific developmental patterning (Ambros and Horvitz, 1984, 1987). This genetic analysis has been complemented by molecular assays revealing regulation of lin-14 by the microRNA-encoding heterochronic gene lin-4, which binds lin-14 mRNA and leads to its degradation during the L1-to-L2 transition (Ha et al., 1996). Indeed, neither lin-14 mRNA nor LIN-14 protein is detectable after the L2 stage (Ruvkun and Giusto, 1989). However, our analysis of lin-14(sa485) mutants has shown defects late in development (L4 and young adult), prompting us to ask when LIN-14 is required for the functions associated with vulval integrity and eversion, and uve specification. We examined the LIN-14 time of action for regulating vulval eversion with a reciprocal temperature shift assay. Figure 4.8 illustrates that LIN-14 is required before the L2 stage for proper vulva-gonad synchrony. We conclude that the lin-14(sa485) defects arise during L1, and that these defects only manifest later in development.

4.4. DISCUSSION

4.4.1. Identification of a new allele of the heterochronic gene lin-14

We report here identification and characterization of a novel allele of the lin-14 locus. Existing mutations in lin-14 are clustered in two regions. Loss-of-function mutations are found in the coding region corresponding to the C-terminus of this transcription factor. In contrast, gain-of-function lesions alter the lin-14 3’ UTR, where
they disrupt miRNA binding and translational regulation (Ruvkun et al., 1991). The allele discussed here—lin-14(sa485)—corresponds to a transition mutation in exon 10 that results in a glycine to aspartic acid substitution at position 379 in the major LIN-14 isoform. The functional domains of LIN-14 have not been elucidated, so the structural impact of the sa485 G379D substitution is unclear. However, the high recovery frequency of reversion alleles affecting the C-terminal region of LIN-14 (encoded by exons 8-11), suggests that this domain is critical for an essential LIN-14 function such as DNA binding or transactivation.

Although the specific impact of the lin-14(sa485) mutation on LIN-14 biochemical function is unresolved, several lines of evidence suggest that it is a hypomorphic allele. As mentioned, this mutation affects LIN-14 protein sequence in a region altered by known lin-14(lf) alleles. Secondly, although no formal dominance tests were conducted, lin-14(sa485) appears recessive to the wild-type allele with respect to the Evl-D phenotype, as well as for the Mab male phenotype in tra-2 mutants. In addition, although lin-14(sa485) animals do not exhibit the robust precocious L2 development phenotype seen in lin-14(n179ts) or lin-14(n540) null mutants, these animals do exhibit weak precocious phenotypes normally associated with reduction of lin-14 activity. For example, occasionally lin-14(sa485) mutants prematurely generate portions of adult cuticle during the L4 larval stage. Finally, despite the delayed vulval eversion detected in the lin-14(sa458) strain, our timecourse experiment revealed that this retarded phenotype is preceded, and likely caused by, a precocious vulval phenotype.
4.4.2. An L4 function for the L1-specific transcription factor LIN-14

lin-14 RNA and protein are present during, and necessary for, the L1 stage. It remains an open question as to how the reduced activity of lin-14(sa85) activity during L1 is carried through to the morphogenetic behavior of vulval precursor cell (VPC) and ventral uterine (VU) progeny during L4. Stage-specific analysis of lin-14 mRNA and protein levels find no persistence of these molecules to the time Uta, uv1 fate, and Evl-D phenotypes arise (Ruvkun and Giusto, 1989). An attractive explanation is that LIN-14, as the initial transcriptional regulator of the heterochronic pathway, sets in motion a temporal cascade that regulates all steps of developmental progression. We interpret that this temporal cascade does not act as discrete larval stage switch mechanism; instead, even modest alterations in the source of this regulatory pathway are faithfully transmitted through developmental time.

The most compelling gene to be considered as a downstream mediator of the lin-14(sa485) phenotypes is lin-29. lin-29 encodes a zinc-finger transcription factor necessary for L4-specific developmental events. Importantly, defects associated with mutations in lin-29 are consistent with every phenotype we observed in lin-14(sa485) mutants. lin-29 is necessary for uv1 cell specification, morphogenesis and gene expression in the vulva, and the integrity of the vulva-uterus-seam connection (Bettinger et al., 1996, 1997; Newman et al., 2000). lin-29 is already known to mediate lin-14 function in late seam cell development (Ambros, 1999), but it was previously not thought to act downstream of lin-14 in the vulva or uterus. In combination with these reports, our data suggest that slightly precocious lin-29 expression is a plausible mechanism for Evl-D and the other lin-14(sa485) phenotypes.
One question arising from this work is why differentiation, but not the initial cell
divisions, of vulval precursor cells (VPCs) is altered in lin-14(sa485) mutants. lin-14(ef)
mutations have been shown previously to shorten the cell cycle of the VPC’s (Euling and
Ambros, 1996). lin-29 mutants exhibit normal execution of VPC cell divisions, but show
defects in vulval maturation (Newman et al., 2000). Together these data support a model
wherein LIN-14 target gene levels involved in cell cycle progression are unaffected,
while targets that play a role in affecting morphogenesis—perhaps through lin-29—are
altered. Another possibility is that the timing of the third round of cell divisions is
affected. These occur around the time of the L3 molt, and we did not track this final
round of division closely. In either case, we can conclude that lin-14(sa485) disrupts
some, but not all, processes previously reported for lin-14(ef) alleles.

4.4.3. lin-14(sa485) mutants exhibit asynchrony between gonadal and vulval
development

Here, we demonstrate that a reduction in lin-14 function can have an effect on the
temporal development of the vulva. Specifically, vulval inversion is completed
prematurely within the L4 larval stage compared to wild type, yet no previous
developmental milestones of the vulva are achieved precociously. One implication of this
finding is that lin-14 not only promotes a stage-specific developmental program (L1), it
can also influence intra-stage progression. This is an important distinction from other
studies of lin-14 function, and supports the idea that intra-stage temporal development is
not regulated as a biological switch. Our data suggest that some temporal events respond
to altered lin-14 function in a more continuous manner. Mechanistically, the G379D
substitution may reduce the DNA binding or transactivation capacities of LIN-14, and such functional alterations could cause quantitative changes in target gene transcription, impacting temporal development in a less dramatic fashion than canonical lin-14(If) alleles. We speculate that LIN-14 activity in lin-14(sa485) animals is sufficient to promote the L1 stage, but leads to developmental changes that ultimately yield a shift in the progression of vulval cell differentiation. This shift manifests at a time-point when precise temporal synchrony between vulval and gonadal tissue is required for proper maturation of the egg-laying system.

To track temporal defects in the egg-laying system, it was necessary to establish a standard measure of time with which to study developmental progression. Due to their slow growth, real time could not be used to compare lin-14(sa485) mutants to wild type. Molting could not be used because the cuticle-secreting seam cells are under the control of lin-14 activity, as are most post-embryonically developing somatic tissues (Ambros and Horvitz, 1987). The early embryonic separation of germline from soma and their inherent developmental differences suggested that the germline could provide the temporal scale to study lin-14 function in the egg-laying system. This choice was not without drawbacks, however, as the early milestones we used to track gonad development involve somatic, rather than germline tissue. For instance, the extent of distal tip cell (DTC) migration does not necessarily reflect germ cell development and maturation. However, germline proliferation and differentiation are dependent in part upon the DTC and other cells of the somatic gonad, indicating that these processes are linked (McCarter et al., 1997; Pepper et al., 2003; Killian and Hubbard, 2005). In addition, temporal regulation by the heterochronic pathway is not reported to directly influence germline
behavior. In turn, this allows us to be fairly confident that the stepwise events of germline maturation (but not necessarily gonad migration) provide a reliable “clock” with which to study alterations in maturation of the soma, including vulval and uterine tissue.

4.4.4. The cellular role of lin-14 in regulating temporal synchrony of the developing egg-laying system

Our findings suggest a model wherein alterations in vulval morphology during the L4 larval stage result from reduced LIN-14 activity during the L1 stage in lin-14(sa485) mutants. These mutants prematurely reach full vulval inversion, and this is coincident with a temporary disruption in uv1 specification. Subsequently, vulval eversion is retarded to a similar extent as uv1 cell differentiation. In more severe lin-14 loss-of-function conditions, uv1 specification is lost altogether and there is collapse of the inverted vulva. We hypothesize that uv1 specification and vulval eversion are parallel consequences of the heterochronic vulval defect. That is, the consecutive nature of these delay phenotypes is not indicative of a cause-and-effect relationship, but instead may reflect two measures of the same primary abnormality. This interpretation is suggested due to the lack of a similar Pvl phenotype in other genetic mutants with defects in uv1 specification. For example, egl-38 is required to promote the uv1 cell fate. In egl-38(n578) mutants, the uv1 cells are transformed to the uterine seam (utse) cell fate, but unlike lin-14 mutants, egl-38 mutants do not generally exhibit a Pvl phenotype (Chamberlin et al., 1997; Rajakumar and Chamberlin, 2007). The gene cog-3 has also been shown to lack uv1 cells, which undergo necrosis prior to differentiation. Like egl-38 mutants, cog-3(ku212) animals do not exhibit a strong Pvl phenotype (Huang and
Hanna-Rose, 2005). These data point to the interpretation that some tissue in addition to uv1 cells is necessary for proper vulval integrity and eversion. One possibility is that vulval integrity is dependent upon the utse, and that the development of the entire π cell equivalence group is delayed in lin-14(sa485) hermaphrodites. Our assays would thus detect delays in uv1 specification directly via ida-1::gfp, and indirectly in the utse via vulval eversion. Alternatively, cell types such as the vulval muscles or seam cells may provide the main structural anchor necessary for vulval integrity, inversion, and eversion. Taken with the correlations between lin-14(sa485) and lin-29 phenotypes mentioned above, we hypothesize that these phenotypes results from the disruption of a LIN-14—LIN-29 transcriptional cascade.
Table 4.1. \textit{sa485} linkage analysis
*Genetic marker strains carried mutations in these genes

<table>
<thead>
<tr>
<th>Genetic Marker*</th>
<th>Marker Chromosome</th>
<th>\textit{sa485} homozygotes segregating marker phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{dpy-5}</td>
<td>I</td>
<td>6/8</td>
</tr>
<tr>
<td>\textit{rol-6}</td>
<td>II</td>
<td>5/8</td>
</tr>
<tr>
<td>\textit{unc-32}</td>
<td>III</td>
<td>6/8</td>
</tr>
<tr>
<td>\textit{unc-5}</td>
<td>IV</td>
<td>10/11</td>
</tr>
<tr>
<td>\textit{dpy-11}</td>
<td>V</td>
<td>4/11</td>
</tr>
<tr>
<td>\textit{lon-2}</td>
<td>X</td>
<td>0/11</td>
</tr>
</tbody>
</table>

Figure 4.1. Genetic map of X chromosome markers used for \textit{sa485} mapping

Table 4.2. Three-point \textit{sa485} mapping
*Heterozygous parents of recombinants carried an AB/C chromosomal orientation. Genetic marker strains carried mutations in these genes; SNP’s carried altered restriction enzyme site sequence.
Table 4.3. Two-point mapping using three snip-SNP’s across the X chromosome
*sa485 homozygous animals from mothers with sa485 and polymorphic chromosomes.

<table>
<thead>
<tr>
<th>snip-SNP marker</th>
<th>SNP genetic position</th>
<th>sa485 homozygotes carrying Hi polymorph*</th>
</tr>
</thead>
<tbody>
<tr>
<td>pkP6010</td>
<td>-17.6</td>
<td>5/7</td>
</tr>
<tr>
<td>pkP6110</td>
<td>-0.8</td>
<td>2/7</td>
</tr>
<tr>
<td>pkP6139</td>
<td>16.0</td>
<td>0/19</td>
</tr>
</tbody>
</table>

Table 4.4. lin-14(n179) fails to complement lin-14(sa485)
*Specific Mab phenotype was a reduced/absent spicule cuticle phenotype.
Figure 4.3. *lin-14(sa485)* mutants exhibit a delay in vulval eversion relative to germline maturation

(A) Nomarski images and corresponding diagrams of the maturing egg-laying system in wild type and *lin-14(sa485)* mutants. In wild-type adults, vulval eversion is always completed by the time the first fertilized embryo enters the uterus. A portion of *lin-14(sa485)* adults that have reached the same stage in germline maturation, however, display vulval morphology indicative of the L4 larval stage. Vulvas are marked by arrows. (B) Vulval eversion as a function of developmental time. Temporal measurements reflect four stages of germline maturation that occur during the L4-to-adult transition: sperm differentiation, oocyte maturation, 1 and 2 fertilized embryos (per gonad arm).

Continued on next page…
Figure 4.3. continued.
Figure 4.4. Intermediate levels of *lin-14* activity in *lin-14(n179)* mutants result in the *Evl-D* phenotype
Vulval eversion progression graph illustrating several timepoints in which *lin-14* activity was disrupted by temperature shift. “15°C” represents constant growth at the permissive temperature. All other labels indicate the time at which animals were shifted from 15°C to 25°C. Failure to complete normal vulval eversion reflects the Pvl phenotype and incomplete vulval eversion. n values for each condition/timepoint combination were 14-18.
Figure 4.5. Precocious completion of vulval inversion precedes Evl-D phenotype
(A) Key for vulval development milestones: 1st and 2nd cell divisions of VPC’s, [1, 2]; vulval inversion, [3]; “pyramid” stage where the AC sits dorsal to the inverting vulva, prior to its fusion with the utse, [4]; “Christmas tree” stage where vulval inversion is complete and vulval cells are extending processes to complete vulval tube formation, [5]; completion of vulval eversion forming the mature vulva, [6]. (B) Key for gonadal/germline development milestones: the AC becomes distinguishable from other somatic gonad nuclei, [A]; dorsal turn of the DTC, [B]; proximal migration of DTC has reached position overlying the vulva, [C]; sperm differentiation is apparent, [D]; production of the 1st fertilized embryo, [E]. (C) Timecourse illustration for vulval and gonadal development milestones of synchronized larvae, listed as hours post feeding. Larval molts are indicated by double vertical lines. Milestone positions mark time at which ≥50% of worms have reached the indicated stage of development.
Figure 4.6. The Pvl phenotype in *lin-14(lf)* mutants is a result of vulval collapse during the L4 larval stage
Nomarski images of vulvas in *lin-14* mutants. (A) Normal eversion in a *lin-14(sa485)* animal. (B) Collapsing vulva in a *lin-14(n179)* L4 larva shifted from permissive to restrictive temperature during the L1 stage. (C) Terminal Pvl phenotype of *lin-14(n179)* mutant grown at restrictive temperature.
Figure 4.7. uv1 specification is delayed in lin-14(sa485) animals, and delay is coincident with Ev1-D phenotype

(A) Proportion of animals expressing the uv1 terminal differentiation marker ida-1::gfp during maturation of the egg-laying system. (*) Vulva and gonad stages with a shared number label are roughly comparable to one another in the real time development of wild-type worms. DIC images are taken in the central plane to illustrate vulval morphology. (B) Nomarski and fluorescent images of wild type and lin-14(sa485) adults carrying ida-1::gfp. Nomarski images were taken in the central focal plane of each animal to show stage of vulval development. Fluorescent images were taken in lateral plane of the same animals to show GFP-expressing uv1 cells (marked with brackets in all images). Additional blurred fluorescence is from expression of ida-1::gfp in the HSN.
Figure 4.7. continued.
Figure 4.8. Abnormal vulval eversion in *lin-14* mutants is due to LIN-14 activity during the L1 larval stage
Final vulval eversion phenotype of *lin-14*(n179) animals subjected to different temperature shift combinations. All animals were scored as early adults. Failure to complete normal vulval eversion reflects the Pvl phenotype and incomplete vulval eversion. n ≥55 for all combinations.
CHAPTER 5

DISCUSSION

5.1. Project relationships

Although the specific genes of interest in the work presented here are distinct, the conclusions we are able to draw from them are relevant to one another. In this way, we have contributed to an interconnected and broad foundation for understanding transcriptional regulation of C. elegans post-embryonic development. Below, we discuss some of the ways in which pertinent connections can be made between my vab-3, egl-38, and lin-14 studies.

5.2. Transcriptional regulation of tissue development and function: lessons from egl-38

Through genome-wide expression analysis, we have uncovered many potential target genes for the transcription factor EGL-38. Our particular cross-referencing approach identified a set of genes robustly enriched for egl-38-dependent expression and egl-38-dependent process function. In addition, this strategy uncovered a novel role for this Pax2/5/8 factor in innate immunity. Overlapping transcript lists derived from three
distinct experiments defined the validated candidate gene set, as well as other sets for future investigation. Cross-referencing of microarray outputs has become a controversial topic in recent years. This approach has been impaired due to relatively low levels of reproducibility in terms of identifying common gene responses to a particular cellular state (rev. in Walker and Hughes, 2008). However, questions about the approach are centered mainly on cross-referencing data from multiple hybridization platforms and/or experimental designs (rev. in Marshall, 2004). Alternatively, cross-referencing of array data generated in parallel has been demonstrated to be successful in many cases (for example, see Fisher and Lithgow, 2006). Taken together, this report and others have shown that the use of overlap as a means of identifying genes of interest from multiple microarray experiments can be successful in cases where statistical data filtration is limited. Furthermore, the multiple research avenues opened up by our experimental dataset demonstrates that this approach may not only be a situational necessity, in some instances it may also be preferable to single experiment datasets passed through stringent and complex statistical filtering.

A question that must always be considered is the applicability of what we have learned to other contexts. There are several reasons that these experiments can be more broadly informative in the context of transcription factor biology. As mentioned above, we have shown that much of the noise inherent to EGL-38 microarray experiments can be eliminated through the use of biological filters in place of statistical filters. There is no clear reason why this should not be the case for the study of other transcription factors, provided there exist a multitude of ways in which to modulate their activity in vivo.

Secondly, the combinatorial nature of control over multiple tissues by Pax transcription
factors is highly conserved. Therefore, the activity of EGL-38 in process-specific
regulation and simultaneous positive and negative regulation of individual targets can be
used as a framework to test for similarities and differences among other Pax factors.

Sequence-specific transcription factors share many regulatory and mechanistic
features, suggesting that the lessons learned here may have applicability to the regulation
of developmental gene expression in general. Conclusions from this microarray data can
also inform the nature of tissue-specificity vs. general target gene regulation. Although
we have not fully validated the distinct sets of candidate genes, nor have we analyzed the
cis-regulatory sequences of these genes, full annotation of the C. elegans genome and an
assortment of informatic tools make it possible to study these features of our dataset.
Using these tools, and other techniques, we can ask whether or not there are different
types of egl-38 response elements among the candidate genes, whether those are direct
EGL-38 binding sequences, and what other elements are present that may account for the
combinatorial control by additional transcription factors. In summary, the microarrays
designed, performed, analyzed, and validated in this work have provided avenues for
further research, as well as insights into related studies in transcription factor biology and
development.

Considering the success of our microarray approach for the study of EGL-38, one
interesting research path laid out by our work would be a set of similar microarray
experiments searching for VAB-3 target genes. Like EGL-38, VAB-3 controls
developmental processes in multiple tissues (Zhang and Emmons, 1995; Chisholm and
Horvitz, 1995; Chamberlin and Sternberg, 1995). Additionally, multiple classes of C.
elegans pax-6 mutations have been characterized: those affecting vab-3, those affecting
the \textit{mab-18} transcript that arises from an internal promoter, and those that affect both \textit{vab-3} and \textit{mab-18} (Cinar and Chisholm, 2004). Thus, as we performed individual microarray experiments for distinct tissue-specific alleles of \textit{egl-38}, one could do the same to assess the similarities and distinctions between target gene profiles of these two PAX-6 isoforms. Furthermore, studies have demonstrated the utility of inducible expression as a means to generate an artificial gain-of-function condition (this work; Park et al., 2006; Wang et al., 2006), suggesting that a \textit{hsp-41::vab-3} strain would also be feasible for such experiments.

\textbf{5.3. The role of temporal interactions in Pax-dependent tissues: lessons from lin-14}

Investigation of a novel allele of the transcription factor LIN-14 has defined a heterochronic function for this protein that manifests in late larval development. Although previous reports of \textit{lin-14} function include mutant defects observed in later post-embryonic development, these are all interpreted as the indirect effects of temporal shifts early in development. In the case of \textit{lin-14(sa485)} hermaphrodites, no early defects are seen, suggesting that \textit{lin-14} can impact late developmental events in the absence of gross first larval stage heterochronic transformations. In addition, our analysis shows that \textit{lin-14} can exert temporal control over time increments less than full larval stages. Combined with the current understanding of \textit{lin-14} function, our work provides one example of the relationship between modular and stepwise developmental timing events. We can conclude that a modular developmental timing switch is the result of dramatic alterations in LIN-14 activity, while small modifications in this activity do not affect the modular switch, but rather induce changes in stepwise developmental progression. This
conclusion is supported by the mutant phenotypes associated with “complex” *lin-14(gf)* intragenic revertant mutations, which cause a mosaic of precocious and retarded events across multiple tissues (Reinhart and Ruvkun, 2001).

Specification and morphogenesis of the cells comprising the *C. elegans* egg-laying system are known to require *egl-38* function (Chamberlin et al., 1997; Rajakumar and Chamberlin, 2007). One goal we set out to understand was how *egl-38* regulates the development of these tissues, and we approached this through the identification of downstream effector genes. In the larger picture, Pax proteins and other transcription factor families are known to control target gene expression in a combinatorial fashion (rev. in Chi and Epstein, 2002; Ogata et al., 2003). Studies of the egg-laying system in *lin-14(sa485)* animals hinted at another example of combinatorial control. Unlike the combinatorial control mediated by spatially distributed transcription factors, our work has shown that the temporal regulator *lin-14* is important for precise cellular interactions between *egl-38*-dependent tissues. For instance, proper specification of the uv1 cells requires both *egl-38* and *lin-14* function. Although LIN-14 itself is not present during uv1 specification, we and others have demonstrated uv1 defects in *lin-14(If)* mutants (Huang and Hanna-Rose, 2006) influence uterine development, presumably through other heterochronic pathway components. Two potential candidate effectors are the transcription factor LIN-29 and the miRNA *let-7*. *lin-29* function is necessary for induction of the uv1 cell precursor π cells (Newman et al., 2000), and *let-7* RNA levels are controlled by the same Notch signaling that induces these cells (Soloman et al., 2008). Thus, *lin-14* activity provides uv1 cell competence through some downstream transcriptional mediator, and *egl-38* helps to specify the uv1 cells. Whether EGL-38 and
heterochronic regulators control expression of the same target genes is unclear. However, these pathways provide an example of combinatorial control at the cellular level, if not the transcript level.

Another emergent theme in our lin-14 and egl-38 work is the recognition that each of these transcription factors not only displays tissue-specific roles, but process specific roles within a single tissue, as well. Canonical lin-14(lef) mutations have been shown to cause precocious VPC cell division, as well as to cause the morphological Pvl defect in the adult vulva (Ambros and Horvitz, 1987; Euling and Ambros, 1996). We found that the lin-14(sa485) mutation preferentially disrupts the morphogenesis of the vulval cells while leaving the cell cycle progression of the VPC’s unaltered. Similarly, egl-38 mutations known to disrupt transcriptional regulation in the rectum have been described (Chamberlin et al., 1997, Johnson et al., 2001). In this work, we have found that although egl-38(n578) mutants do not exhibit the developmental defects in the rectum observed in other mutants, these animals are severely affected in terms of susceptibility to M. nematophilum infection. Therefore, for both lin-14 and egl-38, I have shown that certain mutations can disrupt a subset of biological processes in the same tissue. This implies that these transcription factors can impact distinct groups of target genes within the same tissue. While the formation of a single tissue-specific trans-acting complex can account for control of one set of target genes, it cannot be the sole mechanism for generating control of multiple groups of target genes. Combinatorial control of target gene transcription for both EGL-38 and the lin-14 downstream effector must include sets of transcripts with distinct cis-regulatory elements as well, and therefore, multiple trans-acting factor combinations/complexes.
5.4. Genetic controls of transcription factor expression: lessons from vab-3

Through construction of a transcriptional reporter transgene, we were able to study sensory organ expression of Pax6/vab-3 in the C. elegans male. Genetic analyses indicate that vab-3 expression in these sensory organ precursors is achieved through establishment of an activation domain in the tail, and subsequent refinement of that expression domain through different repressive inputs. Although we have not demonstrated direct activation of vab-3 transcription by any regulatory proteins, cis-elements in the vab-3::gfp sequence include one or more binding sites for the proteins EGL-5, EGL-38, LIN-48, POP-1, and VAB-3 itself (data not shown). Irrespective of the biochemical mechanisms involved, this work has demonstrated a highly efficient logic underlying the restriction of vab-3 to just two blast cells in the worm tail.

An unanswered question following from this work is whether this regulatory logic represents something unique, or whether it could be conserved beyond the system studied. Could domain refinement be reiterated in the control of vab-3 elsewhere in the worm? Are the expression domains of other so-called master control genes limited in this manner? Considering the importance of vab-3 in specifying sensory mating structures, it is not surprising that the method of transcriptional regulation involves a complex and layered series of genetic inputs. Even a single additional cell expressing vab-3, as in the case of U cell expression in egl-38(sy294) mutants, can virtually eliminate the production of a functional mating organ (this work; Chamberlin et al., 1997). It follows that other master control genes are likely to be under precise, layered control, as well. For instance, multiple repressive inputs for the osteogenic master control gene Runx2 have been discovered (Wiper-Bergeron et al., 2007). Although no positive regulatory factors have
been found, it is plausible that these inputs act to limit the activity of some transcriptional activator(s). One can imagine how this conclusion applies to our work on Pax2/5/8/egl-38, as well. Though the regulatory inputs of egl-38 transcription are not known, the powerful organogenic function of this factor would certainly make restriction of this transcription factor crucial. Therefore, it is reasonable to expect that egl-38 expression is dependent upon multiple layers of transcriptional regulation.

In our work describing transcriptional inputs for Pax6/vab-3 expression, we noted that ectopic expression events can be as detrimental to mature organ function as is the loss of normal expression. Given this, it is not surprising to find that more than half of the candidate EGL-38 target genes are predicted to be repressed by egl-38 activity in at least some tissues. Although the three genes identified as egl-38-responsive genes with immune function are all under positive regulation in the rectum, two of the three are under negative regulation in other tissues throughout the body. As for activation of vab-3 expression by egl-5, there must be some additional transcriptional regulator(s) responsible for activating expression of these target genes. In point of fact, there is evidence for such a layered regulatory mechanism. Transcription of egl-17 in the developing vulva is genetically repressed by egl-38 in a subset of cells in which its expression is activated by the LIM homeobox transcription factor lin-11 (Fernandes and Sternberg, 2007). Together, the layering of positive and negative genetic inputs shown by our vab-3 work provides an insightful perspective with which to further understand the regulation of EGL-38-responsive genes we have identified.
A link between the layered regulation of vab-3 and our lin-14 results may be even more direct. While we found the spatial vab-3 expression domain to be limited by multiple repressive inputs, we did not examine the temporal restriction of vab-3 expression. Interestingly, the vab-3 activator egl-5 is expressed in the B cell prior to division, well before vab-3 reporter expression can be observed. Given the role of Wnt regulation and the deterministic nature of asymmetric cell divisions generally (rev. in Mizumoto and Sawa, 2007; Knoblich, 2008), it is likely that B cell division is the trigger releasing vab-3 from its repressed state in this tissue. However, it is conceivable that regulation of that division and/or direct transcriptional repression of vab-3 by LIN-14 plays a critical role in precisely timing the onset of vab-3 expression. In support of this, the late L1 timing of B cell division is coincident with the down-regulation of LIN-14 translation by the miRNA lin-4. Furthermore, inspection of vab-3 regulatory sequence DNA reveals the presence of a consensus LIN-14 binding site (Hristova et al., 2005; data not shown), and division of the B cell in lin-14(sa485) mutants is frequently abnormal (data not shown). Taking this into the broader context of vab-3 regulatory events, these results indicate that beyond spatial domain refinement, temporal transcriptional boundaries are likely to help define the expression of Pax6/vab3 in male sensory organ development.

5.5. Significance of transcription factor specificity and temporal regulation models

Many developmental disorders are the consequence of mutations encoded in transcription factors (for examples, see Nemer, 2008; Lehmann et al., 2003), and transcription factor mutations are among the most frequent in cancers (for examples,
Chumakov, 2007; Oneil and Look, 2007). The ability of these proteins to govern the expression of tens or hundreds of genes, and therefore tissue identity and function, make clear why this is the case. For the same reason, targeting transcription factors for therapeutics is an attractive area of pursuit (rev. in Stanczyk et al., 2008; Heaney, 2007; Penna et al., 2007). Our work has illustrated that such application-driven goals must carefully address the specificity and genetic network complexity of these factors.

The highly conserved nature and pleiotropic regulatory behavior of Pax proteins is both an advantage and a disadvantage to the pursuit and application of Pax-based therapeutics. Specifically, development of a reagent to target one Pax gene/protein may provide a firm basis for targeting other Pax family members. However, achieving the specificity of a therapeutic agent not only between Pax genes, but also between distinct functions of a single Pax gene is a significant challenge. The discovery that egl-38 can regulate multiple processes within the same tissue further complicates the specificity problem. Simply delivering a therapeutic agent to the desired tissue to disrupt one Pax function may unintentionally affect other important cellular processes. This ties into the layered regulation of vab-3, as well, which indicates that the complex regulatory logic underlying transcription factor expression could affect targeting of a Pax factor. If one chose to eliminate EGL-38’s immune susceptibility function in the rectum of males, for example, a necessary repressive input into vab-3 regulation would concurrently be disrupted. That is, while the goal may be to regulate critical targets of Pax factors, the fact that these proteins often act in transcriptional cascades complicates this goal. Although the specific applicability of these projects to higher animals remains unknown,
they have substantiated genetic principles to be considered in ongoing work to target transcription factor activities for biomedical purposes.

LIN-14 is a novel transcription factor with no clear vertebrate orthologs (Hong et al., 2000). Many of the components of the C. elegans heterochronic pathway, however, are conserved in higher organisms, and play roles in temporal regulation. A lin-28 ortholog is temporally expressed along the axis of epithelial maturation in the intestine of vertebrates (Yang and Moss, 2003). This process, as well as the circadian clock mechanism involving the heterochronic gene lin-42, can be viewed as stepwise temporal progression events. Alternatively, they can also be construed as modular processes, with each cycle representing the initiation point of a new temporal module. The blurring of stepwise and modular forms of temporal regulation is a key discovery from our work with lin-14. The fact that lin-14 can play a role in both types of control mechanisms suggests that they need not be two processes, but rather different representations of a shared temporal control.

5.6 Transcriptional control over C. elegans immunity: lessons from everyone

While innate immunity is “hard-wired”, the engagement of multiple signaling pathways allows for pathogen-specific responses. As mentioned previously, different bacterial pathogens have been shown to elicit distinct gene expression changes in C. elegans (Wong et al., 2007). The implication of these findings is that individual pathogens activate various combinations of signaling pathways, and the combined action of pathway transcription factors leads to specific sets of target gene expression changes. Individual signaling pathways can induce a set of immune response genes on their own
(Troemel et al., 2006). However, the complexity of different immune responses suggests that pathogen-specific expression changes are established through synergistic, as well as additive, transcriptional regulation. That is, some target genes are likely induced in response to two or more transcriptional inputs, but not by any one pathway alone.

In this work, we found that appropriate vab-3 expression was dependent upon multiple transcriptional inputs. Specifically, expression appears to be activated in all cells across the tail region not under the influence of a repressive genetic control. The pathogen-specific nature of immunity described above suggests that immune effector gene expression is also precisely controlled through layered transcriptional regulation. Beyond synergistic activation by multiple signaling events, one can imagine a mode of effector gene activation similar to that of vab-3 regulation. A tissue-specific activator may allow for expression, while one or more repressive inputs may prevent expression. Upon pathogen exposure, the cognate response pathway(s) is activated, and the repression is overcome. There is precedent for regulatory suppression of an immune response in invertebrates, where the Drosophila gene caspar acts as an inhibitory factor the Imd-type immune response (Kim et al., 2006). While the involvement of negative regulatory inputs into immune effector gene expression need not fit into the transcriptional control model observed for vab-3, parallels can be drawn regarding the action of combinatorial regulation as a means to precisely control transcription important for animal health.

Setting aside the regulatory mechanisms of immune effector expression, it is clear that certain tissues are relied upon heavily to fight the establishment of infection. The primary barrier to infection by pathogenic organisms is the outer layer of the animal
body. Human skin, for instance, is capable of releasing anti-microbial peptides to ward
off infection (rev. in Goodarzi et al., 2007). By virtue of this protection, the predominant
mode of host infection is through natural entry points such as the mouth. In *C. elegans*,
there are three entry points large enough to allow bacterial invasion: the mouth, the vulva,
and the rectum. Therefore, it is reasonable that the expression of some immune effectors
be concentrated in these tissues. To illustrate this point, the pharynx expresses the Toll-
like receptor, TOL-1, as well as anti-microbial proteins like astacins and defensins (Tenor
and Aballay, 2008; Mohrion et al., 2003).

While the full expression pattern of egl-38 is not known, it is noteworthy that egl-
38 functions in at least two of the three tissues exposed to pathogens. The correlation of
egl-38 expression and exposed tissues may be coincidental. This is supported by
expression of the gene *C05C9.1*, which we found causes susceptibility to *M.
nematophilum* infection. *C05C9.1* is downregulated in the egl-38 mutant rectum,
however it is simultaneously upregulated in the vulva, arguing against a conserved role
for egl-38-regulated immune function in the vulva. Alternatively, expression of egl-38 in
exposed tissues may explain the evolutionary recruitment of egl-38 into the regulation of
genes with immune function. If this were the determined to be the case, one might hope
to identify other transcription factors with immune function in part by such an expression
pattern-exposed tissue correlation.

Whether exposed directly to the environment, or serving as internal sites of
pathogen infection, cells and tissues must be capable of mounting a controlled immune
response via expression of effector molecules. Loss of important effector gene
expression can critically reduce cellular competence to respond to pathogen presence (for
example, see Sang et al., 2005). Misexpression of immune effectors in the absence of a pathogen can also be detrimental (for example, see Mrug et al., 2008). While part of that competence arises from spatially established expression patterns, temporal competence can also impact the immune response. For instance, in our C. elegans—M. nematophilum studies, we exposed synchronized populations to the pathogen at a fixed time-point. While the chosen time corresponds to the L4 stage of wild-type animals, some slow-growing strains tested had slightly younger animals at the same time-point. These strains tended to show higher levels of infection. Moreover, males were noted to have higher infection levels than hermaphrodites. This could be due to the fact that the male rectum is undergoing morphogenesis during the L4 stage, and the cells have not yet fully differentiated.

I have shown that even partial loss of lin-14 activity can lead to significant developmental consequences. Therefore, it would not be surprising to find that mutations in lin-14 and other heterochronic genes can affect susceptibility to pathogens via temporal competence. Specifically, the heterochronic pathway was initially characterized by alterations in external cuticle structure and development of the hypodermal cells that produce it (Ambros and Horvitz, 1984). The cuticles of C. elegans are composed of many types of molecules, and several of these have been implicated in susceptibility to biofilms as well as M. nematophilum (Darby et al., 2007; O’Rourke et al., 2006). Whether through cuticle alterations or other abnormally developing tissue, the temporal nature of cellular competence is likely to impact the ability of C. elegans to fight bacterial infection.
APPENDIX A

OLIGONUCLEOTIDE PRIMERS
<table>
<thead>
<tr>
<th>Primer</th>
<th>amplicon</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR24</td>
<td>vab-3' intron 2-exon 3</td>
<td>ggagcaaacaggcttgc</td>
</tr>
<tr>
<td>PR25</td>
<td>vab-3' intron 2-exon 3</td>
<td>ccagcttccagtttgtcagcttcacaaattac</td>
</tr>
<tr>
<td>PR367</td>
<td>vab-3' intron 2-exon 3</td>
<td>aaagcgcagctagggcctag</td>
</tr>
<tr>
<td>PR368</td>
<td>fragment with snip-SNP pkP601</td>
<td>ccaacacatctacatagctcag</td>
</tr>
<tr>
<td>PR369</td>
<td>fragment with snip-SNP pkP601</td>
<td>tttctgacacacgcttgcag</td>
</tr>
<tr>
<td>PR370</td>
<td>fragment with snip-SNP pkP601</td>
<td>ccaacacatctacatagctcag</td>
</tr>
<tr>
<td>PR371</td>
<td>fragment with snip-SNP pkP601</td>
<td>aatcttcgagcgccttag</td>
</tr>
<tr>
<td>PR372</td>
<td>fragment with snip-SNP pkP601</td>
<td>aatcttcgagcgccttag</td>
</tr>
<tr>
<td>PR1515</td>
<td>fragment with snip-SNP pkP601</td>
<td>cgacagcttcttccagc</td>
</tr>
<tr>
<td>PR1516</td>
<td>fragment with snip-SNP pkP601</td>
<td>acacactgaacacctccacacctcc</td>
</tr>
<tr>
<td>PR1829</td>
<td>fragment with snip-SNP pkP601</td>
<td>aacagccagctagggcctag</td>
</tr>
<tr>
<td>PR1830</td>
<td>fragment with snip-SNP pkP601</td>
<td>aacagccagctagggcctag</td>
</tr>
<tr>
<td>PR1895</td>
<td>fragment with snip-SNP pkP601</td>
<td>taactttcaactgttccacag</td>
</tr>
<tr>
<td>PR1896</td>
<td>fragment with snip-SNP pkP601</td>
<td>ccaacacatctacatagctcag</td>
</tr>
<tr>
<td>PR1897</td>
<td>fragment with snip-SNP pkP601</td>
<td>ccaacacatctacatagctcag</td>
</tr>
<tr>
<td>PR1898</td>
<td>fragment with snip-SNP pkP601</td>
<td>tcagcagcttcttccagc</td>
</tr>
<tr>
<td>PR1899</td>
<td>fragment with snip-SNP pkP601</td>
<td>tcagcagcttcttccagc</td>
</tr>
<tr>
<td>PR1900</td>
<td>fragment with snip-SNP pkP601</td>
<td>agccgggtattacaagag</td>
</tr>
<tr>
<td>PR1901</td>
<td>fragment with snip-SNP pkP601</td>
<td>cctccgagcgccttag</td>
</tr>
<tr>
<td>PR1902</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1903</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1904</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1905</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1906</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1907</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1908</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1909</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1910</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1911</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1912</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1913</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1914</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1945</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1946</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1947</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1948</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1949</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1950</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
<tr>
<td>PR1951</td>
<td>fragment with snip-SNP pkP601</td>
<td>aagacgcgagcgccttag</td>
</tr>
</tbody>
</table>

**TABLE A.1. Primer names, sequences, and description of PCR products**

Continued on next page…
TABLE A.1. continued.

| PR1952 | *phat-*promoter; | ggagatagtgacgacgatgatgag |
| PR1953 | *phat*::GFP fusion | catctttgtgtgtgacccagaag |
| PR1954 | **F11E6**promoter | tcaatgtttttttctctgtgctc |
| PR1955 | **F08F3**promoter | cattctctcggctcctcattactgtg |
| PR1956 | **F08F3**promoter | cattttttgtgtgacacggaagca |
| PR1957 | fusion | ttatggatagatcacgaggctttc |
| PR1958 | lact-promoter | catcttttttatgcaagacggagaag |
| PR1959 | lact-promoter | cattttttttatgcaagacggagaag |
| PR1960 | T21H3promoter | atcagtttaaggcttccgcttct |
| PR1961 | C05C9promoter | catttttccaaacagtgatcttt |
| PR1962 | C05C9promoter | catttttccaaacagtgatcttt |
| PR1963 | H13N06promoter | gtgaaggctacggtgatgtctcc |
| PR1964 | H13N06promoter | cataggaagacacagaattctgaat |
| PR1965 | F11E6.8::GFP fusion | actagcggtacggtcttttgag |
| PR1966 | F11E6.8::GFP fusion | cattttttttatgcaagacggagaag |
| PR1967 | F11E6.8::GFP fusion | cattttttttatgcaagacggagaag |
| PR1968 | F18E3.1::GFP fusion | ttatgctttgtgcttcaaatgtc |
| PR1969 | F18E3.1::GFP fusion | cattttttttatgcaagacggagaag |
| PR1970 | gst-2promoter | cccattttttgaaaggaagctactg |
| PR1971 | gst-2promoter | cattttttttatgcaagacggagaag |
| PR1972 | F41D3promoter | gccacatgtaatccacacat |
| PR1973 | F41D3promoter | cattttttttatgcaagacggagaag |
| PR2028 | C41H0::GFP fusion | cccaaagcttccagcccccaactc |
| PR2029 | F11E6::GFP fusion | ggttaaggctagttcagatgatgc |
| PR2030 | lact-B::GFP fusion | tgaaacgtaaaacagaataag |
| PR2031 | T21H3::GFP fusion | ccgcccaacacttaagcccaag |
| PR2032 | C05C9::GFP fusion | tcgctattgtgacatagttc |
| PR2033 | H13N06::GFP fusion | ttgaaagttgtgtaaccaag |
| PR2034 | F32D8::GFP fusion | ggctagctgctgctcaat |
| PR2035 | gst-2::GFP fusion | ggaagttgagatagacagggag |
| PR2036 | F41D3::GFP fusion | gcatggcgttgacctgtag |
| PR2057 | C05C9cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2058 | C05C9cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2059 | F11E6cDNA | gcgcctagctgcatgcctctcaacactttgcat |
| PR2060 | F11E6cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2063 | gst-2cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2064 |gst-2cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2065 | phat-cDNA | gcgggtctgcatgcctctcaacactttgcat |
| PR2066 | phat-cDNA | gcgggtctgcatgcctctcaacactttgcat |
APPENDIX B

C. ELEGANS GENETIC STRAINS
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Transgene description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL5717</td>
<td>inIs179(II); him-8(e1490)</td>
<td>ida-1::gt</td>
</tr>
<tr>
<td>CB4856</td>
<td>wild type-Hawaiin isolate</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM549</td>
<td>unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM551</td>
<td>lin-48(sa469) unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM553</td>
<td>lin-44(e1792); unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM554</td>
<td>lin-17(n671); unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM598</td>
<td>lin-48(gv4) unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM654</td>
<td>unc-119(e2498); egl-38(sy294); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM748</td>
<td>unc-119(e2498); guIs21</td>
<td>hsp-41::egl-5</td>
</tr>
<tr>
<td>CM836</td>
<td>lin-48(sa469); unc-119(e2498); egl-38(sy294); him-guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM837</td>
<td>lin-48(gv4) unc-119(e2498); egl-38(sy294); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM973</td>
<td>egl-5(n486) unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM1131</td>
<td>lin-17(n671); egl-5(n486) unc-119(e2498); him-5(e1490); guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM1132</td>
<td>guEx356</td>
<td>vab-3::gt</td>
</tr>
<tr>
<td>CM1179</td>
<td>him-5(e1490); unc-78(e1217) lin-18(e620); lon-2(e678)</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM1294</td>
<td>him-5(e1490); lin-2(e1309)</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM1421</td>
<td>egl-38(sy294)</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM1452</td>
<td>egl-38(n578)</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM1453</td>
<td>inIs179(II); him-5(e1490); lin-14(sa485)</td>
<td>ida-1::gt</td>
</tr>
<tr>
<td>CM1455</td>
<td>him-5(e1490); dpy-6(e14) lin-14(sa485)</td>
<td>n.a.</td>
</tr>
<tr>
<td>CM1466</td>
<td>unc-119(e2498); guEx1035</td>
<td>H19N07op::gt</td>
</tr>
<tr>
<td>CM1468</td>
<td>unc-119(e2498); egl-38(n578); guEx1035</td>
<td>H19N07op::gt</td>
</tr>
<tr>
<td>CM1469</td>
<td>unc-119(e2498); guEx1037</td>
<td>lact-8::gt</td>
</tr>
</tbody>
</table>

Table B.1. Strain names, genotypes, and description of relevant transgenes

Continued on next page...
| CM1471 | unc-119(e2498); egl-38(n578); guEx1037 | lact-8::gf |
| CM1472 | unc-119(e2498); guEx1039 | F18E3.11::g |
| CM1474 | unc-119(e2498); egl-38(n578); guEx1039 | F18E3.11::g |
| CM1500 | him-5(e1490); lin-14(sa485) | n.a. |
| CM1501 | unc-119(e2498); egl-38(n578)/nT1 | n.a. |
| CM1502 | unc-119(e2498); guEx1050 | C05C9.1::g |
| CM1503 | unc-119(e2498); egl-38(n578); guEx1050 | C05C9.1::g |
| CM1504 | unc-119(e2498); guEx1051 | F11E6.8::g |
| CM1505 | unc-119(e2498); egl-38(n578); guEx1051 | F11E6.8::g |
| CM1507 | unc-119(e2498); egl-38(n578); guEx1052 | H13N06.2::g |
| CM1508 | unc-119(e2498); guEx1053 | phat-3::gf |
| CM1509 | unc-119(e2498); egl-38(n578); guEx1053 | phat-3::gf |
| CM1510 | unc-119(e2498); guEx1054 | F08F3.4::g |
| CM1512 | unc-119(e2498); egl-38(n578); guEx1054 | F08F3.4::g |
| CM1515 | unc-119(e2498); guEx1057 | gst-21::gf |
| CM1516 | unc-119(e2498); egl-38(n578); guEx1057 | gst-21::gf |
| CM1517 | unc-119(e2498); guEx1058 | F41D3.3::g |
| CM1518 | unc-119(e2498); egl-38(n578); guEx1058 | F41D3.3::g |
| DR441 | lin-14(n179ts) | n.a. |
| DR466 | him-5(e1490) | n.a. |
| JK987 | tra-2(q276)/mnC1 | n.a. |
| JT8174 | him-5(e1490); lin-14(sa485) | n.a. |
| MT3751 | dpy-5(e61); rol-6(e187); unc-32(e189) | n.a. |
| MT464 | unc-5(e53); dpy-11(e224); lon-2(e678) | n.a. |
| N2 | wild type-Bristol isolate | n.a. |
| RH10 | unc-119(e2498) | n.a. |
| SP224 | +/-mnDp1(V;X); let-4(mn105) unc-3(e151) | n.a. |
| SP392 | +/-mnDp1(V;X); unc-3(e151) let-41(mn146) | n.a. |
| SP928 | dpy-6(e14) unc-9(e101) | n.a. |
APPENDIX C

MALE PHENOTYPES OF LIN-14(SA485)
FIGURE C.1. ABNORMAL B CELL DIVISION AND SPICULE FORMATION IN LIN-14(sa485) MALES

(A) lin-14(sa485) males exhibit a male abnormal (Mab) phenotype. Specifically, these mutants show a severe reduction in the production of spicule material. This phenotype was the basis of the genetic screen in which lin-14(sa485) was isolated. (B-C) Nomarski images of the tail region of L2 males. B cell daughters are labeled and outlined (dotted black lines). (B) In wild-type males, initial division of the B cell is asymmetric, producing a larger anterior daughter, B.a, which goes on to produce spicule-secreting socket cells. (C) In lin-14(sa485) males, the asymmetric cytokinesis of B is frequently abnormal, yielding daughter cells of equivalent size, or even a larger posterior daughter cell. (D) Quantification of the asymmetric cytokinesis defect and expression of the B.a cell fate marker vab-3::gfp in lin-14(sa485) mutants.
APPENDIX D

M. NEMATOPHILUM INFECTION DATA
Figure D.1. *EGL-38 MUTANTS SHOW REDUCED INFECTION BY M. NEMATOPHILUM*

(A-L) Nomarksi and fluorescent images of rectums from hermaphrodites grown on *E. coli* OP50 or MBL plates. (A,D,G,J; first column) Light micrographs of uninfected animals. (B,E,H,K; second column) Light micrographs of animals plated on MBL plates. (C,F,I,L; third column) Fluorescent images revealing bacterial SYTO13 staining. These images are from the same animals shown in (B,E,H,K). Genotypes pictured are: (A-C) wild type; (D-F) *egl-38(n578)*; (G-I) *egl-38(sy294)*; (J-L) *egl-38(sy287)*. (M) Graph depicting proportions of a given population that display altered *M. nematophilum* infection levels. Wild-type animals typically have a basal/moderate level of rectal infection, and few animals show reduced or enhanced SYTO13 staining. Three reduced *egl-38* function conditions exhibit significant proportions of animals with reduced levels of bacterial infection.
Figure D.1. continued.
LIST OF REFERENCES


Hristova, M., Birse, D., Hong, Y., & Ambros, V. (2005). The caenorhabditis elegans heterochronic regulator LIN-14 is a novel transcription factor that controls the developmental timing of transcription from the insulin/insulin-like growth factor gene ins-33 by direct DNA binding. Molecular and Cellular Biology, 25(24), 11059-11072.


155


