IDENTIFICATION AND CHARACTERIZATION OF FACTORS FUNCTIONING WITH EGL-38 PAX TO REGULATE LIN-48 IN CAENORHABDITIS ELEGANS

DISSERTATION

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

By
Rong-Jeng Tseng, M.S.

*****

The Ohio State University
2008

Dissertation Committee: Approved by
Dr. Amanda A. Simcox, Adviser
Dr. David M. Bisaro
Dr. Harald Vaessin
Dr. Mark A. Seeger

The Ohio State University
2008

Dissertation Committee: Approved by
Dr. Amanda A. Simcox, Adviser
Dr. David M. Bisaro
Dr. Harald Vaessin
Dr. Mark A. Seeger

Graduate Program in Molecular, Cellular, and Developmental Biology
ABSTRACT

*Pax* genes are important in organogenesis during animal development. These genes affect the development of a range of different cell types, and promote cell proliferation and cell survival. Different functions are mediated through the regulation of different target genes, suggesting that cellular context plays an important role in Pax protein function. Pax proteins may function in a combinatorial manner with other proteins acting through separate enhancers. In *C. elegans*, genetic studies have shown that the *Pax2/5/8* gene *egl-38* functions in the development of several different cells and organ types including the egg-laying system, the hindgut, and the development of male mating spicules.

*lin-48* is a direct target for EGL-38 in hindgut cells. To identify additional genes important for *lin-48* expression, I have performed a genetic screen for altered *lin-48::gfp* expression pattern and identified mutant candidates that might mediate the response to EGL-38. Two major categories of mutant candidates were isolated. One category of mutants has reduced *lin-48::gfp* expression in the hindgut, which is a phenocopy of *egl-38*. The other category showed enhanced and ectopic GFP expression. The two categories represent candidates for positive and negative regulatory factors, respectively.
Most mutants in both categories have pleiotropic mutant phenotypes but exhibit moderate
to low penetrance for all defects. In addition, the newly isolated *egl-38*(gu22) allele was
compared to previously described alleles to extend the analysis of each allele for a range
of functions. Each *egl-38* allele disrupts certain functions of *egl-38* but they cannot be
ordered into an allelic series. Instead, the alleles exhibit cell or tissue preferential defects
compared to each other.

In many organisms, repetitive DNA serves as a trigger for gene silencing. However, some gene expression is observed from repetitive genomic regions such as heterochromatin, suggesting mechanisms exist to modulate the silencing effects. In this work, I identified mutations in two genes important for expression of repetitive sequences: *lex-1* and *tam-1*. Here I show that *lex-1* encodes a protein containing an ATPase domain and a bromodomain. LEX-1 is similar to the yeast Yta7 protein, which maintains boundaries between silenced and active chromatin. *tam-1* has previously been shown to encode a RING finger/B-box protein that modulates gene expression from repetitive DNA. I found that *lex-1*, like *tam-1*, acts as a class B synthetic multivulva (synMuv) gene. However, since *lex-1* and *tam-1* mutants have normal P granule localization, it suggests they act through a mechanism distinct from other class B synMuvs. I observed intragenic (interallelic) complementation with *lex-1* and a genetic interaction between *lex-1* and *tam-1*, data consistent with the idea that the gene products function in the same biological process, perhaps as part of a protein complex. I propose that LEX-1 and TAM-1 function together to influence chromatin structure and to promote expression from repetitive sequences.
Dedicated to my parents
ACKNOWLEDGMENTS

I wish to thank my adviser, Dr. Helen M. Chamberlin, for her guidance and support all through my graduate study.

I thank my committee members, Dr. Amanda A. Simcox, Dr. David M. Bisaro, Dr. Harald Vaessin, and Dr. Mark A. Seeger, for their valuable suggestions and willingness to spend time with me during their tight schedule.

I am grateful to fellow lab members, Dr. Xiaodong Wang, Dr. DonHa Park, Guojuan Zhang, Vandana Rajakumar, Sama Sleiman, Ryan Johnson, Hongtao Jia, Kristin Armstrong, and Xin Li, for their friendship, support and making the lab experience a fun and enjoyable one.

Finally, I would like to thank my family, especially my parents, for their support, patience, and love.
VITA

1993 .......................................................... B.S. Biology
National Cheng Kung University
Tainan, Taiwan

1995 .......................................................... M.S. Genetics
National Yang Ming University
Taipei, Taiwan

1997 - 1999 ............................................... Research Assistant
Division of Molecular and Genomic Medicine
National Health Research Institutes
Taipei, Taiwan

1999 – 2006 .............................................. Graduate Teaching and Research Associate
The Ohio State University
Columbus, Ohio, USA

PUBLICATIONS


FIELD OF STUDY

Major field: Molecular, Cellular, and Developmental Biology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td>ii</td>
</tr>
<tr>
<td><strong>Dedication</strong></td>
<td>iv</td>
</tr>
<tr>
<td><strong>Acknowledgments</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>Vita</strong></td>
<td>vi</td>
</tr>
<tr>
<td><strong>List of Tables</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>List of Figures</strong></td>
<td>xi</td>
</tr>
</tbody>
</table>

## Chapters:

1. **Introduction**                                                     | 1    |

2. **A genetic screen for factors affecting lin-48 expression**         | 15   |
   2.1. **Introduction**                                                 | 15   |
   2.2. **Materials and Methods**                                        | 19   |
   2.3. **Results**                                                      | 24   |
   2.4. **Discussion**                                                   | 35   |

3. **gu22 is a new allele of egl-38**                                    | 59   |
   3.1. **Introduction**                                                 | 59   |
   3.2. **Materials and Methods**                                        | 62   |
   3.3. **Results**                                                      | 63   |
   3.4. **Discussion**                                                   | 68   |
4. The bromodomain protein LEX-1 acts with TAM-1 to modulate gene expression in \textit{C. elegans} ................................................................. 75

4.1. Abstract ................................................................. 75
4.2. Introduction .......................................................... 77
4.3. Materials and Methods ............................................ 81
4.4. Results ................................................................. 86
4.5. Discussion ............................................................ 92

5. Conclusion ............................................................... 115

List of References .......................................................... 124
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Penetrance of phenotypes in isolated mutant strains</td>
<td>39</td>
</tr>
<tr>
<td>2.2</td>
<td>Complementation test</td>
<td>40</td>
</tr>
<tr>
<td>2.3</td>
<td>Complementation test</td>
<td>41</td>
</tr>
<tr>
<td>2.4</td>
<td>Pleiotropic effects observed in mutants</td>
<td>42</td>
</tr>
<tr>
<td>2.5</td>
<td>Lethality associated with mutants</td>
<td>43</td>
</tr>
<tr>
<td>2.6</td>
<td>Category 1 mutants affect <em>lin-48</em> activity in male tail development</td>
<td>44</td>
</tr>
<tr>
<td>3.1</td>
<td><em>egl-38</em> alleles differentially affect expression of <em>lin-48</em></td>
<td>70</td>
</tr>
<tr>
<td>3.2</td>
<td><em>egl-38</em> alleles preferentially affect different functions in vivo</td>
<td>71</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of genes recovered in a screen for altered <em>lin-48::gfp</em> expression</td>
<td>99</td>
</tr>
<tr>
<td>4.2</td>
<td>Genetic mapping of <em>lex-1</em> and <em>tam-1</em></td>
<td>100</td>
</tr>
<tr>
<td>4.3</td>
<td><em>lex-1</em> is a class B synMuv gene</td>
<td>101</td>
</tr>
<tr>
<td>4.4</td>
<td>Alleles of <em>lex-1</em> exhibit intragenic complementation and allele-specific</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>intergenic interference with <em>tam-1</em></td>
<td></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><em>lin-48::gfp</em> expression pattern in <em>C. elegans</em></td>
<td>45</td>
</tr>
<tr>
<td>2.2</td>
<td>F2 genetic screen using EMS</td>
<td>46</td>
</tr>
<tr>
<td>2.3</td>
<td>Pleiotropics associated with mutants</td>
<td>47</td>
</tr>
<tr>
<td>2.4</td>
<td><em>lin-48::gfp</em> expression is reduced in Category 1 mutants</td>
<td>48</td>
</tr>
<tr>
<td>2.5</td>
<td>Hindgut and male tail defects associated with Category 1 mutants</td>
<td>51</td>
</tr>
<tr>
<td>2.6</td>
<td>Semi-quantitative RT-PCR of endogenous <em>lin-48</em></td>
<td>52</td>
</tr>
<tr>
<td>2.7</td>
<td>Altered <em>lin-48::gfp</em> expression in the hindgut of Category 2 mutants</td>
<td>53</td>
</tr>
<tr>
<td>2.8</td>
<td>Ectopic neuronal extension from the hindgut</td>
<td>54</td>
</tr>
<tr>
<td>2.9</td>
<td>Altered <em>lin-48::gfp</em> expression in excretory duct</td>
<td>55</td>
</tr>
<tr>
<td>2.10</td>
<td>Ectopic <em>lin-48::gfp</em> expression in vulva</td>
<td>56</td>
</tr>
<tr>
<td>2.11</td>
<td>Multi-vulva phenotype in <em>gu36</em></td>
<td>57</td>
</tr>
<tr>
<td>2.12</td>
<td>Altered growth of <em>lin-48::gfp</em> expressing cell</td>
<td>58</td>
</tr>
<tr>
<td>3.1</td>
<td>Structure of a paired domain bound to DNA</td>
<td>72</td>
</tr>
<tr>
<td>3.2</td>
<td>Sequence of the paired domain from EGL-38 and other PAX proteins</td>
<td>73</td>
</tr>
<tr>
<td>3.3</td>
<td>Gene organization for <em>egl-38</em></td>
<td>74</td>
</tr>
<tr>
<td>4.1</td>
<td><em>lin-48::gfp</em> expression intensity is altered in <em>tam-1</em> and <em>lex-1</em> mutants</td>
<td>103</td>
</tr>
</tbody>
</table>
4.2  *nac-2::gfp* expression intensity is reduced in *tam-1* and *lex-1* mutants …… 104

4.3  *tam-1* alleles isolated in the genetic screen ........................................... 106

4.4  Molecular cloning of *lex-1* ........................................................................... 107

4.5  Anti-PGL-1 staining detects P granule localization in wild type
    and mutants ....................................................................................................... 110

4.6  Amino acid sequence alignment of LEX-1-related proteins ...................... 111
CHAPTER 1

INTRODUCTION

1.1 Gene regulation in organ development

Development of an organism depends upon finely-tuned and accurate control of mRNA transcription. Appropriate regulation of mRNA transcription is central to the differentiation and functions of eukaryotic cells, and to the development of complex organisms. Gene transcription requires the combined activity of the RNA polymerase machinery, protein complexes that act to modulate the chromatin environment, and sequence-specific DNA binding proteins that mediate selective transcriptional repression or activation. mRNAs are synthesized by the coordinated action of a set of general transcription and mRNA modification factors. These factors and the fundamental mechanisms involved in transcription are conserved among eukaryotes, including *C. elegans*.

Eukaryotic mRNAs are synthesized by Pol II through an intricate multistep process (Lemon and Tjian, 2000; Orphanides and Reinberg, 2002). mRNA transcription initiates at a fixed start site that is located adjacent to the promoter. Transcription then must be maintained during an elongation phase, in which Pol II progresses more distally...
along the gene. Remarkably, these events are mechanistically coupled to the series of steps that process the transcript into a mature mRNA. Studies in various systems have revealed that this apparatus is not controlled through a simple on/off "switch" at the promoter, and that the factors and mechanisms involved in transcription are instead subject to regulation at a surprising number of different levels.

Interactions between nuclear proteins are involved in every step of transcriptional regulation, including selection of target genes, regulation of DNA binding ability, regulation of transcriptional activity and turnover of transcription factors. Functional transcription is often dependent upon the selection of suboptimal nucleotide sequence(s) present in the target gene. The lineage-specific and inducible expression of eukaryotic genes is controlled by assembly of multipartite complex of transcription factors on regulatory regions composed of sequence-specific protein binding sites.

During development many thousands of genes are expressed to control patterning of the developing embryo. In the early phase of development, very rapid cell proliferation and differentiation occurs, but this rapid growth is under the control of a host of developmental genes. Relatively little is yet known about the processes that control patterning and how the precisely regulated number of cells required to form the complete organism is determined. Several gene families control patterning of specific organs, and regulate the decision-making process of cell life and death during development. One such family, called Pax genes (Dahl et al., 1997), derives its name from a highly conserved DNA-binding domain called the paired domain, a conserved 128 amino acid domain in the amino-terminal portion of the protein (Treisman et al., 1991). In this study, I used
*Caenorhabditis elegans* (*C. elegans*) as a model organism to study how a given Pax transcription factor participates in the development of different organs through regulating target genes in different cell types.

### 1.2 Pax genes and animal development

*Pax* genes are important in organogenesis during animal development. These genes affect the development of a range of different cell types, and promote cell proliferation and cell survival (reviewed by Chi and Epstein, 2002). *Pax* genes have been identified in a wide variety of species, including jellyfish (Sun 1997), the worm *C. elegans* (Chamberlin et al., 1997), insects such as *Drosophila* (Quiring et al., 1994), zebrafish (Krauss et al., 1991), chickens (Nohno et al., 1993), and many species of mammal (Noll 1993). Nine *Pax* genes, *Pax1*-*Pax9*, have been characterized in humans and mice, which exhibit highly restricted temporal and spatial expression patterns (Dahl et al., 1997).

Based on sequence similarities within the paired domain, as well as the presence or absence of additional motifs (such as the homeodomain or octapeptide), PAX proteins can be subdivided into four subfamilies (Chalepakis et al., 1993; Noll, 1993; Dahl et al., 1997). For example, the *Pax2*, *Pax5* and *Pax8* genes are organized together (Dahl et al., 1997) because their primary sequence is very similar, and each encodes a paired domain, an octapeptide, and a partial homeodomain (Ward et al., 1994). The paired domain and homeodomains encode DNA binding domains within the PAX proteins. An additional
domain in the *Pax* genes is the transactivation domain within the carboxyl terminus of each PAX protein, which is a serine- and threonine-rich domain responsible for transcriptional activation of target genes (Ward et al., 1994). Thus, each PAX protein is able to act as a transcription factor regulating the expression of a range of downstream genes (Mansouri et al., 1996).

PAX proteins function as transcription factors and play an essential role in embryonic pattern formation, cell proliferation and cell differentiation (Epstein et al., 1994; Dahl et al., 1997; Dohrman et al., 2000; Chi and Epstein, 2002). Mutations in PAX genes cause significant developmental abnormalities in a broad spectrum of organisms from flies to humans (Dahl et al., 1997; Chi and Epstein, 2002). Examples of target organs and tissues for the expression of PAX proteins during organogenesis are the skeleton (PAX1 and 9), central nervous system (PAX2, 3, 5, 6, 7, and 8) kidney (PAX2 and 8), B-cells (PAX5), thyroid (PAX8), pancreas (PAX4 and 6) and skeletal muscle (PAX3 and 7). Mouse transgenic models that have mutations disrupting the expression of these specific *Pax* genes exhibit abnormal development or agenesis of the corresponding target tissues (reviewed by Chi and Epstein, 2002; Lang et al., 2007). The analysis of mouse mutants has revealed their crucial role in the formation of a variety of tissues. In particular, they are involved in the regulation of early steps in organ development. They act to define the regional specification of distinct germ layers.

Recent studies have also discovered a role for PAX proteins in tissue-specific stem cell or progenitor cell populations, including melanocytes, muscle, and B-cells. Although expression of *Pax* genes is critical for cell survival during normal animal
development, constitutive expression or ectopic expression can lead to cancer (reviewed by Robson et al., 2006; Lang et al., 2007). This is supported by the fact that expression of PAX proteins is deregulated by different mechanisms in several tumor types including Wilms' tumors, rhabdomyosarcomas, brain tumors and lymphomas. An emerging hypothesis is that PAX proteins play an essential role in maintaining tissue specific stem cells by inhibiting terminal differentiation and apoptosis and that these functional characteristics may facilitate the development and progression of specific cancers. Understanding the normal developmental pathways regulated by PAX proteins may shed light on potentially parallel pathways shared in tumors, and ultimately result in defining new molecular targets and signaling pathways for the development of novel anti-cancer therapies.

Genetic studies have shown that each Pax protein acts in the development of several different cell and organ types (reviewed by Chi and Epstein, 2002). Different functions are mediated through the regulation of different target genes, suggesting that cellular context plays an important role in Pax protein function. Thus other factors may allow the same Pax protein to regulate one set of target genes in one cell type, and a different set of targets in a different cell type. Pax proteins may function in a combinatorial manner with other proteins acting through separate enhancers (e.g. Brophy et al., 2003). In addition, examples in which Pax proteins form complexes or bind DNA synergistically with other transcription factors have also been described (Fitzsimmons et al., 1996; Roberts et al., 2001; Jin et al., 2002; Miranda et al., 2002; Sigvardsson et al., 2002; Di Palma et al., 2003). While these studies suggest possible mechanisms for combinatorial control, dissecting how a given Pax factor participates in target gene
expression in different cell types in vivo remains an important question.

1.3 Pax genes in *C. elegans*

There are five *Pax* genes in *C. elegans*. Based on sequence similarities within the paired domain, they can be classified into four subfamilies as in other animals. The *Pax-6* gene *vab-3*, identified based on its involvement in head and tail development (Chamberlin and Sternberg, 1995; Chisholm and Horvitz, 1995; Zhang and Emmons, 1995). Another gene, *egl-38*, which was identified based on its role in uterine and tail development (Chamberlin et al., 1997), is a member of the *Pax-2/5/8* family. The *egl-38* gene was recently duplicated to yield the highly related, *pax-2* gene (Wang et al., 2004). *egl-38* null mutants die as young larvae, whereas *pax-2* null mutant is viable and fertile. Even though these proteins possess an overall identity of 75% and 100% identity in the Paired domain, the phenotypes associated with the *egl-38* mutant suggest that these genes act non-redundantly, possibly in different tissues. The remaining two *Pax* genes in *C. elegans* are a single member of the *Pax-1/9* family, K07C11.1, and a single member of the *Pax-3/7* family, F27E5.2; neither of these genes has yet been genetically analyzed.

1.4 The roles of *egl-38* and *lin-48* in organogenesis

To better understand how a single Pax transcription factor can regulate different target genes in different cells, our lab has characterized the functions of the
Caenorhabditis elegans Pax2/5/8 gene egl-38. Genetic studies have shown that egl-38 is an essential gene and egl-38 functions in the development of several different cells and organ types including the egg-laying system, the hindgut, and the development of male mating spicules (Chamberlin et al., 1997). The egl-38(s1775) allele disrupts a splice donor site within intron 1, the mutation strongly or completely eliminates egl-38 function. egl-38(s1775) mutants die as embryos or soon after hatching (Chamberlin et al., 1997). Reduction-of-function but non-null alleles of egl-38 have been identified, allowing its function to be assessed in a range of larval and adult tissues (Chamberlin et al., 1997; Chamberlin et al., 1999). These alleles are missense mutations (sy287, sy294, n578) of egl-38 and affect the paired domain coding sequence. These different non-null alleles preferentially disrupt different functions of EGL-38. For example, the egl-38(sy287) and egl-38(sy294) alleles preferentially disrupt development of the male tail and mutants fail to express lin-48 in hindgut cells. In contrast, the egl-38(n578) allele preferentially disrupts development of the egg-laying system, and disrupts hindgut development to a minimal extent.

Moreover, egl-38 also affects the expression of different genes in different cell types. For example, expression of the epidermal growth factor gene lin-3 in the hermaphrodite vulval cells is dependent on egl-38, whereas the Ovo zinc finger transcription factor gene lin-48 is a target for EGL-38 in hindgut cells (Chang et al., 1999; Johnson et al., 2001). Genetic analysis has shown that egl-38 and lin-48 affect the development of the same subset of hindgut cells, and act to make those cells different from other hindgut cells. However, lin-48 is not expressed in the egg-laying system, indicating it is a target in one cell type, but not the other. Molecular and biochemical
experiments indicate that \textit{lin-48} is a direct target for EGL-38 in the hindgut. In addition, two redundant elements in the \textit{lin-48} promoter are important for \textit{lin-48} expression in the hindgut cells, and mediate the response to EGL-38, although EGL-38 binds with high affinity to only one of the elements (Johnson et al, 2001).

\textit{lin-48} encodes an OVO-like C2H2 zinc finger protein and is required for the fate of specific cell types in \textit{C. elegans}. Previous studies in our lab have shown that \textit{lin-48} is important for normal development of the hindgut, the excretory duct, and the male spicule (Chamberlin et al, 1999; Wang and Chamberlin, 2002). GFP tagged LIN-48 is expressed in a subset of hindgut cells, the excretory duct cell, and some neuronal support cells in head and tail. Normal \textit{lin-48::gfp} expression in the excretory duct cell and a subset of hindgut cells requires the activity of EGL-38 Pax transcription factor (Johnson et al., 2001). However, \textit{lin-48::gfp} expression in the head and phasmid cells is EGL-38 independent, suggesting that other regulatory factors or pathways are important for \textit{lin-48} expression in other cells. Moreover, EGL-38 functions and is expressed in cells of the hindgut and excretory duct (Chamberlin et al, 1997; Zhang et al, 2005), but also other cells where \textit{lin-48} is not expressed, such as cells in the egg-laying system (Rajakumar and Chamberlin, 2007). How could EGL-38 regulate \textit{lin-48} in such a tissue-specific manner? We predict EGL-38 functions in combination with another factor to regulate \textit{lin-48} specifically in the hindgut and excretory duct. On the other hand, some negative regulatory mechanisms might prevent EGL-38 from activating \textit{lin-48} expression in cells where \textit{lin-48} is not expressed. In addition, we hypothesize that there might be additional factors that regulate \textit{lin-48} expression in cells where expression is not EGL-38-dependent. To test these hypotheses, I wanted to identify additional genes and factors that are
important for the regulation of *lin-48* expression. Thus, I carried out a genetic screen for mutations that altered *lin-48::gfp* expression (refer to Chapter 2). From this genetic screen in *C. elegans*, I have identified mutations of a new synMuv B gene: *lex-1* (refer to Chapter 4). LEX-1 is similar to the yeast Yta7 protein, which maintains boundaries between silenced and active chromatin. Data suggests that LEX-1 modulates chromatin to establish and maintain gene activity in the context of repetitive DNA sequence.

1.5 Chromatin regulation in development

Proper regulation of gene expression is of utmost importance for development and survival of all forms of life, as regulated transcription of complex genomes is necessary for cellular energy economy and the precise patterns of gene expression required for cellular differentiation. Transcription of genomes is regulated at multiple levels, which can involve large regions of chromosomes, individual nucleosomes, or direct effects on the transcription machinery at individual promoters (Kimura and Horikoshi 2004). The activities of RNA polymerase and sequence-specific DNA binding proteins are dependent on proper establishment and maintenance of chromatin structure since chromosomal DNA is packaged in chromatin.

Within the eukaryotic nucleus, DNA is packaged into chromatin, a periodic nucleoprotein complex consisting of repeating nucleosomes. Each nucleosome contains four core histone proteins, H2A, H2B, H3 and H4, whose flexible N-terminal tails are subject to non-random N-acetylation, a reversible process that is governed by the
opposing actions of histone acetyltransferase (HAT) and histone deacetylase (HDAC) (Roth et al., 2001; Grozinger and Schreiber, 2002). Together with other modifications, such as methylation and ubiquitination, histone acetylation forms part of the histone code, a complex modification language that is fundamental to the regulation of chromatin structure and function (Strahl and Allis, 2000; Turner, 2000).

At the level of chromosome structure, the chromatin is organized into regions with different degrees of accessibility to the transcriptional machinery. Chromatin is typically repressive for nuclear processes like transcription (Horn and Peterson, 2002). The more condensed or protected regions of chromatin are relatively inaccessible and so transcriptionally "silent," whereas "active" regions are more open and accessible (reviewed by Rusche et al., 2003; Vermaak et al., 2003). Acetylation loosens chromatin structure and thus correlates with gene activation. Specialized boundary zones have been found between certain silent and active regions; these zones prevent the invasion of either region into the other (reviewed by Bi and Broach, 2001; Labrador and Corces, 2002).

The silent, active and boundary regions all represent stably maintained and heritable epigenetic states of localized chromatin organization. The cell therefore requires machineries that establish, maintain, and ensure the faithful replication of all the epigenetic states on chromosomes. This appears to be accomplished through the interplay of different protein-DNA complexes formation and modification (Margueron, 2005). Strategies for establishing and maintaining the silent and active regions of chromatin include packaging the DNA with appropriately modified histones (Jenuwein and Allis, 2001; Vermaak et al., 2003), the use of energy-driven chromatin remodeling machinery
(Becker and Horz, 2002), and the addition of specific chromatin-associated proteins (Meneghini et al., 2003; Rusche et al., 2003; Mizuguchi et al., 2004).

1.6 Chromatin factors for gene regulation in *C. elegans*

In the past decade, many studies have shown the importance of chromatin modifications and chromatin remodeling in gene regulation (reviewed in Jenuwein and Allis, 2001). These include modification of histone tails by acetylation, deacetylation, and methylation, binding of factors to modified histones, and nucleosome remodelling. Chromatin boundary elements (Donze 2004; Donze and Kamakaka 2002; Kimura and Horikoshi 2004; West et al. 2002) are also involved in the regulation mechanism, as boundaries can either prevent a distal enhancer from activating gene expression inappropriately (insulators), or prevent the unrestricted spreading of heterochromatin (barriers). A large number of different complexes are known to have activities in modifying chromatin, but it is poorly understood how they function in a developmental context. The *Caenorhabditis elegans* vulva represents a simple developmental system in which to study the function and mechanism of chromatin regulation. The so-called synthetic multivulval (synMuv) genes, many of which encode chromatin-associated proteins, inhibit vulval development by antagonizing Ras signaling through an unknown mechanism (Fay and Han, 2000). Identification of the factors involved in this regulation and study of their functions should lead to an understanding of how these work together to bring about correct cellular development.
In general, synMuv genes prevent cells that do not receive inductive signal from adopting vulval fates. At the time of vulval induction, the LIN-3 EGF ligand is produced by the gonadal anchor cell and stimulates Ras pathway signaling in the neighboring cells P(5–7).p. The high levels of Ras pathway activity in P(5–7).p exceed a predetermined threshold that is required for the adoption of vulval cell fates. Because of their distance from the anchor cell, P3.p, P4.p, and P8.p receive little to no LIN-3 signal, and the unstimulated level of Ras pathway activity in these cells is insufficient to induce vulval fates.

Ras signaling during vulval development acts through the LIN-1 ETS and LIN-31 winged helix transcription factors (Miller et al., 1993; Beitel et al., 1995; Tan et al., 1998). It is unclear whether chromatin remodeling and modification by the synMuv genes impacts transcription by LIN-1 and LIN-31 or whether the synMuv and Ras pathways converge in a less direct manner. No matter how they converge, the synMuv genes likely act to limit the extent of Ras-mediated vulval induction. Chromatin remodeling and modification may be used as a general mechanism to limit inductive signaling in developing organisms.

One possible mechanism is that the synMuv genes may establish the threshold for cells to respond to the inductive signal. In synMuv mutants, this threshold is reduced in P(3–8).p to a level that can be exceeded even by unstimulated, that is, LIN-3-independent, Ras pathway activity. Alternatively, instead of increasing the threshold for a response to Ras pathway activity, the synMuv genes may decrease the absolute level of Ras pathway activity in P(3–8).p. In synMuv mutations, the repression of Ras pathway activity is
eliminated. In turn, the basal Ras pathway activity is boosted, such as that in P3.p, P4.p, and P8.p of wild-type animals or in P(3–8).p of *lin-3* mutants, to a level that exceeds an unchanged threshold for vulval fates.

### 1.7 Goal of this study

The goal of this dissertation is to understand how a single Pax transcription factor, EGL-38, can regulate different target genes in different cells. I hypothesize that EGL-38 functions in combination with other factors to archive the tissue-specific regulation. Give the fact that the expression profiles of *egl-38* and its target gene, *lin-48*, do not fully overlap, there might be additional factors that regulate *lin-48* expression in cells where expression is not EGL-38-dependent. To test these hypotheses, I wanted to identify additional genes and factors that are important for the regulation of *lin-48* expression. Thus, I carried out a genetic screen for mutations that altered *lin-48::gfp* expression.

In the following chapters, I present the genetic screen result as well as the characterization and identification of the isolated mutant alleles. Two major categories of mutant candidates were isolated as described in Chapter 2. The two categories represent candidates for positive and negative regulatory factors, respectively. In Chapter 3, I compared the newly isolated *egl-38(gu22)* allele to previously described alleles. To extend the analysis of each allele for a range of functions, the phenotypes associated with mutant animals were examined. Finally, two identified genes from the screen were characterized as described in Chapter 4. Data suggests that the bromodomain protein
LEX-1 and the RING finger protein TAM-1 function together to influence chromatin structure and to promote expression from repetitive sequences.
CHAPTER 2

A GENETIC SCREEN FOR FACTORS AFFECTING \textit{lin-48} EXPRESSION

Parts of the following chapter were originally published in:
Zhang, G., Sleiman, S. F., \textbf{Tseng, R. J.}, Rajakumar, V., Wang, X. and Chamberlin, H. M.

2.1 INTRODUCTION

\textit{Pax} genes are important in organogenesis during animal development. These genes affect the development of a range of different cell types, and promote cell proliferation and cell survival (reviewed by Chi and Epstein, 2002). \textit{Pax} genes encode transcription factors that are identified based on the presence of a conserved DNA binding domain called the paired domain. Although four main classes of Pax proteins have been identified, all Pax proteins share several features common to the paired domain. The paired domain is comprised of two helix-turn-helix domains which can interact with DNA across a relatively large stretch of DNA of more than 20 bp. The bipartite domains can bind to DNA independent of each other (Xu et al., 1995; Jun and Desplan, 1996; Kozmik et al., 1997; Jun et al., 1998; Xu et al., 1999a). Due to the relatively large size of the DNA contacted by the paired domain, and the altered binding properties as a
monomer or as part of a complex with other proteins, the consensus DNA sequence for Pax-responsive elements is notably degenerate (Czerny et al., 1993).

Genetic studies have shown that each Pax protein acts in the development of several different cell and organ types (reviewed by Chi and Epstein, 2002). Different functions are mediated through the regulation of different target genes, suggesting that cellular context plays an important role in Pax protein function. Thus other factors may allow the same Pax protein to regulate one set of target genes in one cell type, and a different set of targets in a different cell type. Pax proteins may function in a combinatorial manner with other proteins acting through separate enhancers (e.g. Brophy et al., 2003). In addition, examples in which Pax proteins form complexes or bind DNA synergistically with other transcription factors have also been described (Fitzsimmons et al., 1996; Roberts et al., 2001; Jin et al., 2002; Miranda et al., 2002; Sigvardsson et al., 2002; Di Palma et al., 2003). While these studies suggest possible mechanisms for combinatorial control, dissecting how a given Pax factor participates in target gene expression in different cell types in vivo remains an important question.

To better understand how a single Pax transcription factor can regulate different target genes in different cells, our lab has characterized the functions of the Caenorhabditis elegans Pax2/5/8 gene egl-38. Genetic studies have shown that egl-38 functions in the development of several different cells and organ types including the egg-laying system, the hindgut, and the development of male mating spicules (Chamberlin et al, 1997). egl-38 also affects the expression of different genes in the different cell types. For example, expression of the epidermal growth factor gene lin-3 in
the hermaphrodite vulval cells is dependent on egl-38, whereas the Ovo zinc finger transcription factor gene lin-48 is a target for EGL-38 in hindgut cells (Chang et al., 1999; Johnson et al., 2001). lin-48 is not expressed in the egg-laying system, indicating it is a target in one cell type, but not the other. Molecular and biochemical experiments indicate that lin-48 is a direct target for EGL-38 in the hindgut. In addition, two redundant elements in the lin-48 promoter are important for lin-48 expression in the hindgut cells, and mediate the response to EGL-38, although EGL-38 binds with high affinity to only one of the elements (Johnson et al, 2001).

lin-48 encodes an OVO-like zinc finger protein and is required for the fate of specific cell types in C. elegans. LIN-48 contains four C2H2 zinc-finger repeats. Within the 130 amino acids of the zinc-finger domain, LIN-48 is 73% identical to Drosophila OVO and 66% identical to mouse mOVO1 (Mevel-Ninio et al., 1991; Dai et al., 1998). OVO proteins in Drosophila and mouse act as transcription factors and play an important role in the development of several distinct cell types (Oliver et al., 1987; Payre et al., 1999; Dai et al., 1998). Drosophila OVO binds DNA in a sequence-specific manner, and can act as a transcriptional activator and as a repressor, depending on the isoform (Lu et al., 1998; Lee and Garfinkel, 2000; Andrews et al., 2000). Work with ovo genes in Drosophila and mouse has focused on their roles in fertility and epidermal development (Oliver et al., 1987; Payre et al., 1999; Dai et al., 1998). Although lin-48 plays no apparent role in fertility or development of the epidermis, ovo genes in mouse, Drosophila and C. elegans exhibit parallels in that they all play a role in the differentiation and maintenance of specific cell types. In addition, C. elegans and mouse ovo genes are similar in that they play a role in urogenital development. Mouse Ovo1 is
important in development of the genital tract and kidney (Dai et al., 1998), and \textit{lin-48} plays a role in development of the hindgut (which develops into the adult male cloaca) and the excretory system.

Previous studies in our lab have shown that \textit{lin-48} is important for normal development of the hindgut, the excretory duct, and the male spicule (Chamberlin et al, 1999; Wang and Chamberlin, 2002). GFP tagged LIN-48 is expressed in a subset of hindgut cells, the excretory duct cell, and some neuronal support cells in head and tail (Figure 2.1). Normal \textit{lin-48::gfp} expression in the excretory duct cell and a subset of hindgut cells requires the activity of EGL-38 Pax transcription factor (Johnson et al., 2001). However, \textit{lin-48::gfp} expression in the head and phasmid cells is EGL-38 independent, suggesting that other regulatory factors or pathways are important for \textit{lin-48} expression in other cells. Moreover, EGL-38 functions and is expressed in cells of the hindgut and excretory duct (Chamberlin et al, 1997; Zhang et al, 2005), but also other cells where \textit{lin-48} is not expressed, such as cells in the egg-laying system (Rajakumar and Chamberlin, 2007). How could EGL-38 regulate \textit{lin-48} in such a tissue-specific manner? We predict EGL-38 functions in combination with another factor to regulate \textit{lin-48} specifically in the hindgut and excretory duct. On the other hand, some negative regulatory mechanisms might prevent EGL-38 from activating \textit{lin-48} expression in cells where \textit{lin-48} is not expressed. In addition, we hypothesize that there might be additional factors that regulate \textit{lin-48} expression in cells where expression is not EGL-38-dependent. To test these hypotheses, I wanted to identify additional genes and factors that are important for the regulation of \textit{lin-48} expression. Thus, I carried out a genetic screen for mutations that altered \textit{lin-48::gfp} expression.
2.2 MATERIALS AND METHODS

2.2.1 Strains

Strains were grown and maintained as described (Brenner 1974; Sulston and Hodgkin, 1988). Unless stated otherwise, experiments were conducted at 20°C. Mutations used were listed in the following:

LG I:  
smg-1(r861), dpy-5(e61), lin-35(n745) 

LG II: 
rol-6(e187), lin-8(n111), lin-38(n751), trr-1(n3630) 

LG III: 
unc-93(e1500), lin-48(sa469), dpy-17(e164), unc-32(e189), 
unc-119(e2498) 

LG VI: 
dpy-13(e184), unc-5(e53), unc-24(e138), egl-38(sy294), dpy-20(e1282). 
sDf9, sDf65, sDf69, sDf82. 

LG V: 
unc-62(s472), unc-46(e177), dpy-11(e224), him-5(e1490) 

LG X: 
lon-2(e678), smg-2(e2008), lin-15B(n744), lin-15A(n767) 


2.2.2 Isolation of mutations

Mutations were isolated as part of a genetic screen for mutants with an altered lin-48::gfp expression pattern. The lin-48::gfp transgene contained in sals14 includes regulatory sequences, but not coding sequences for lin-48. However, previous work suggests that is serves as a reliable transcriptional reporter for the gene. L4
hermaphrodites from the strain CM3 (unc-119(e2498); him-5(e1490); saIs14) or CM117 (unc-119(e2498); sal-14) were mutagenized with 50 mM EMS for 4 hr (Sulston and Hodgkin 1988), placed individually on agar plates, and allowed to produce self-progeny for two generations.

18,000 mutagenized gametes were screened in this work (Figure 2.2). By using a dissecting microscope with epifluorescence, 661 F2 hermaphrodites with any altered GFP patterns were recovered and placed individually on a fresh plate and allowed to self-cross. Three hundred and seventy-six of the recovered hermaphrodites were able to produce progeny. Among these strains, 24 mutant lines were able to breed true.

### 2.2.3 Genetic mapping

All mutations were backcrossed at least three times to unmutagenized parental strains, CM117 or CM3. Mapping strains with markers for each chromosome, MT3751 [dpy-5(e61) I; rol-6(e187) II; unc-32(e189) III] and MT464 [unc-5(e53) IV; dpy-11(e224) V; lon-2(e678) X], were used to assign mutations to specific linkage groups. Mutant alleles (gu22, gu24, gu34, gu39, gu48) that exhibited linkage to unc-5(e53) and not to the other markers were assigned to linkage group IV. gu44 exhibited linkage to dpy-11(e224) and was assigned to linkage group V. Once assigned to a linkage group, the mutations were further mapped by 3-factor mapping, deficiency mapping, and SNP mapping (more detail in Chapter 4).
2.2.4 Complementation test

The mutations were tested first for dominance by crossing with the parental strains CM3 or CM117. Recessive mutant alleles that mapped to Linkage Group IV were tested for complementation with each other and with *egl-38(sy294)* which is also on Linkage Group IV. Mutant males with mating ability from each strain were crossed directly to hermaphrodites homozygous for the test allele and *dpy-11(e224)* on Linkage Group V. The *lin-48::gfp* expression of non-Dpy cross-progeny was scored by comparing the expression level to that of wild-type animals. Alleles involved in crosses in which greater than 30% of cross progeny exhibited a mutant phenotype were interpreted to fail to complement.

2.2.5 Analysis of *lin-48::gfp* expression

*lin-48::gfp* expression as assayed under the dissecting microscope at x100 magnification was subjectively quantified in adult animals on a four point scale, with 4 being wild-type expression level in all hindgut cells, 3, robust expression (but not quite wild-type), 2, significantly reduced expression level, and 1, no expression. For the data of Table 2.3, animals scored 3 or higher were classified as "wild-type", whereas animals scored 2 or lower were classified as "mutant". Cellular expression of *lin-48::gfp* (Figure 2.4) was assayed using Nomarski optics and epifluorescence on a Zeiss compound microscope at x1000 magnification as described (Johnson et al., 2001). Expression was subjectively quantified in larval animals as showing high, low, or absence of expression.
The difference between the numbers obtained for Table 2.3 and those for Figure 2.4 reflects the difference between the sensitivity and magnification of the dissecting microscope and the compound microscope. In each case wild-type animals are included as a control for interpretation within the assay type.

2.2.6 Egg-laying assay

Adult hermaphrodites were scored as egg-laying defective (Egl) if they retained more than the normal single row of eggs in the uterus and the retained embryos had developed past the gastrulation stage.

2.2.7 Male tail defect

Young adult males were observed using Nomarski optics at 1000x magnification. The morphology of their tail structures was compared to wild type, and the extent of lateral alae differentiation was scored according to the method of Kenyon (1986), using the landmarks of Wrischnik and Kenyon (1997). Differentiation of ectopic spicule socket cells in adult \textit{lin-48} mutant males was scored as previously described (Chamberlin et al., 1999). The male animals were also scored for distended intestinal lumen (Con) and mating efficiency.
2.2.8 Viability

Viability was tested at 20°C. L4 homozygous mutant hermaphrodites were placed singly on fresh plates and then transferred to fresh plates every 24 hr until they stopped producing self-progeny. Twenty-four hours after transferring the parent, unhatched eggs and dead L1 stage larvae were counted and removed from the plate. Forty-eight hours after removing the dead eggs, the remaining progeny were counted and assessed for stage of development. Because dying larvae disintegrate rapidly, dead larvae, especially L1s, may be underrepresented when counted by this method.

2.2.9 Hypersensitized assay for endogenous lin-48 activities

Males heterozygous for lin-48 and homozygous for lex-1 and tam-1 alleles were produced by crossing lin-48(sa469) unc-32(e189)/+ ; him-5(e1490); sals14/+ males to unc-93(e1500) dpy-17(e164); [lex-1 or tam-1] him-5(e1490); sals14 hermaphrodites. Non-Dpy non-Unc cross hermaphrodites were selected and allowed to self-cross. From F1 parents segregating the lin-48(sa469) unc-32(e189) chromosome, non-Dpy non-Unc Lex-1 or Tam-1 F2 hermaphrodites were selected and self-crossed. These animals were confirmed to be genotype lin-48(sa469) unc-32(e189)/unc-93(e1500) dpy-17(sa469); [lex-1 or tam-1] him-5(e1490); sals14 based on phenotypes observed among the F3 progeny. Non-Unc non-Dpy (heterozygous) F3 males were selected and examined at 1,000x magnification using Nomarski optics.
2.3 RESULTS

2.3.1 A genetic screen isolated three categories of mutant alleles affecting

*lin-48::gfp* expression pattern

To identify additional factors that influence the tissue-specific expression of *lin-48*, I performed an F2 genetic screen to isolate mutants with altered *lin-48::gfp* expression pattern. Following EMS mutagenesis, 24 mutant lines were isolated from a screen of 18,000 mutagenized gametes (Figure 2.2, Table 2.1). Based on the observed patterns of altered GFP, these mutants were classified into three major categories.

The first category contains 14 mutants with reduced GFP expression in hindgut cells compared to cells in the head, which is a phenocopy of *egl-38*. However, in contrast to *egl-38* mutants, most category 1 mutants exhibit only a moderate to weak reduction of GFP expression in the hindgut. In general, GFP expression is still visible in the hindgut cells of these mutants. In addition, some of the animals retain a wild-type GFP expression pattern. These observations indicate mutants of category 1 are not fully penetrant for the reduced GFP phenotype. Since EGL-38 is important for *lin-48* expression in the hindgut, the responsible genes in this class might include cofactors that function with *egl-38* in the hindgut, genes important for EGL-38 expression or activity, or *egl-38* itself with new mutations. However, genes regulating *lin-48* expression through other mechanisms are also possible.

In the second category, nine mutants have ectopic or altered GFP expression patterns. All nine mutations are pleiotropic (exhibiting altered expression in several
tissues, and exhibiting phenotypes in addition to altered *lin-48::gfp*). However, a moderate to very low percentage of animals show altered GFP expression in each strain. Thus, these strains exhibit low penetrance of the GFP phenotypes. The mutant genes in this second category might include negative regulators of *egl-38* or factors that repress *lin-48* expression through other pathways.

The third category of mutants contains only one mutant, *gu35*. *gu35* affects the growth of *lin-48* expressing cells in the head, which in turn alters the GFP pattern. The GFP expression level and the number of GFP expressing cells are not changed. I interpret that the responsible gene is not involved in the regulation of *lin-48* expression. However, this work indicates that *lin-48::gfp* might be a good marker for studying morphogenesis of the *lin-48* expressing cells, as well as the development defect associated with the mutants.

2.3.2 Category 1: Reduced *lin-48::gfp* expression in the hindgut

Genetic mapping experiments were initiated as a first step in the molecular identification of the genes affected in the mutant strains. By using mapping strains with markers on each chromosome, three category 1 mutants (*gu22, gu24, gu48*) were assigned to linkage group IV and one category 1 mutant (*gu44*) was assigned to linkage group V (Table 2.1). Other mutants were later assigned to chromosome IV or V by 2-factor and/or 3-factor mapping.
Since the mapping strains do not carry the integrated transgene \textit{lin-48::gfp}, some mutant homozygotes with heterozygous \textit{lin-48::gfp} were obtained during the mapping procedures. These animals contain one half dose of \textit{lin-48::gfp} compared to homozygotes, but I found that the GFP mutant phenotype is similar in both cases. This phenomenon was observed among all category 1 mutants and other crossing procedures. Thus, category 1 mutants are not sensitive to transgene copy number to exhibit the mutant GFP phenotype.

In addition to genetic mapping, the dominance of each mutant allele was tested. Hermaphrodites (\textit{unc-119(e2498); dpy-20(e1282); sals14}) with the same genotype as the parental strain CM117 but with \textit{dpy-20(e1282)} were crossed with males carrying each mutant allele, respectively. Since \textit{gu22} and \textit{gu47} males were not able to mate, mutant hermaphrodites with \textit{dpy-11(e224)} were crossed with parental strain CM117. The self-progeny will remain dumpy (\textit{Dpy}) while cross-progeny will not (Non-Dpy). \textit{lin-48::gfp} of all Non-Dpy cross progeny exhibited a wild-type expression pattern, indicating that all category 1 mutations are recessive. This result is consistent with the informal observations I made while crossing with mapping strains.

2.3.2.1 Three complementation groups in category 1

Category 1 mutants exhibit a moderate to low percentage of animals with weak defects (low penetrance), except for the \textit{gu22} mutant which has very high percentage of animals with strong reduced GFP phenotype (high penetrance). As \textit{egl-38} is important for
lin-48 expression in the hindgut, it is possible that some category 1 mutants are new egl-38 alleles. In addition, some mutants were mapped to chromosome IV, the same as the egl-38 locus. Since all category 1 mutants are recessive, the complementation test was applied to determine whether any of the new mutants contain alleles of egl-38. Complementation with egl-38 was tested by crossing mutant males with egl-38(sy294) dpy-20(e1282) hermaphrodites, and observing whether there were egl-38-phenotypes among the non-Dpy cross-progeny. A new egl-38 allele, gu22, was identified by this method, confirming that the screen can identify mutants with Egl-38 phenotype (see Chapter 3). Inter se complementation tests with the rest of the strains identified two more complementation groups (Table 2.2, Table 2.3). One complementation group consists of three alleles (gu24, gu47, gu48) and the other group consists of eight alleles (gu20, gu25, gu41, gu43, gu44, gu45, gu49, gu50).

Among the isolated category 1 mutants, egl-38(gu22) is the only allele to confer a high penetrance for the reduced lin-48::gfp expression phenotype. This defect reflects the fact that lin-48 expression in hindgut cells depends on EGL-38 activity. In contrast, the responsible genes of the two new complementation groups might simply modulate lin-48 expression.

2.3.2.2 Pleiotropies associated with mutants

Although lin-48 mutant males exhibit spicule morphological defects, most of the identified category 1 mutants have no significant morphological alteration. The exception
is *egl-38(gu22)* mutant animals that exhibit an abnormal spicule or protrusion of the hindgut phenotype (Figure 2.3). However, other pleiotropic phenotypes associate with some mutants in category 1. For instance, some mutants exhibit a low to modest level of egg laying defect, male mating defect, growth retardation, or lethality (Table 2.4, Table 2.5, and data not shown). In addition, fused rays, crumpled spicules, and abnormal male tail morphology were occasionally found in male *gu24* and *gu47* animals (Figure 2.3).

### 2.3.2.3 Mutants affect expression of *lin-48*

Although category 1 mutants exhibited reduced *lin-48::gfp* expression, many animals still exhibited detectable GFP in hindgut cells. To quantify these observations, I assessed GFP expression in mutants using the higher resolution of a compound microscope (Figure 2.4). Since preliminary observations using the dissecting microscope suggested that expression in the hindgut diminishes as the animals mature, data were collected for both L1 and L4 animals. I found that although most category 1 mutants exhibit decreased expression of *lin-48::gfp*, most hindgut cells retain a diminished, but detectable level of reporter gene expression. This is in contrast to mutants showing more specific *lin-48* regulatory defects, such as *egl-38* mutants (Johnson et al., 2001). In addition, *lin-48::gfp* expression in other cell types can be diminished. *gu24, gu47, gu48* mutant strains show a reduction in the number of head cells and the proportion of excretory duct cells and phasmid cells that express the reporter. The defect in *gu44* mutant animals is more specific to the hindgut cells.
Since the genetic screen utilized a reporter transgene \textit{lin-48::gfp}, the recovered mutants could be defective in expression of both endogenous and transgenic \textit{lin-48::gfp}, or only the transgene. In an attempt to avoid mutations in genes that primarily affect transgenes, we selected the mutants based on a relative reduction of transgene expression in hindgut cells compared to other cells. However, the analysis above suggested the potential for broader defects in the mutants. Mutants for genes such as \textit{tam-1} have often been isolated from GFP-reporter screens. However, \textit{tam-1} has been shown to influence silencing of transgene expression compared to wild type, and has only modest effects on the expression of endogenous genes (Hsieh et al., 1999). Thus we suspected that our mutants could be defective primarily in multi-copy transgene expression rather than the expression of endogenous \textit{lin-48}. To test whether category 1 mutants affect endogenous \textit{lin-48} expression, we reasoned that the mutants should exhibit defects associated with \textit{lin-48} mutants. We tested this possibility by characterizing the hindgut cells and the structures that develop from them in category 1 mutant males.

The expression of \textit{lin-48} in the hindgut cells is required for their normal development. In male animals, two \textit{lin-48}-expressing cells (U and F) normally divide to produce male-specific cells, and also provide a signal that patterns the development of spicule precursor cells (Sulston et al., 1980; Chamberlin et al., 1993). Consequently, in \textit{lin-48} mutant males, the spicules are abnormal. In addition, the mutant U cell can develop in a similar fashion to the normal spicule precursor, B, resulting in additional spicule cells that can be detected as an ectopic cluster of spicule socket cells (Chamberlin et al., 1999; Jiang and Sternberg, 1999). Homozygous category 1 mutant males except for \textit{gu22} mutants have phenotypically wild-type spicules (Figure 2.5 A) and, with the
exception of gu47, are capable of siring cross-progeny (Table 2.4). Thus the mutants do not share male tail phenotypes with lin-48 mutants. However, since the effect on the reporter transgene is to reduce but not eliminate expression, it is possible that endogenous lin-48 activity is reduced, but is sufficient for normal hindgut development. Consequently, I tested whether a mutation in category 1 enhances male defects in lin-48 heterozygotes, to provide a more sensitive assay for lin-48 activity. Normally, the lin-48(sa469) allele is recessive to the wild-type allele, so that heterozygous males are phenotypically wild type. However, I found that in a category 1 mutant background, heterozygous males can exhibit the lin-48-related defects of crumpled spicule including ectopic spicule cells (Table 2.6, Figure 2.5). Thus a role for category 1 genes in affecting endogenous lin-48 function can be detected in a sensitized genetic background. Overall, however, the effect on endogenous lin-48 gene activity is modest. I also find that mutants exhibit a constipated phenotype in which the intestine becomes abnormally distended with food (Figure 2.4 C), consistent with the idea that other features of hindgut development or function may be affected in category 1 mutants.

In order to more directly test whether category 1 mutants affect endogenous lin-48 expression, semi-quantitative RT-PCR was utilized (Figure 2.6). gu22 shows significant reduction of endogenous lin-48 which is consistent with its high penetrance of reduced lin-48::gfp expression phenotype. As gu47 mutant animals have modest but broad defect of lin-48::gfp expression, the endogenous lin-48 also slightly reduced in gu47 mutant. In contrast, gu44 has similar level of endogenous lin-48 as wild-type. This might results from the fact that gu44 mutant defect is modest and more specific to the hindgut. Thus, the effect on the overall level of endogenous lin-48 is not significant.
2.3.3 Category 2: Ectopic or altered \textit{lin-48::gfp} expression

Mutants in category 2 have ectopic \textit{lin-48::gfp} expression in the vulva or have altered \textit{lin-48::gfp} expression of hindgut or excretory duct (Table 2.1). As shown in the table, mutants in category 2 usually have pleiotropic defects. For instance, \textit{gu30}, \textit{gu34}, \textit{gu36}, and \textit{gu39} have all three mutant expression phenotypes, but with varying penetrance levels.

In wild-type animals, only four hindgut cells (U, F, K, K’) express \textit{lin-48::gfp}, an expression pattern that is visible as two bright GFP spots when animals are observed under a dissecting microscope (Figure 2.7 C). Altered GFP expression in the hindgut of mutants usually shows as an irregular pattern. Among category 2 mutants, \textit{gu21}, \textit{gu36}, and \textit{gu38} often exhibit more than four GFP-expressing cells in the hindgut (Figure 2.7 D). Moreover, some cells appear have higher GFP levels than wild type. Since the morphology of the hindgut cells is also changed in these mutants, it is difficult to recognize which hindgut cells are expressing \textit{lin-48::gfp}. In rare cases of \textit{gu36} animals, a ventral GFP line extending from the U cell region toward the middle of animal body is observed (Figure 2.8), suggesting that a tail neuron is inappropriately expressing the reporter gene.

Normal \textit{lin-48::gfp} expression in the excretory duct cell usually is present in the nucleus and in a short anterior branch along the duct (Figure 2.9 B). In mutant animals, especially \textit{gu34} mutants which have the strongest phenotype, the excretory duct cell shows a different \textit{lin-48::gfp} pattern. There are three to four long GFP expressing branches extending from the nucleus toward the posterior (Figure 2.9 D). In addition, the
altered lin-48::gfp expression in the excretory duct cell is brighter than wild type. It is possible that branch extension is altered in the mutant, resulting in alteration in the morphology of the excretory duct cell. Alternatively, the higher expression level might simply enhance the amount of GFP in the distal part of the excretory duct cell, making the distal branches easier to see.

In addition to the altered lin-48::gfp expression in the hindgut and the excretory duct where lin-48::gfp is normally expressed, six category 2 alleles (gu30, gu32, gu34, gu36, gu39, gu40) confer ectopic lin-48::gfp expression in the vulva (Figure 2.10). In general, all six alleles confer a very low penetrance for ectopic expression in vulva. Furthermore, the penetrance appears to fluctuate from generation to generation. As observed under a compound microscope, the ectopic lin-48::gfp is expressed at the top of the vulva, next to the uterus. It is possible that the expressing cells are vulF, but this has not been confirmed. Since EGL-38 is expressed in vulF, and is necessary for normal vulF development (Chamberlin et al., 1997; Rajakumar and Chamberlin, 2007), the mutant phenotype suggests category 2 mutants may disrupt genes that normally act to repress lin-48 expression in cells where EGL-38 is present. However, the broad range of pleiotropic defects suggest that the affected genes identified as category 2 mutants act in processes in addition to lin-48 regulation.

In addition to the ectopic or altered lin-48::gfp expression phenotypes, mutant animals of category 2 usually also exhibit some morphological or physiological defects. For example, the L4 vulva shown in Figure 2.10 A has an abnormal posterior half. Some animals also have a multivulva phenotype and egg-laying defect (Figure 2.11 A, B). In
general, mutant strains of category 2 exhibit moderate to low growth retardation and reduced viability. Irregular egg shapes and unhatched eggs are frequently observed among category 2 mutant strains (Figure 2.11 B).

2.3.4 Category 3: Altered growth of lin-48::gfp-expressing cells

In the head region, lin-48::gfp is expressed in the neuronal support cells of the labial sensory structures and in some additional unidentified cells. Observed under the epifluorescence microscope, the nuclei of these GFP-expressing cells surround the anterior bulb of the pharynx. Each neuronal support cell has a long extension from the cell body that ends at the tip of the animal’s nose. Usually two of the neuronal support cells have very bright lin-48::gfp expression and can be easily observed under the dissecting microscope (Figure 2.12 B).

There is only one mutant, gu35, in category 3. In gu35 mutant animals (Figure 2.12 F), the bright lin-48::gfp expression at the head is significantly shorter than that of wild-type animals. Moreover, this altered GFP phenotype is fully penetrant, which is different from all other mutant alleles isolated in this genetic screen. However, the expression level and the number of lin-48::gfp expressing cells is not changed. The morphology, mating efficiency, and viability of gu35 mutants are also normal.

To better understand the defect in gu35 mutants, I observed affected cells under the compound microscope. I found that the two brightest GFP-expressing neuronal support cells are shorter than the neuronal support cells in wild type, and that the cell
nuclei in the mutant animals are positioned more anterior than in wild type (Figure 2.12). Since the affected cells are shorter and closer to the anterior, it is possible that $gu35$ affects the cell extension or affects the nucleus position or cell migration. This indicates that the defect in $gu35$ mutants is not one of cell fate specification, but rather cell differentiation or morphogenesis.

To test the dominance of $gu35$, mutant homozygotes were crossed with wild-type animals. The cross progeny (heterozygotes) exhibited a GFP phenotype intermediate between $gu35$ homozygotes and wild-type animals (Figure 2.12 D). All three GFP patterns (long : intermediate : short) were observed among the self-progeny from these heterozygotes. Six self-progeny representing each GFP pattern were then placed individually in a new plate. Hermaphrodites with the intermediate GFP pattern segregated offspring exhibiting all three GFP patterns while hermaphrodites with the other two GFP patterns only segregated offspring with the same phenotype as the parent. Since the mutant phenotype in $gu35$ homozygotes is fully penetrant, this observation suggests that animals with the intermediate GFP pattern are heterozygous, and the $gu35$ mutation is semi-dominant. Consequently, it is possible that either the $gu35$ allele represents a gain-of-function allele of the affected gene, or that the locus is haploinsufficient.


2.4 DISCUSSION

Our lab’s characterization of *lin-48* indicates that EGL-38 has tissue-restricted targets that are expressed in only a subset of EGL-38-expressing cells (Johnson et al, 2001). One way EGL-38 may have different targets in different tissues is to act in a combinatorial manner with one or more additional transcription factors. In this model, both EGL-38 and the second factor would be necessary for the hindgut expression of *lin-48*. Consequently, if both EGL-38 and an additional factor are required, then the second factor must meet one of the following criteria. It could act through a DNA element. This raises the possibility that EGL-38 and the second factor would physically interact. Alternatively, the second factor may not act through a discrete site, but act in a manner different from EGL-38. For example, it might influence accessibility of the *lin-48* regulatory regions.

To identify additional genes important for *lin-48* expression, I have performed a genetic screen for altered *lin-48::gfp* expression pattern and identified mutant candidates that might mediate the response to EGL-38. This genetic screen was designed to recover mutations in genes that influence *lin-48* expression in a cell-specific manner, and did indeed recover a new allele of *egl-38*, gu22, confirming that the screen can identify mutants with Egl-38 phenotype.

In addition to the new *egl-38* allele, two major categories of mutant candidates were isolated. One category of mutants has reduced *lin-48::gfp* expression in the hindgut, which is a phenocopy of *egl-38*. The other category showed enhanced and ectopic GFP expression. The two categories represent candidates for positive and negative regulatory
factors, respectively. Most mutants in both categories have pleiotropic mutant phenotypes but exhibit moderate to low penetrance for all defects. Based on the mutant phenotypes, I interpret that the affected genes for both categories are likely modulators of lin-48 expression rather than genes that have an ‘on/off’ effect of lin-48 expression. Since EGL-38 is important for lin-48 expression in a tissue-restricted manner, the responsible genes might include cofactors that function with EGL-38, genes important for EGL-38 expression or activity. However, genes regulating lin-48 expression through other mechanisms are also possible, such as chromatin modifying enzymes or general transcriptional machinery.

2.4.1 Genetic and molecular characterization of category 1 mutants

Most category 1 mutants do not exhibit significant morphological alteration in the hindgut or in male structures that develop from hindgut cells. These tissues are essentially normal in category 1 mutants (data not shown). In contrast, other genes that affect endogenous lin-48 gene expression share lin-48 mutant phenotypes in the hindgut, excretory system, or both (Johnson et al. 2001; Wang et al. 2006; Zhang et al. 2005). One possible explanation for the observation is that the mutations are not null alleles. On the other hand, the mutant gene might normally function as a positive modulator for lin-48 expression. In both cases, lin-48 expression would be reduced, but the level would be higher than the threshold for normal functions. This possibility is consistent with the observation that category 1 mutants exhibit modest penetrance of the reduced lin-48::gfp expression phenotype, which does not completely diminish GFP in the hindgut cells. One
way I tested this possibility was by introducing the Category 1 mutations into a more sensitive genetic background, \textit{lin-48(sa469)/+} (Table 2.6). Although these results suggested that Category 1 mutations can affect hindgut cell development when \textit{lin-48} gene dose is reduced, it is possible that the mutations primarily alter transgene expression, not endogenous gene expression. This possibility was tested by semi-quantitative RT-PCR for endogenous \textit{lin-48} expression (Figure 2.6). Indeed, both results suggest that the alleles of category 1 were recovered based on their influence on multicopy transgene expression, and that they have a modest role in regulating the endogenous \textit{lin-48} gene product. The responsible genes of the two newly identified complementation groups will be discussed further in Chapter 4.

Previous experiments have identified two promoter elements important for \textit{lin-48} expression, and one of these (lre2) binds EGL-38 with high affinity. Genetic results indicate that both of these elements mediate the EGL-38 response. Specifically, both elements must be mutant to mimic the altered \textit{lin-48} expression pattern observed in \textit{egl-38} mutants, and single lre1 or lre2 mutant transgenes are equally sensitive to the \textit{egl-38(sy287)} and \textit{egl-38(n578)} mutant backgrounds (Johnson et al, 2001). I speculate that the screen failed to identify new genes that significantly affect \textit{lin-48} gene expression due to functional redundancy in \textit{lin-48} transcriptional regulation (Johnson et al. 2001; Wang and Chamberlin 2002) or to the essential roles of the trans-acting factors.
2.4.2 Enhanced *lin-48* expression in category 2 mutants might change cell fate determination or cell growth.

Mutants of category 2 exhibit pleiotropic defects, including ectopic and altered *lin-48::gfp* expression patterns, and morphological and physiological defects. Since category 2 mutants exhibit broadly enhanced *lin-48::gfp* expression which is not necessarily limited to *egl-38* functioning tissues, the responsible genes of category 2 might be involved in pathways which are not *egl-38*-specific or are not only involved in the *egl-38* pathway. Moreover, the broad range of pleiotropic defects suggests that the affected genes identified as category 2 mutants act in processes in addition to *lin-48* regulation. Because mutant strains exhibit some growth retardation and lethality, genes in category 2 might be essential.

Since most of the defects in Category 2 mutants exhibit low penetrance, one simple explanation is that these mutants are not null alleles. Considering that half of the mutant candidates originally selected in the screen died before giving any progeny, the null mutants of these genes might have not been recovered. Another explanation is that these mutated genes are negative modulators for *lin-48* expression. The normal role of these modulators might be to act as a suppressor which reduces the activity of factors for *lin-48* expression or reduces the accessibility of factors to the *lin-48* promoter region. Category 2 mutants may disrupt genes that normally act to repress *lin-48* expression in cells where EGL-38 is present. In this case the modulators just regulate factor activity, or the threshold for *lin-48* expression. Consequently, the mutant does not exhibit a dramatic or wide-spread increase in *lin-48* expression.
<table>
<thead>
<tr>
<th>Mutant Strain</th>
<th>Hindgut</th>
<th>Vulva</th>
<th>Excretory</th>
<th>Head</th>
<th>Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1: Reduced lin-48::gfp expression in the hindgut</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM129 (gu20)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM131 (gu22)</td>
<td>+++</td>
<td></td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM133 (gu24)</td>
<td>++</td>
<td></td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM134 (gu25)</td>
<td>+</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM159 (gu41)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM160 (gu42)</td>
<td>+</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM161 (gu43)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM162 (gu44)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM163 (gu45)</td>
<td>+</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM164 (gu46)</td>
<td>+</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM192 (gu47)</td>
<td>++</td>
<td></td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM193 (gu48)</td>
<td>++</td>
<td></td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM194 (gu49)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CM195 (gu50)</td>
<td>++</td>
<td></td>
<td></td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Category 2 : Ectopic or Altered lin-48::gfp expression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM130 (gu21)</td>
<td>++</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>CM146 (gu30)</td>
<td>+</td>
<td>+/-</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>CM148 (gu32)</td>
<td>+/-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM149 (gu33)</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>CM150 (gu34)</td>
<td>+</td>
<td>+/-</td>
<td>++</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM152 (gu36)</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>CM154 (gu38)</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM155 (gu39)</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>CM156 (gu40)</td>
<td>+/-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 3: Altered growth of lin-48::gfp expression cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM151 (gu35)</td>
<td></td>
<td></td>
<td>+++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Penetrance of phenotypes in isolated mutant strains. “+++”, “++”, “+” indicate the penetrance level: high, moderate, low, respectively. “+/-” indicates very low penetrance, less than 5% animals expressing the phenotype.
### Table 2.2: Complementation test.

“+” indicates the two alleles can complement each other. “-” indicates the two alleles fail to complement each other. The *egl-38* allele used in the complementation tests is *sy294*.

<table>
<thead>
<tr>
<th>Hermaphrodite</th>
<th>Male</th>
<th>gu20</th>
<th>gu22</th>
<th>gu24</th>
<th>gu44</th>
<th>gu48</th>
<th>gu50</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>egl-38</em></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>gu20</em></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><em>gu22</em></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>gu24</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>gu44</em></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><em>gu47</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>gu48</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.3: Complementation test.

The table shows percentage of wild-type *lin-48::gfp* expression in the indicated transheterozygous animals; number of animals scored is indicated in parentheses.
Table 2.4: Pleiotropic effects observed in mutants. The table shows percentage of animals with wild-type egg-laying function; number of animals scored is indicated in parentheses. Male mating ability is indicated as: “+” mate well, “+/-” difficult to mate, “-” unable to mate.
<table>
<thead>
<tr>
<th>Developmental Stages (%)</th>
<th>e/L1(died)</th>
<th>L1/L2/L3</th>
<th>L4/Adult</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>4.3</td>
<td>0.2</td>
<td>95.5</td>
<td>1615</td>
</tr>
<tr>
<td>gu20</td>
<td>37.7</td>
<td>47.8</td>
<td>14.4</td>
<td>324</td>
</tr>
<tr>
<td>gu44</td>
<td>10.4</td>
<td>0.4</td>
<td>89.2</td>
<td>537</td>
</tr>
</tbody>
</table>

**Table 2.5: Lethality associated with mutants.** Numbers indicate the percentage of animals that arrest at each of the indicated developmental stages. N indicates the number of animals assayed.
<table>
<thead>
<tr>
<th>Genotype</th>
<th>Abnormal spicule (%)</th>
<th>Ectopic spicule (%)</th>
<th>Defective (total; %)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>56</td>
</tr>
<tr>
<td>$\text{lex-1(gu24)}$</td>
<td>25</td>
<td>5</td>
<td>25</td>
<td>44</td>
</tr>
<tr>
<td>$\text{lin-48(+) }$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{lex-1(gu47)}$</td>
<td>23</td>
<td>14</td>
<td>23</td>
<td>57</td>
</tr>
<tr>
<td>$\text{lin-48(sa469)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{lex-1(gu48)}$</td>
<td>24</td>
<td>14</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>$\text{tam-1(gu44)}$</td>
<td>29</td>
<td>26</td>
<td>29</td>
<td>42</td>
</tr>
<tr>
<td>$\text{tam-1(cc567)}$</td>
<td>25</td>
<td>20</td>
<td>28</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Abnormal spicule (%)</th>
<th>Ectopic spicule (%)</th>
<th>Defective (total; %)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>77</td>
</tr>
<tr>
<td>$\text{lex-1(gu24)}$</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>66</td>
</tr>
<tr>
<td>$\text{lin-48(+) }$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{lex-1(gu47)}$</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>$\text{lin-48(+) }$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{lex-1(gu48)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53</td>
</tr>
<tr>
<td>$\text{tam-1(gu44)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>76</td>
</tr>
<tr>
<td>$\text{tam-1(cc567)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
</tbody>
</table>

**Table 2.6: Category 1 mutants affect *lin-48* activity in male tail development.**

Numbers indicate the percentage of animals that have defect in male tail. N indicates the number of animals assayed.
Figure 2.1: *lin-48::gfp* expression pattern in *C. elegans*. Arrowheads indicate the *lin-48::gfp* expressing cells (left to right): neuronal support cells, excretory duct, hindgut, phasmid.
**F2 Screen**

CM117 \([unc-119(e2498); saIs14(lin-48::gfp; unc-119)]\)

EMS mutagenesis (50 mM EMS, 20°C, 4 Hrs)

---

**Figure 2.2: F2 genetic screen using EMS.** \(lin-48::gfp\) animals were mutagenised using EMS to induce random mutations in the germ cells of P0 hermaphrodites. Since hermaphrodites are self-fertilizing, the sperm and eggs it produce will carry a copy of the mutation in the F1 generation. In the F2 generation, a quarter of animals from a mutation-bearing F1 will be homozygous for the mutation. Therefore, F2 generation animals were screened for presence of altered GFP expression patterns. 661 F2 animals with altered \(lin-48::gfp\) patterns were isolated from 18,000 haploid genomes. The animals with altered GFP thus obtained were self-crossed for one more generation to confirm the presence of the mutation. Only 376 of F2 animals were able to give progeny and 24 mutant lines were able to breed true.
Fig 2.3: Pleiotropics associated with mutants. Wild-type hermaphrodite (A). gu22 mutant hermaphrodite with protrusion of hindgut (arrowhead, B). Long and straight spicule in wild-type male (arrowhead, C). Abnormal male spicule of gu22 (arrowhead, D). Wild-type male tail (E). Fused ray of gu24 male tail (arrowhead, F). A wild-type male tail is tapered at the tip (G), whereas gu24 mutants can exhibit swelling at the tail tip (arrowhead, H).
Figure 2.4: *lin-48::gfp expression is reduced in Category 1 mutants*. Expression of the *lin-48::gfp* transgene *saIs14* was assayed for expression in the four hindgut cells (U, F, K, K'), the excretory duct cell, and head cells in wild type and mutants during L1 larval stage (A) and L4 larval stage (B). Hindgut cell and excretory duct cell expression is represented as percent of cells expressing normal levels of GFP (*black bar*), expressing very low, but detectable levels of GFP (*gray bar*) or not expressing GFP (*white bar*). Head cell expression is indicated as average number of expressing cells per animal. N indicates the number of animals assayed. (C) Category 1 mutants can be constipated. A *gu47* L4 hermaphrodite exhibits a distended intestine (*arrow, C2*) compared to wild-type L4 (C1). Defecation defect is represented as percent of animals with full distended intestine (*black bar*), one third to half of intestine is distended (*gray bar*) or no distended intestine (*white bar*).
Figure 2.4

A. Quantitative analysis of *lin-48::gfp* in L1 mutant animals.

<table>
<thead>
<tr>
<th></th>
<th>Expression in hindgut (percent of cells)</th>
<th>Expression in excretory duct (percent of cells)</th>
<th>Head cell expression (average no)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>7</td>
<td>54</td>
</tr>
<tr>
<td>gu24</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>gu47</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>5</td>
<td>38</td>
</tr>
<tr>
<td>gu48</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>gu44</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>7</td>
<td>31</td>
</tr>
</tbody>
</table>

B. Quantitative analysis of *lin-48::gfp* in L4 mutant animals.

<table>
<thead>
<tr>
<th></th>
<th>Expression in hindgut (percent of cells)</th>
<th>Expression in excretory duct (percent of cells)</th>
<th>Head cell expression (average no)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>10</td>
<td>69</td>
</tr>
<tr>
<td>gu24</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>6</td>
<td>73</td>
</tr>
<tr>
<td>gu47</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>gu48</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>7</td>
<td>51</td>
</tr>
<tr>
<td>gu44</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>9</td>
<td>32</td>
</tr>
</tbody>
</table>

Continued
C. Category 1 mutants can be constipated.

### Expression in phasmid (percent of cells)

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>gu24</th>
<th>gu47</th>
<th>gu48</th>
<th>gu44</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>69</td>
<td>73</td>
<td>35</td>
<td>51</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>

### Defecation defect (percent of animals)

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>gu24</th>
<th>gu47</th>
<th>gu48</th>
<th>gu44</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>59</td>
<td>36</td>
<td>35</td>
<td>51</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.5: Hindgut and male tail defects associated with Category 1 mutants. Mutations in Category 1 confer an enhanced spicule defect on \textit{lin-48(sa469)/lin-48(+)} mutants. \textit{lin-48(sa469)/lin-48(+)} males generally have wild-type spicules (arrow, A), whereas in a \textit{gu47} mutant background \textit{lin-48(sa469)/lin-48(+)} males can have abnormal spicule development (arrows, B). Data are summarized in Table 2.6.
Figure 2.6: Semi-quantitative RT-PCR of endogenous *lin-48*. *act-2* was used as an internal control to normalized the amount of PCR products. ‘x’ indicates mock control. *Arrow* indicates the PCR product of *lin-48*, *arrowhead* indicates the PCR product of *act-2*. 
Figure 2.7: Altered *lin-48::gfp* expression in the hindgut of Category 2 mutants. GFP expression in the hindgut is restricted in U, F, K, K’ cells in wild-type animals (*arrowhead*, C). The GFP is altered and irregular in *gu36* mutant (*arrowhead*, D).
Figure 2.8: Ectopic neuronal extension from the hindgut. Occasionally, Category 2 mutant animals have ectopic neuronal extension (arrowheads) from the region near U cell (arrow).
Figure 2.9: Altered *lin-48::gfp* expression in excretory duct. Normal *lin-48::gfp* expression in excretory duct cell usually is present in the nucleus and in a short anterior branch along the duct (arrowhead A, B). In mutant animals, especially *gu34* which has the strongest phenotype, the excretory duct cell shows a different *lin-48::gfp* pattern. There are three to four long GFP expressing branches extending from the nucleus toward the posterior (arrowhead C, D).
Figure 2.10: Ectopic lin-48::gfp expression in vulva. The ectopic lin-48::gfp is expressed in two cells (arrowheads) at the top of the vulva, next to the uterus. It is possible that the expressing cells are vulF, but this has not been confirmed.
Figure 2.11: Multi-vulva phenotype in *gu36*. *Arrow* indicates the original vulva. Ectopic vulva (*arrowheads*) and irregular egg shapes (B) are frequently observed among category 2 mutant strains.
Figure 2.12: Altered growth of *lin-48::gfp* expressing cell. In *gu35* mutant animals, the bright *lin-48::gfp* expression at the head is significantly shorter (the distance from *arrow* to *arrowhead*) than that of wild-type animals. *Arrow* indicates the position of the nucleus of the neuronal support cell. *Arrowhead* indicates the tip of the animal’s nose.
CHAPTER 3

gu22 IS A NEW ALLELE OF EGL-38

Table 3.1 was done in collaboration with Dr. Chamberlin; Table 3.2 was done by Vandana Rajakumar and Xiaodong Wang.

3.1 INTRODUCTION

In the genetic screen for additional factors affecting the expression of EGL-38 target gene lin-48 (Chapter 2), I isolated a mutation, gu22, which affects lin-48 expression in hindgut cells. gu22 mutants exhibit a strong reduction of lin-48::gfp phenotype in contrast to the other mutants in category 1. Moreover, gu22 mutants exhibit defective hindgut and male tail development similar to that seen in egl-38 mutants. As a first step toward molecular identification of the affected gene, and to test whether gu22 is a new allele of egl-38, genetic mapping and complementation tests with egl-38 were performed. Genomic DNA sequencing of gu22 was utilized as well to further confirm the mutation in egl-38.
It has been shown that different missense alleles of *egl-38* preferentially disrupt different functions of *egl-38* in different organs. Each of these characterized missense alleles has a mutation that affects the DNA binding domain of the encoded EGL-38 protein. This observation suggests that the mutations affect the ability of EGL-38 to bind to the DNA of different target genes in the different tissues (Chamberlin, 1997). In particular, the mutations in *egl-38* do not form an allelic series. The *sy294* mutation causes a substitution of valine at an invariant glycine positioned at the end of the first alpha helix of the paired domain (Xu et al., 1995) and thus could result in a destabilization or disruption of the folded protein (Figure 3.1). The developmental consequences of this mutation are a complete disruption of male tail development, a disruption of viability and a mild disruption of egg-laying function. In contrast, the *n578* mutation is a substitution of glutamic acid at an invariant glycine that normally contacts the minor groove of the DNA (Xu et al., 1995). The consequences of this mutation are a strong disruption of egg-laying function, with only minor effects on viability and male tail development. These results suggest that there are differences between the target genes for EGL-38 in the egg-laying system and in the tail. This may reflect differences in the relative importance of different DNA targets, or that certain genes are targets in one cellular context, but not in others.

Although *egl-38* is an essential gene, its different functions have been characterized using non-null, reduction-of-function alleles that allow homozygotes to be viable. These mutant alleles affect *egl-38* in vivo function in different ways, and each allele corresponds to a missense mutation that alters an amino acid in the EGL-38 DNA binding domain (Figure 3.2) (Chamberlin et al., 1997, 1999; this work). Consequently, to
investigate the relationship between Pax proteins and their in vivo targets, we have tested
the in vivo and in vitro properties of these mutant EGL-38 variants. In this chapter, I
describe the characterization of a new *egl-38* allele, *gu22*, that also disrupts a subset of
EGL-38 functions, and that causes a distinct missense substitution in the DNA binding
domain.
3.2 MATERIALS AND METHODS

3.2.1 DNA sequencing

We used PCR to amplify *egl-38* genomic DNA from mutant animals according to the single worm PCR method of Barstead *et al.* (1991). PCR products were sequenced directly by the Plant-Microbe Genomics Facility of the Ohio State University using an Applied Biosystems 3730 DNA Analyzer and BigDye™ cycle sequencing terminator chemistry. The PCR products covered all predicted exons, introns and at least 300bp beyond the first and last exons.

3.2.2 *lin-48::gfp* expression

To characterize the expression of *lin-48::gfp* in mutant animals (Table 3.1), the transgene *saIs14* was used. This transgene contains the pTJ1157 reporter gene with the upstream regulatory sequences from *lin-48* that include both identified EGL-38-responsive elements (lre1 and lre2). In wild-type animals, these elements exhibit functional redundancy, but transgenes containing a single response element exhibit enhanced sensitivity in different *egl-38* genetic backgrounds (Johnson *et al.*, 2001). Thus, the most robust EGL-38 response is mediated through both elements. Trans-heterozygous animals were created by crossing *saIs14*-bearing males with *unc-22(s7) egl-38(x)* hermaphrodites. Male offspring (*unc-22(s7) egl-38(x) / + + *) were crossed to *dpy-20(e1282) egl-38(y) / DnT1* hermaphrodites. Twi (Unc-22 heterozygotes) non-Unc (non-DnT1) animals were scored. *egl-38(s1775)* is a lethal allele that behaves as a genetic null for *egl-38* (Chamberlin *et al.*, 1997).
3.3 RESULTS

3.3.1 Genetic mapping of gu22

\textit{gu22} was mapped to linkage group IV by crossing animals bearing the mutation with animals from mapping strains containing markers from each linkage group: MT3751 (\textit{dpy-5(e61)} I; \textit{rol-6(e187)} II; \textit{unc-32(e189)} III) and MT464 (\textit{unc-5(e53)} IV; \textit{dpy-11(e224)} IV; \textit{lon-2(e678)} X). \textit{gu22} exhibited linkage with \textit{unc-5(e53)}, but not the other markers. Since \textit{gu22} males are not able to mate, complementation with \textit{egl-38} was tested by crossing \textit{gu22} heterozygous males (\textit{gu22} / \textit{dpy-20(e1282)}) with \textit{dpy-20(e1282)} \textit{egl-38(sy294)} hermaphrodites, and observing \textit{egl-38}-phenotypes among the non-Dpy cross-progeny (\textit{gu22} / \textit{dpy-20(e1282)} \textit{egl-38(sy294)}). Indeed, the non-Dpy cross progeny showed \textit{egl-38}-phenotypes, suggesting \textit{gu22} failed to complement \textit{egl-38}. Thus, I identified a new \textit{egl-38} allele, \textit{gu22}.

3.3.2 A point mutation was found in the new \textit{egl-38} allele by sequencing

To further confirm \textit{gu22} is a new \textit{egl-38} allele and to identify the nucleotide mutation, the entire \textit{egl-38} gene from \textit{gu22} mutants was sequenced. A single point mutation was found in the coding region corresponding to the EGL-38 paired domain (Figure 3.1; Figure 3.2). The \textit{gu22} missense transition mutation changes the methionine (\textit{ATG}) to isoleucine (\textit{ATA}) in the first $\alpha$-helix of the N-terminus DNA binding motif of paired domain. Moreover, previous work has identified cDNA of two classes of \textit{egl-38} transcripts (Chamberlin et al., 1997). The predicted product from the longer transcript
includes a complete paired domain, whereas the shorter transcript would produce a protein with a disrupted amino-terminal portion of the paired domain (Figure 3.3). The gu22 point mutation also disrupts the start codon of the shorter alternative splicing form of egl-38. The next possible start codon is 65 codons downstream which would eliminate most part of paired domain.

### 3.3.3 The longer transcript mediates all known functions of egl-38

Previous work has identified cDNA of two classes of egl-38 transcripts (Chamberlin et al., 1997). While both are possible to contribute to egl-38 function, clarifying their contribution will further help to understand egl-38 regulation.

Based on DNA sequence, gu22 causes an amino acid change in the paired domain of the product from the longer transcript, and also eliminates the first start codon from the shorter transcript. The missense mutation, causing a substitution of isoleucine for methionine, is fairly conservative, and might not be expected to significantly impact the function of the mutant protein. These observations raised the possibility that elimination of the product from the shorter transcript, rather than alteration of the product from the longer transcript, is responsible for the gu22 phenotypes. In the assay of lin-48::gfp expression in the hindgut for egl-38 alleles, gu22/s1775 transheterozygotes exhibited an enhanced defect, consistent with the result you might expect from testing a hypomorph in trans to a null. However, the s1775 allele specifically affects the splicing of the first exon of the longer transcript, whereas the shorter transcript is intact a non-mutant. If the
primary effect of gu22 is to disrupt the product of the shorter transcript, then one might predict transheterozygotes would exhibit intra genetic complementation, with a less severe, rather than more severe phenotype.

As indicated above, the s1775 mutation disrupts splicing of the first exon of the longer transcript, but does not affect the shorter transcript. s1775 is a lethal allele and when it is transheterozygous with other egl-38 alleles, the animals exhibit enhanced reduction of all identified egl-38 functions (Chamberlin et al., 1997; Chamberlin et al., 1999). Altogether, the results with gu22 and s1775 suggest there is no characterized function that requires the shorter transcript or its product. If it has a function, either (1) the s1775 mutation can also affect its structure or production in some way, or (2) it has a function not identified in our genetic analysis, or (3) its function is dependent on the functional product from the longer transcript. For example, either the RNA or the protein product might play a negative regulatory role by interfering or competing with the longer product. However, our current results are consistent with the longer transcript mediating all the functions of egl-38.

3.3.4 Characterization of gu22 and comparison with other egl-38 alleles

The paired domain of EGL-38 functions as a DNA binding domain. Three previously isolated missense mutations (sy287, sy294, n578) of egl-38 affect the paired domain coding sequence, and preferentially disrupt different functions of EGL-38 (Chamberlin et al., 1997; Chamberlin et al., 1999). The isolation of this fourth mutation
affecting the paired domain might help to further explore the roles of the paired domain in EGL-38 function, and in the regulation of EGL-38 target genes. To compare the new egl-38(gu22) allele to the previously isolated alleles, and to extend the analysis of each allele for a range of functions, a series of phenotypic characterizations were performed to analyze the level of egl-38 function associated with each mutant allele. Each allele was tested for its effect on egg-laying ability and male tail morphology, traits associated with egl-38 mutants that have been described before (Chamberlin et al., 1997). In addition, egg-laying function was characterized at the cellular level by scoring the morphology of the vulF cells in L4 larval stage hermaphrodites, and the excretory system function was assessed by observing the position of the excretory duct cell opening in L3 larval stage animals. Since three other missense egl-38 alleles are reduction-of-function mutants and preferentially disrupt different functions of egl-38, the new allele gu22 was put in trans with another allele as a dosage assay. Expression of lin-48::gfp in hindgut cells was scored in animals homozygous for each allele and in animals bearing different transheterozygous allelic combinations (Table 3.1). These assays showed how the point mutation in the paired domain contributes to the alterations of egl-38 functions.

Most gu22 mutant males have developmental defects of male spicules and tail (Figure 2.5 D) which in turn affects the ability of gu22 males to mate. The male tail function is disrupted in sy287 and sy294, but is largely retained in n578 males (Chamberlin et al., 1997; Chamberlin et al., 1999). Thus with respect to male tail development function, gu22 clusters with sy287 and sy294. This is consistent with the result that gu22 fail to complement other alleles, but not n578 in the assay for hindgut expression of lin-48::gfp (Table 3.1). In contrast, egg-laying ability of gu22
hermaphrodites is less affected (Table 2.4; Table 3.2) while sy294 and n578 have moderate and server egg-laying deficiency, respectively (Chamberlin et al., 1997; Chamberlin et al., 1999). The excretory system function was only affected in sy294 mutant. Overall, gu22 is most similar to sy287, and distinct from n578 and sy294 with respect to the egl-38 functions disrupted in egl-38(gu22) mutants.

Although each allele disrupts certain functions of egl-38, we have found that they cannot be ordered into an allelic series, going from weakest to strongest allele. Instead, the alleles exhibit cell or tissue preferential defects compared to each other. One possible interpretation of the loss-of-function genetics is that there are differences in tissue-specific stability, or subtle differences in protein activity for the different protein variants.
3.4 DISCUSSION

3.4.1 Alterations to a Pax protein DNA binding domain can affect its function in a tissue-preferential manner

DNA binding affinity is a key component of transcription factor function. Our results with EGL-38 indicate that mutations in the DNA binding domain of Pax proteins can alter their DNA binding and their in vivo activity. Many mutations that affect the DNA binding domain of Pax proteins have been characterized previously (Tang et al., 1997; DeStefano et al., 1998; Vilain et al., 2001; Jumlongras et al., 2004). However, since we have characterized several alleles and functions of egl-38, we have been able to demonstrate that the alleles that affect the DNA binding domain of EGL-38 can preferentially affect its activity in certain functions, but not others. Our results with EGL-38 suggest one of two possible models to explain the tissue-preferential effects. One possibility is that the different mutant alleles affect the ability of EGL-38 to bind to certain sequences, but not others. This model predicts that either there are one or a few key targets in each cell type, or that there are specific response element sequences that are associated with all the targets for EGL-38 in a particular cell type. An alternative model is that, although the mutations affect DNA binding, their tissue preferential effects result from their impact on EGL-38 function in combination with other proteins that influence target sequence recognition in vivo. This model predicts that there would be cell-specific co-factors that function with EGL-38 and participate in identifying which sequences correspond to an appropriate DNA binding site or that enhance EGL-38 transactivation activity. It also suggests that the interaction between the co-factor and EGL-38 would be mediated, or otherwise affected, by the EGL-38 DNA binding domain. Although we
favor the second model, clearly distinguishing between the two will require identification of proteins that act with EGL-38 and a large collection of cell-specific EGL-38 target genes.

3.4.2 The relationship between DNA binding and Pax target gene transactivation

Previous studies have shown that although Pax proteins can bind DNA as monomers, they can also form complexes with other transcription factors. For example, Pax5 and Ets transcription factors form a complex on the mb-1 promoter that results in improved DNA binding affinity compared to that seen with either protein alone (Fitzsimmons et al., 1996). Complex formation between Pax5 and Ets is mediated through the β-hairpin domain of the EGL-38 DNA binding domain (Wheat et al., 1999). In vitro, EGL-38 and C. elegans Ets proteins can also form ternary complexes on DNA (Fitzsimmons et al., 2001). Although genetic experiments have not identified a role for C. elegans Ets genes in regulating lin-48 (H.M.C., unpublished), these studies demonstrate that the DNA binding domain can play roles in both DNA binding and protein interaction. In experiments published with the genetic analysis in this chapter, Jennifer Zhang showed that the DNA binding ability of the different EGL-38 mutant proteins is affected (Zhang et al., 2005). However, the tissue-preferential features may also reflect disruption of specific protein interactions mediated by the DNA binding domain. We speculate that the sub-optimal features of endogenous Pax binding sites might indicate a common role for other factors in the recognition of in vivo Pax response elements. This combinatorial control would allow for the recognition of Pax targets in both a cell-specific and sequence-specific manner.
Percent of hindgut cells expressing *lin-48::gfp* in trans-heterozygotes

<table>
<thead>
<tr>
<th><em>egl-38</em> allele</th>
<th><em>egl-38</em> allele</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>s1775</em></td>
</tr>
<tr>
<td><em>sy294</em></td>
<td>0</td>
</tr>
<tr>
<td><em>n578</em></td>
<td>73</td>
</tr>
<tr>
<td><em>sy287</em></td>
<td>5</td>
</tr>
<tr>
<td><em>gu22</em></td>
<td>12</td>
</tr>
<tr>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: *egl-38* alleles differentially affect expression of *lin-48*
Table 3.2: *egl-38* alleles preferentially affect different functions in vivo

<table>
<thead>
<tr>
<th>egl-38 allele</th>
<th>DBD change</th>
<th>Percent wild-type</th>
<th>Microns (± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>None</td>
<td>100 100 100</td>
<td>13.07 (±1.55)</td>
</tr>
<tr>
<td>sy287</td>
<td>G9S</td>
<td>100 90 10</td>
<td>13.96 (±1.29)</td>
</tr>
<tr>
<td>gu22</td>
<td>M29I</td>
<td>94 80 2</td>
<td>13.23 (±1.54)</td>
</tr>
<tr>
<td>sy294</td>
<td>G33V</td>
<td>27 67 0</td>
<td>27.53 (±3.16)</td>
</tr>
<tr>
<td>n578</td>
<td>G69E</td>
<td>0 4 74</td>
<td>13.53 (±1.62)</td>
</tr>
</tbody>
</table>
Figure 3.1: Structure of a paired domain bound to DNA. Cylinders represent α helices; arrows represent β strands. Helices 1-6 are labeled; residue numbers indicate termini of the corresponding secondary structure elements (Modified from Xu et al., 1999). The location of each allele is indicated with an asterisk (*).
Figure 3.2: Alignment of amino acid Sequences of the paired domain from EGL-38 and other PAX proteins. The paired domain from the C. elegans Pax2/5/8 protein EGL-38 (Chamberlin et al., 1997), human PAX2 (Eccles et al., 1992), the C. elegans Pax6 protein VAB-3 (Chisholm and Horvitz, 1995), and the Drosophila Paired protein (Bopp et al., 1986) are shown. Identical amino acids are shaded with black, conserved amino acids are gray. Sequences were aligned using ClustalW 1.8 (Jeanmougin et al., 1998), and shaded using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Structural features associated with each domain are indicated below the sequence, with arrows indicating β-sheet and boxes indicating α-helical structure (Xu et al., 1995, 1999b). The amino acid changes associated with each mutant allele are indicated above. Genomic DNA sequence analysis identified that gu22 corresponds to a missense transition mutation of ATG (methionine) to ATA (isoleucine).
Figure 3.3: Gene organization for *egl-38*. Each box represents an exon. The regions coding for the Paired domain is in **black** and Octapeptide is **shaded**. The untranslated regions are dotted. There are two classes of *egl-38* transcripts (Chamberlin et al., 1997). The larger transcript has seven exons that can be translated to produce a protein of 289 amino acids. The smaller transcript results from the *C. elegans* SL1 splice leader *trans*-spliced to the second exon. Conceptual translation of this transcript results in a truncated protein that lacks the amino terminus of the larger protein, including the first 28 amino acids of the paired domain. The gu22 point mutation corresponds to a missense transition mutation of ATG (methionine), indicated with an asterisk (*), to ATA (isoleucine) in the larger transcript. Moreover, gu22 also disrupts the start codon of the smaller alternative splicing transcript of *egl-38*.  

**ATG**  
**ATG**  
**TAG**  
Paired domain  
Octapeptide
CHAPTER 4

THE BROMODOMAIN PROTEIN LEX-1 ACTS WITH TAM-1 TO MODULATE GENE EXPRESSION IN C. elegans

This chapter is published in:
Figure 4.2 and 4.5 were done by Kristin R. Armstrong and Xiaodong Wang, respectively.

4.1 ABSTRACT

In many organisms, repetitive DNA serves as a trigger for gene silencing. However, some gene expression is observed from repetitive genomic regions such as heterochromatin, suggesting mechanisms exist to modulate the silencing effects. From a genetic screen in C. elegans, we have identified mutations in two genes important for expression of repetitive sequences: lex-1 and tam-1. Here we show that lex-1 encodes a protein containing an ATPase domain and a bromodomain. LEX-1 is similar to the yeast Yta7 protein, which maintains boundaries between silenced and active chromatin. tam-1 has previously been shown to encode a RING finger/B-box protein that modulates gene expression from repetitive DNA. We find that lex-1, like tam-1, acts as a class B
synthetic multivulva (synMuv) gene. However, since lex-1 and tam-1 mutants have normal P granule localization, it suggests they act through a mechanism distinct from other class B synMuv's. We observe intragenic (interallelic) complementation with lex-1 and a genetic interaction between lex-1 and tam-1, data consistent with the idea that the gene products function in the same biological process, perhaps as part of a protein complex. We propose that LEX-1 and TAM-1 function together to influence chromatin structure and to promote expression from repetitive sequences.
4.2 INTRODUCTION

A number of mechanisms in eukaryotes act to establish and maintain different chromatin states (reviewed by Lam et al. 2005). Chromatin can be in relatively “closed” configurations, which correlate with low levels of gene activity, or in more active, “open” configurations. A critical feature of chromatin structure and its relationship to gene activity is that it is both dynamic and stable. Alterations to chromatin structure in response to developmental or environmental conditions are an important mechanism to alter gene activity. At the same time, chromatin states can be maintained over long periods of time, and through cell division. It is becoming apparent that cells have extensive mechanisms to establish, maintain, and modify the epigenetic state of chromosomal regions. Identifying the responsible molecules, and characterizing how they function to influence gene activity, are important areas of inquiry.

In many organisms, repetitive DNA serves as a trigger for gene silencing. A variety of transcriptional and posttranscriptional mechanisms intersect to mediate the silencing of repetitive sequences, but one feature associated with repetitive DNA is a relatively closed chromatin structure (reviewed by Bernstein and Allis 2005; Hsieh and Fire 2000). With respect to this work, repetitive transgenes introduced into *C. elegans* that include multiple gene copies are subject to silencing. In somatic cells, they produce fewer transcripts per gene than do single-copy genes, and they are extensively silenced in the germline (Kelly et al. 1997; Mello and Fire 1995).

Thus there is both silencing of genes in repetitive DNA in *C. elegans*, and there are differences in the response between germline and somatic cells. Recent work has
focused on the germline/soma divergence with respect to silencing. Germline cells exhibit a chromatin organization which is distinct from that in somatic cells, and which requires the Polycomb Group related MES proteins (Fong et al. 2002; Kelly and Fire 1998; Schaner et al. 2003). The activity of MES proteins is normally restricted to the germline by a set of genes called class B synMuv genes (Wang et al. 2005). Mutants for these synMuv B genes exhibit reduced expression from repetitive DNA transgenes in somatic cells. However, since the somatic cells of synMuv B mutants adopt germline-like features, it is interpreted that the expression changes are a secondary effect of disrupted MES regulation and the germline/soma distinction, rather than a direct effect on repetitive DNA expression. This interpretation is consistent with the evidence that mechanistically, synMuv B factors act as transcriptional repressors. Included among the synMuv B class are the \textit{C. elegans} genes for Rb, E2F, and DP transcription factors, and components of the DRM and NuRD repressor complexes (Ceol and Horvitz 2001; Harrison et al. 2006; Korenjak et al. 2004; Unhavaithaya et al. 2002).

Thus it is unlikely that these proteins directly promote gene transcription from repetitive sequences. Indeed, it is possible that the silencing of repetitive transgenes in synMuv B mutants results from a post-transcriptional mechanism, such as enhanced somatic cell RNAi, a phenotype also observed in these mutants. synMuv B genes are so-called due to a synthetic phenotype observed when mutations are combined with those in a second category of gene (the synMuv A genes; Ferguson and Horvitz 1989). Animals doubly mutant for one synMuv A and one synMuv B gene exhibit a synthetic multivulva (synMuv) phenotype, in which an excess of ventral epidermal cells divide to produce vulval tissue (reviewed by Sternberg 2005). Animals mutant for a single synMuv gene,
however, are phenotypically wild type. At least some synMuv double mutants exhibit an excess of *lin-3* transcripts, the gene for the EGF signal that promotes vulval development (Cui et al. 2006a). Current models propose that the synMuv phenotype results from a failure to transcriptionally repress this signaling gene, leading to inappropriate, excessive production of vulval tissue.

In contrast to the factors that influence the germline/soma differences in gene silencing, less is known about *C. elegans* factors that specifically influence silencing in somatic cells, or factors that act to counter silencing and allow for expression of genes contained in repetitive sequences. One gene, *tam-1*, has been identified that normally counters the silencing effects of repetitive DNA (Hsieh et al. 1999). *tam-1* mutants exhibit reduced expression from multicopy transgenes, and are genetically class B synMuvs.

However, *tam-1* does not affect the germline/soma distinction (Wang et al. 2005), and thus may influence gene activity by a mechanism distinct from that of other synMuv B genes. *tam-1* encodes a RING finger/B-box factor. Other RING finger/B-box proteins act as E3 ubiquitin protein ligases (reviewed by Joazeiro and Weissman 2000). However, how TAM-1 influences gene activity and silencing is not known.

In this study, we describe a new gene, *lex-1*. We identified *lex-1* alleles in a screen for mutations that alter the expression of a multicopy transgene containing a reporter for the *C. elegans* Ovo-related gene, *lin-48::gfp*. We show that *lex-1* functions similarly to *tam-1*, and is a class B syn-Muv gene that affects transgene expression but not germline-related functions such as P granule localization. *lex-1* encodes a protein
containing an ATPase domain and bromodomain, domains that suggest LEX-1 associates with acetylated histones and modulates chromatin structure. LEX-1 is similar to the yeast Yta7 protein which is implicated in establishing and maintaining chromatin boundaries (Jambunathan et al. 2005; Tackett et al. 2005). We hypothesize that LEX-1 and TAM-1 modulate chromatin to establish and maintain gene activity in the context of repetitive DNA sequence.
4.3 MATERIALS AND METHODS

4.3.1 Genetic strains

Strains were grown and maintained as described (Brenner 1974; Sulston and Hodgkin 1988). Unless stated otherwise, experiments were conducted at 20°C. Mutations used:

LG I: \textit{smg-1(r861), dpy-5(e61)}

LG II: \textit{rol-6(e187)}

LG III: \textit{unc-93(e1500), lin-48(sa469), dpy-17(e164), unc-32(e189), unc-119(e2498)}

LG VI: \textit{dpy-13(e184), unc-5(e53), unc-24(e138), egl-38(sy294), dpy-20(e1282). sDf9, sDf65, sDf82}

LG V: \textit{unc-62(s472), unc-46(e177), dpy-11(e224), him-5(e1490)}

LG X: \textit{lon-2(e678), smg-2(e2008), lin-15B(n744), lin-15A(n767)}

4.3.2 Isolation of mutations

Mutations were isolated in a screen for mutants affecting \textit{lin-48} expression pattern. CM117 hermaphrodites (\textit{unc-119(e2498); saIs14(lin-48::gfp; unc-119(+)}; Sewell et al. 2003) or CM3 hermaphrodites (same genotype as CM117, but with \textit{him-5(e1490)}) were mutagenized with 50 mM EMS for 4 h (Sulston and Hodgkin 1988). F2 populations were observed using a dissecting microscope with epifluorescence. Hermaphrodites with an
altered GFP expression pattern were recovered and allowed to self-cross. The mutations described in this paper were recovered from a screen of 18,000 mutagenized gametes.

4.3.3 Genetic mapping

All mutations were backcrossed at least three times to unmutagenized parental strains, CM117 or CM3. The following markers were used to assign mutations to specific chromosomes: dpy-5 I, rol-6 II, unc-32 III, unc-5 IV, dpy-11 V, and lon-2 X. The assigned mutations were further mapped by three-factor mapping, deficiency mapping, and SNP mapping (Table 4.2). The gu22 allele was assigned to the egl-38 locus on the basis of linkage, non-complementation, and DNA sequencing (Zhang et al. 2005).

4.3.4 Molecular identification of genes

Once a gene was genetically mapped to a small interval, cosmid clones were microinjected into the mitotic germline of hermaphrodites according to the method of Mello et al. (1991). A total of 100 ng/L of plasmid containing the rol-6(su1006) allele (pRF4) were co-injected as a marker with 10 ng/L of each cosmid into the mutant animals. Heritable lines were examined for restoration of the wild-type GFP pattern. Deficiency mapping placed lex-1 to the right of the sDf82 right breakpoint (outside of sDf82) and the left of the sDf9 right breakpoint (inside of sDf9; Table 4.2), a region of less than 0.1 map units and 100 kbp sequence. We tested the following cosmids from that
region (in three different pools) for rescue: ZK822, ZK199, C37F2, ZK829, ZK617, R13A9, T12G3, K08E4. Only the pool containing K08E4, or the cosmid alone, robustly rescued \textit{lex-1} mutants (Figure 4.4a). Subclones of the cosmid were prepared using standard methods (Ausubel et al. 2000). pRT6 is a 7.3 kb \textit{SacI/PstI} subclone containing a single gene, F11A10.1, which rescued all three of the \textit{lex-1} mutants. Transgenes containing the cosmid T05B10 or the subclone pRT11 (a 13.4 kb \textit{SalI/AvrII} subclone containing a single gene, F26G5.9) rescued the \textit{tam-1} allele, \textit{gu44}. The other \textit{tam-1} alleles were not tested.

4.3.5 DNA sequencing and cDNA analysis

To identify the mutations associated with alleles, we used PCR to amplify the genomic DNA of the gene in question from mutant animals according to the single worm PCR method of Barstead et al. (1991). PCR products were sequenced directly by the Plant-Microbe Genomics Facility of the Ohio State University using an Applied Biosystems 3730 DNA Analyzer and BigDye cycle sequencing terminator chemistry. The PCR products covered all predicted exons, introns and at least 300 bp beyond the first and last exon. To characterize the mRNA product from wild-type \textit{lex-1} and \textit{lex-1(gu47)}, we used RT-PCR and primers that hybridize to sequences including the predicted start and stop codons from F11A10.1. The RT-PCR products were cloned into pBluescript II or pGEX-4T-1, and sequenced. The GenBank accession number for wild-type \textit{lex-1} cDNA is DQ140399. The nucleotide changes associated with the mutant alleles have been submitted to WormBase (http://www.wormbase.org/).
4.3.6 Analysis of lex-1 and tam-1 mutant phenotypes

*lin-48::gfp (saIs14)* expression as assayed under the dissecting microscope at 100x magnification was subjectively quantified in adult animals on a four-point scale, with 4 being wild-type expression level in all hindgut cells, 3, robust expression (but not quite wild type), 2, significantly reduced expression level, and 1, no expression. For the data of Table 4.4, animals scored 3 or higher were classified as “wild type”, whereas animals scored 2 or lower were classified as “mutant”. *nac-2::gfp (guIs1, Sewell et al. 2003)* expression was assayed under the compound microscope at 1,000x magnification. Expression was subjectively quantified in larval animals as showing high, low, or absence of expression (Figure 4.2). PGL-1 was detected in L1 larvae following the protocol of Wang et al. (2005). Wild type, *lin-15B(n744), lex-1(gu24), lex-1(gu47), lex-1(gu48), tam-1(cc567), and tam-1(gu44)* strains were all subject to PGL-1 staining, with representative animals and genotypes shown in Figure 4.5.

4.3.7 Construction of double mutants and scoring of the synMuv phenotype

*lex-1* or *tam-1* double mutants with *lin-15* (A or B) were constructed by crossing *lex-1* or *tam-1* males (carrying the *lin-48::gfp* transgene *saIs14*) with *lin-15* hermaphrodites. F1 males were then back-crossed with *lin-15* hermaphrodites, and the resulting transgene-bearing F2 hermaphrodite offspring were self-crossed. Since *lin-15* is X-linked, these animals are all homozygous for the *lin-15* mutant allele. F3 hermaphrodites with reduced *lin-48::gfp* were picked as double mutants. Strains were
grown at 25°C for at least two generations before scoring the Multivulva phenotype, as
*tam-1* mutants show a synMuv phenotype primarily at 25°C (Hsieh et al. 1999). L4
hermaphrodites were observed to infer the number of P(3–8).p cells that adopt vulval cell
fate at 1,000x magnification using Nomarski optics (Han et al. 1990).
4.4 RESULTS

4.4.1 A genetic screen identifies mutations that alter expression of a \textit{lin-48::gfp} transgene

The \textit{lin-48} gene is expressed and functions in a small number of cells in \textit{C. elegans} (Johnson et al. 2001; Wang and Chamberlin 2002). Normal expression in two distinct types of cells—the excretory duct cell and a subset of hindgut cells—requires the activity of the Pax2/5/8-related transcription factor EGL-38, whereas expression in head and phasmid cells is EGL-38-independent. To identify new genes that function with \textit{egl-38} to regulate \textit{lin-48} expression, we utilized a genetic strain bearing a \textit{lin-48::gfp} reporter transgene, and performed an F2 EMS mutagenesis screen for mutants with reduced expression in hindgut cells (Figure 4.1). We mapped and characterized 12 recessive mutations that correspond to three genes (Table 4.1). We identified one allele of \textit{egl-38} itself (\textit{gu22}; Zhang et al. 2005), confirming that the screen can identify mutants with the \textit{egl-38} phenotype. The remaining alleles identify two other genes, \textit{tam-1} and \textit{lex-1}.

\textit{tam-1} was previously identified as a gene that promotes expression of transgenes with repetitive gene copies similar to the transgene used in our genetic screen. To determine whether \textit{lex-1} is \textit{lin-48}-specific, or broadly alters transgene expression as does \textit{tam-1}, we tested the expression of a second transgene, \textit{nac-2::gfp} (Figure 4.2; Sewell et al. 2003). We find that \textit{nac-2::gfp} and \textit{lin-48::gfp} expression is disrupted a similar extent. We observe a variable silencing of transgene expression with both \textit{lin-48::gfp} and \textit{nac-2::gfp}. This variability is observed between animals (Figure 4.2 G), but also within
animals, with some cells showing robust expression while others lack detectable GFP (e.g., Figure 4.2 D, F). In addition, the cells showing altered expression vary from one animal to the next. Thus the transgenes exhibit a variable rather than a uniform silencing, even in the presumed null mutant \textit{tam-1}(cc567). We conclude that \textit{lex-1} and \textit{tam-1} are phenotypically similar with respect to their impact on the expression of repetitive transgenes.

Eight alleles of the gene \textit{tam-1} were recovered from the screen. The alleles were assigned to the \textit{tam-1} locus on the basis of genetic mapping, transformation rescue, complementation, and sequencing of mutant alleles (Table 4.2, Figure 4.3). \textit{tam-1} encodes a RING-finger/B-box-containing protein (Hsieh et al. 1999). In contrast to the previously reported nonsense alleles, all eight alleles identified in our screen are missense. Four alleles (\textit{gu41}, \textit{gu43}, \textit{gu44}, \textit{gu49}) result in missense changes in the RING finger domain of TAM-1, and in vivo these four alleles exhibit a stronger defect than the other missense alleles (data not shown). \textit{gu41}, \textit{gu43}, and \textit{gu49} all alter highly conserved amino acids (cysteines and a conserved proline, respectively). \textit{gu44} alters a proline that is contained in a TAM-1-specific (non-conserved) 85-amino acid loop between cysteine five and six of the RING finger. We selected one of these stronger alleles (\textit{gu44}) and the canonical nonsense allele (\textit{cc567}) for the analysis below.

Three alleles define a new gene we term \textit{lex-1} (\textit{lin-48} expression abnormal). We mapped \textit{lex-1} to linkage group IV, and used genetic mapping, transformation rescue, and sequencing of mutant alleles to establish that \textit{lex-1} corresponds to the gene F11A10.1 (Table 4.2, Figure 4.4). To characterize the transcript made from the \textit{lex-1} locus, we
performed RT-PCR and sequenced the product. We find that \textit{lex-1} has eight exons, including an exon (exon six) that was not previously predicted using bioinformatic methods (Figure 4.4 B; http://www.wormbase.org/). This gene codes for a protein with an AAA+ ATPase domain (Figure 4.4 C) and a bromodomain (Figure 4.4 D), and it is similar to the product of the yeast \textit{yta7} gene, as well as proteins from mammals and other organisms (Figure 4.4 and Figure 4.6). Since the bromodomain can allow proteins, including Yta7, to interact with acetylated histones (Dhalluin et al. 1999; Jambunathan et al. 2005; Owen et al. 2000), the LEX-1 protein is predicted to interact with chromatin. In addition, many AAA+ ATPase proteins are part of large multiprotein complexes that hydrolyze ATP to direct translocation or remodeling of target substrates (reviewed by Erzberger and Berger 2006). The two domains together suggest a protein that (as part of a larger complex or “machine”) can recognize specific chromatin modifications, and translate them into chromatin structural features.

\textit{lex-1(gu24)} is a missense mutation that alters the ATPase domain, and \textit{lex-1(gu48)} is a missense mutation that alters the bromodomain. These two alleles confirm the functional importance of these two conserved domains. \textit{lex-1(gu47)} affects the splice acceptor site for the last intron. We performed RT-PCR on RNA from \textit{lex-1(gu47)} animals, and identified a product that utilizes an alternative splice site one nucleotide downstream of the normal site. This altered splicing shifts the coding frame, and would result in replacement of the last 45 amino acids of the carboxyl terminus with 46 amino acids from the alternate frame. As we only sequenced one RNA product from \textit{lex-1(gu47)}, it is possible that other splice variants are also produced from the \textit{lex-1(gu47)} allele. However, since \textit{lex-1(gu47)} alters splicing, we tested whether the mutant phenotype
resulted from reduced mRNA abundance due to nonsense-mediated decay. We were unable to detect phenotypic suppression of \textit{lex-1(\textit{gu47})} by mutations in \textit{smg-1} or \textit{smg-2}, genes required for nonsense mediated decay (data not shown; Hodgkin et al. 1989; Pulak and Anderson 1993). This result suggests sequences in the carboxyl terminus of LEX-1 are important for its function or stability.

Homozygous mutants for each of our three \textit{lex-1} alleles show similar defects with respect to altered transgene expression. However, \textit{lex-1(\textit{gu47})} mutants exhibit more of the growth defects and abnormalities observed in animals depleted for \textit{lex-1} using double-stranded RNA interference (data not shown; Kamath et al. 2003; Piano et al. 2002). We have found that each of our mutant alleles is recessive to wild type, and that \textit{\textit{gu24}} and \textit{\textit{gu48}} are viable in trans to deficiency (data not shown and Table 4.2; \textit{\textit{gu47}} was not tested in the deficiency assay). Since RNAi depleted animals exhibit significant embryonic lethality, it is predicted that the null phenotype for \textit{lex-1} is lethal. Altogether, we suggest that our alleles are non-null, reduction of function alleles of \textit{lex-1}.

\section*{4.4.2 \textit{lex-1} is a class B synthetic Multivulva gene}

\textit{tam-1} was previously shown to be a class B synthetic Multivulva (synMuv) gene, as double mutants with class A genes, but not other class B genes, exhibit a synthetic multivulva phenotype when grown at 25°C (Hsieh et al. 1999). Since \textit{lex-1} and \textit{tam-1} exhibit similar phenotypes with respect to transgene expression, we tested whether \textit{lex-1} is also a synMuv gene. We find that \textit{lex-1; lin-15A} double mutants exhibit a modest
Multivulva phenotype at 25°C, whereas \textit{lex-1;} \textit{lin-15B} mutants are wild type (Table 4.3). Thus, \textit{lex-1} and \textit{tam-1} exhibit similar weak and temperature-sensitive synMuv B defects. Although \textit{tam-1} mutants exhibit a synthetic Multivulva phenotype in combination with class A synMuv mutants, they are distinct from some other synMuv B genes due to pleiotropic phenotypes associated with the mutants. Specifically, the somatic cells of \textit{tam-1} mutants do not exhibit germline features such as P granule localization or enhanced RNAi effect (Wang et al. 2005). As alleles of both \textit{tam-1} and \textit{lex-1} were identified in our screen, we tested mutants for P granule localization using an anti-PGL-1 antibody (Figure 4.5; Kawasaki et al. 1998). We find that the P granules are restricted to the germline in \textit{lex-1} mutants as in wild type, and in contrast to class B synMuv mutants such as \textit{lin-15B}(n744). Thus \textit{tam-1} and \textit{lex-1} influence multicopy transgene expression and the vulval development process either more specifically than these other class B synMuv genes, or through a distinct mechanism.

\textbf{4.4.3 \textit{lex-1} and \textit{tam-1} exhibit genetic interactions}

In the course of completing genetic complementation tests with the alleles recovered in our screen, we identified genetic interactions between \textit{lex-1} alleles, as well as between \textit{tam-1} and \textit{lex-1} alleles. \textit{lex-1}(gu24) and \textit{lex-1}(gu48) are missense mutations affecting the ATPase domain and the bromodomain, respectively (Figure 4.4 C, D). In complementation tests, we find that \textit{lex-1}(gu47) fails to complement the other two alleles, whereas \textit{lex-1}(gu24) and \textit{lex-1}(gu48) exhibit intragenic complementation (Table 4.4). This result indicates that LEX-1 function can be restored if the two conserved domains
are present in separate proteins. One interpretation of complementation of this type is that
the LEX-1 protein functions as part of a complex that includes more than one LEX-1
protein (Yook 2006). Alternatively, the result suggests the activities associated with the
ATPase domain and the bromodomain can be physically or temporally separated. We
find that the \textit{lex-1}(gu48) allele also exhibits an extragenic non-complementation
interaction with \textit{tam-1} alleles. In crosses between \textit{lex-1}(gu48) hermaphrodites and \textit{tam-1}
mutant males, 35\% of cross progeny exhibit a mutant phenotype (compare to 13\% when
crossed with \textit{lex-1}(gu24) mutant males; Table 4.4). The reciprocal cross of \textit{tam-1} mutant
males and \textit{lex-1}(gu48) hermaphrodites yields 0–2\% mutant phenotype, indicating that the
interaction exhibits a maternal effect. The sensitivity of \textit{tam-1} heterozygotes to maternal
\textit{lex-1}(gu48) (but not other \textit{lex-1} genotypes) suggests maternal LEX-1 can interfere with
the activity of zygotic TAM-1. Along with the shared phenotypic profile, the genetic
interaction between \textit{lex-1} and \textit{tam-1} is consistent with the idea that the two genes
function in the same biological process.
4.5 DISCUSSION

In this paper we report the genetic and molecular characterization of the *C. elegans* gene *lex-1*. *lex-1* encodes a protein with a bromodomain and an ATPase domain similar to yeast Yta7, a protein important for maintaining chromatin states and repressing the inappropriate spreading of heterochromatin silencing. We find that *lex-1* shares phenotypic features with a previously described gene, *tam-1*, in that both genes affect expression from multicopy transgenes, and act as synMuv B genes. synMuv B genes are best known for their protein products, which include components of the *C. elegans* DRM and NuRD transcriptional repressor complexes, and which act to suppress germline features in somatic cells. However, we suggest *lex-1* and *tam-1* are functionally different from other synMuv B genes, as the mutants do not exhibit germline features in somatic cells. We propose a model in which LEX-1 and TAM-1, a RING-finger/B-box protein, are part of a complex that enhances transcription from repetitive DNA sequences by modulating chromatin structure.

4.5.1 Molecular and genetic characterization of the *lex-1* gene

Three alleles of *lex-1* were identified in a genetic screen for altered expression of a *lin-48::gfp* transgene. This genetic screen was designed to recover mutations in genes that influence *lin-48* expression in a cell-specific manner, and did indeed recover a new allele of the *Pax2/5/8* gene *egl-38* (Zhang et al. 2005). However, our results suggest that the *lex-1* alleles were recovered based on their influence on multicopy transgene
expression, and that they have, at best, a modest role in regulating the endogenous \textit{lin-48} gene product. Specifically, we have shown that \textit{lex-1}, like \textit{tam-1}, alters expression from a non-\textit{lin-48}-related transgene. In addition, other genes that affect endogenous \textit{lin-48} gene expression share \textit{lin-48} mutant phenotypes in the hindgut, excretory system, or both (Johnson et al. 2001; Wang et al. 2006; Zhang et al. 2005). In contrast, these tissues are essentially normal in \textit{lex-1} mutants (data not shown). The genetic interactions between \textit{lex-1} and \textit{tam-1}, a gene previously identified for its role in promoting expression from multicopy transgenes (Hsieh et al. 1999), are also consistent with shared functions for \textit{lex-1} and \textit{tam-1}. We speculate that our screen failed to identify new genes that affect \textit{lin-48} gene expression due to functional redundancy in \textit{lin-48} transcriptional regulation (Johnson et al. 2001; Wang and Chamberlin 2002) or to the essential roles of the trans-acting factors.

The \textit{lex-1} gene encodes a protein with a bromodomain and an ATPase domain. LEX-1 shares sequence similarity with proteins found throughout eukaryotes (Figure 4.4), including the Yta7 protein from yeast. Yta7 is part of a protein complex that associates with chromatin (Tackett et al. 2005). It co-purifies with histones that are enriched for distinct patterns of posttranslational modification, and functions to maintain the boundary between transcriptionally active and transcriptionally silenced regions of the chromosome (Jambunathan et al. 2005; Tackett et al. 2005). Altogether, the data suggest that Yta7-related proteins associate with particular chromosomal regions, and act to modulate or stabilize chromatin states. Our data with LEX-1 are consistent with it functioning in a similar manner in \textit{C. elegans}. For example, LEX-1 complexes may associate with transgene sequences and stabilize expression, counterbalancing the factors that act to
repress transcription from repetitive sequences. According to this model, we propose that LEX-1 would normally act to enhance expression from (unknown) endogenous loci that are proximal to repetitive or heterochromatic sequences, and that our repetitive transgenes serve as a reporter for LEX-1 activity. However, we cannot rule out the alternative possibility that LEX-1 influences expression from repetitive transgenes through an indirect mechanism.

Other AAA+ ATPase proteins like LEX-1 multimerize into large complexes or arrays to direct translocation or remodeling of target substrates (reviewed by Erzberger and Berger 2006). We observe intragenic complementation between \textit{lex-1}(gu48) and \textit{lex-1}(gu24) alleles, a result consistent with the idea that LEX-1 proteins multimerize. However, it is interesting that these alleles are recessive to wild type (rather than antimorphic and dominant), and can compensate for each other (rather than interfere). This suggests that if LEX-1 is indeed part of a multimer complex, it is not necessary for all molecules in the complex to be fully functional.

4.5.2 LEX-1 and TAM-1 function together to modulate gene expression

Our results indicate that \textit{lex-1} and \textit{tam-1} share in vivo functions. Both genes act to maintain wild-type expression levels from multicopy transgenes, and our genetic screen identified multiple alleles of each of these genes, but not alleles of other genes that are known to influence the expression from transgenes. In addition, both \textit{lex-1} and \textit{tam-1} act as weak synMuv B genes, but do not exhibit ectopic P granules as do some other synMuv
B mutants. The genetic interaction between \textit{lex-1}(gu48) and \textit{tam-1} mutants is also consistent with the idea that they function in a similar process in vivo. Specifically, we have shown that \textit{tam-1/+} embryos are sensitive to the presence of \textit{lex-1}(gu48) when it is the only variant of \textit{lex-1} provided by the mother, whereas they are not similarly sensitive to other \textit{lex-1} mutant genotypes. This suggests that LEX-1 may be important for TAM-1 function, such that defective maternal LEX-1 interferes with full function of zygotic TAM-1. One hypothesis that accounts for the observed genetic interaction is that the two proteins interact with each other or are part of a complex, with LEX-1 responsible for recruiting TAM-1 to appropriate locations on the chromatin through its bromodomain. We favor this hypothesis over others (such as one in which maternal LEX-1 participates in the zygotic transcription of \textit{tam-1}), as it accounts for the observed genetic interaction with the allele that affects the bromodomain (\textit{gu48}), yet the absence of a similar interaction with the alleles that affect the ATPase domain (\textit{gu24}) of the protein product, or splicing of the last exon (\textit{gu47}).

How might LEX-1 and TAM-1 function together to influence gene transcription? The presence of a bromodomain and an ATPase domain in LEX-1 suggests it may recognize acetylated histones, and use the hydrolysis of ATP to alter chromatin structure. In other words, it may act as an interpreter of the histone code. We also speculate that LEX-1 is important for the recruitment of TAM-1 to chromatin. TAM-1 contains a RING finger and a B-box, domains that suggest it functions as an E3 ubiquitin protein ligase (reviewed by Joazeiro and Weissman 2000). Although a common role for E3 ligases is to attach the poly-ubiquitin chains that target proteins for proteasomal degradation, they can also mediate reversible mono- and poly-ubiquitinations that alter protein activity,
subcellular localization, or interactions (reviewed by Schnell and Hicke 2003). For example, dynamic ubiquitination and de-ubiquitination of histone H2B is important for gene transcription activation, and ubiquitination of histone H2A is associated with chromatin silencing (de Napoles et al. 2004; Henry et al. 2003). In the context of our genetic results, we favor a model in which TAM-1 is a ubiquitin E3 ligase that modifies histones, and thereby influences chromatin structure and gene expression. Future genetic, biochemical and protein localization experiments should clarify the molecular relationship between LEX-1 and TAM-1, and their role in modulating gene expression.

4.5.3 LEX-1 and TAM-1 and class B synMuv genes

Although *C. elegans* class B synMuv genes are defined by their synthetic Multivulva phenotype, they can be divided into different categories based on associated pleiotropies. *lin-35/Rb* and related genes (the “germline/soma synMuv B” genes) act to restrict a germline organization of chromatin to germline cells (Wang et al. 2005). Mutants share a range of pleiotropic phenotypes: reduced transgene expression, ectopic P granules, enhanced RNAi, and functional interactions with genes of the SWI/SNF complex (Cui et al. 2004; Wang et al. 2005). Previous studies show that the synthetic multivulva phenotype of these synMuv B mutants, as well as other pleiotropies, is dependent on *mes-4* (Cui et al. 2006b; Wang et al. 2005) *mes-4* encodes a histone H3methyltransferase required for normal germline development (Bender et al. 2006). In the germline/soma synMuv B mutants, the synthetic multivulva phenotype results from de-repression of the EGF gene *lin-3* (Cui et al. 2006a). Thus a simple model for how
these synMuv B genes impact vulval development is that the alteration of somatic cell chromatin to a more germline-like state, in absence of other repressive mechanisms (provided by the synMuv A genes), permits inappropriate expression of \textit{lin-3}. Notably, in this model the synMuv phenotype (as well as other pleiotropies such as silencing of repetitive transgenes) is a secondary consequence of the primary germline/soma defect.

In contrast, other synMuv B mutants, including \textit{lin-36}, \textit{lin-61}, \textit{tam-1}, and now \textit{lex-1}, lack at least some of the pleiotropic defects that result from a disruption of the germline/soma distinction (Cui et al. 2004; Harrison et al. 2007; Wang et al. 2005). Consequently, the mechanism(s) whereby these genes influence vulval development may be different from the germline/soma synMuv B genes. \textit{lin-61} encodes an MBT-repeat protein predicted to act as a transcriptional repressor (Harrison et al. 2007). \textit{lin-61} mutants do not exhibit germline/soma pleiotropies, but also do not exhibit the silencing of repetitive transgenes observed in \textit{tam-1} and \textit{lex-1} mutants. Based on the identity of LIN-61 as an MBT repeat containing protein, it may also influence \textit{lin-3} transcription, although the current model would predict its effect to be \textit{mes-4}-independent. This work focuses on the genes \textit{tam-1} and \textit{lex-1}, and shows that they share similar mutant phenotypes and exhibit a genetic interaction. Although further experiments will be required to determine how TAM-1 and LEX-1 influence vulval development, our results predict that their synMuv effect would also be \textit{mes-4}-independent. Sequence features and phenotypic analysis suggest TAM-1 and LEX-1 influence gene transcription. However, current data do not point to a clear transcriptional activation or repression mechanism. Indeed, in contrast to the germline/soma synMuv B proteins, LEX-1 and TAM-1 may act to regionally maintain transcriptional states or inhibit the spread of silencing, rather than
to repress transcription. Thus LEX-1 and TAM-1 could act to promote the transcription from repetitive DNA in a direct, rather than indirect, manner.
<table>
<thead>
<tr>
<th>Gene</th>
<th>LG</th>
<th>Alleles isolated</th>
<th>No. other alleles</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>egl-38</td>
<td>IV</td>
<td>gu22</td>
<td>4</td>
<td>Trent et al. (1983); Chamberlin et al. (1997, 1999); Zhang et al. (2005)</td>
</tr>
<tr>
<td>lex-1</td>
<td>IV</td>
<td>gu24, gu47, gu48</td>
<td>0</td>
<td>This work.</td>
</tr>
<tr>
<td>tam-l</td>
<td>V</td>
<td>gu20, gu25, gu41, gu43, gu44, gu45, gu49, gu50</td>
<td>&gt;18</td>
<td>Hsieh et al. (1999)</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of genes recovered in a screen for altered *lin-48::gfp* expression
<table>
<thead>
<tr>
<th>Gene</th>
<th>Heterozygote</th>
<th>Recombinants</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>lex-1</td>
<td>lex-1(gu24) / dpy-13 unc-5</td>
<td>Dpy non-Unc</td>
<td>dpy-13 (3) lex-1 (0) unc-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-Dpy Unc</td>
<td>dpy-13 (3) lex-1 (0) unc-5</td>
</tr>
<tr>
<td>lex-1(gu24)</td>
<td>unc-24 dpy-20</td>
<td>Unc non-Dpy</td>
<td>unc-24 (7) lex-1 (0) dpy-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-Unc Dpy</td>
<td>unc-24 (6) lex-1 (0) dpy-20</td>
</tr>
<tr>
<td>lex-1(gu48)</td>
<td>unc-24 dpy-20</td>
<td>Unc non-Dpy</td>
<td>unc-24 (5) lex-1 (0) dpy-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-Unc Dpy</td>
<td>unc-24 (3) lex-1 (0) dpy-20</td>
</tr>
<tr>
<td>lex-1*</td>
<td>sDf82; him-5 / +</td>
<td></td>
<td>sDf82 does not delete lex-1</td>
</tr>
<tr>
<td>lex-1*</td>
<td>sDf65; him-5 / +</td>
<td></td>
<td>sDf65 deletes lex-1</td>
</tr>
<tr>
<td>lex-1*</td>
<td>sDf9; him-5 / +</td>
<td></td>
<td>sDf9 deletes lex-1</td>
</tr>
<tr>
<td>tam-1</td>
<td>tam-1(gu44) him-5 / unc-46 dpy-11</td>
<td>Unc non-Dpy</td>
<td>unc-46 (0) tam-1 (13) dpy-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-Unc Dpy</td>
<td>unc-46 (0) tam-1 (39) dpy-11</td>
</tr>
<tr>
<td>tam-1(gu20)</td>
<td>unc-46 / CB4856</td>
<td>Tam-1 non-Unc-46</td>
<td>snp_C13D9[12] (2) snp_F14F9[1] (0) unc-46</td>
</tr>
</tbody>
</table>

*Tested with both gu24 and gu48

**Table 4.2: Genetic mapping of lex-1 and tam-1**
<table>
<thead>
<tr>
<th>Genotype</th>
<th>VPCs induced</th>
<th>% Muv animals</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>lex-1</em>(gu24)</td>
<td>3.00 (n=47)</td>
<td>0</td>
</tr>
<tr>
<td><em>lex-1</em>(gu47)</td>
<td>3.02 (n=56)</td>
<td>2</td>
</tr>
<tr>
<td><em>lex-1</em>(gu48)</td>
<td>3.02 (n=54)</td>
<td>2</td>
</tr>
<tr>
<td><em>tam-1</em>(cc567)</td>
<td>3.00 (n=54)</td>
<td>0</td>
</tr>
<tr>
<td><em>tam-1</em>(gu44)</td>
<td>3.00 (n=63)</td>
<td>0</td>
</tr>
<tr>
<td><em>lin-15A</em>(n767)</td>
<td>3.05 (n=41)</td>
<td>5</td>
</tr>
<tr>
<td><em>lex-1</em>(gu24) ; <em>lin-15A</em>(n767)</td>
<td>3.13 (n=64)</td>
<td>13</td>
</tr>
<tr>
<td><em>lex-1</em>(gu47) ; <em>lin-15A</em>(n767)</td>
<td>3.43 (n=53)</td>
<td>26</td>
</tr>
<tr>
<td><em>lex-1</em>(gu48) ; <em>lin-15A</em>(n767)</td>
<td>3.35 (n=49)</td>
<td>31</td>
</tr>
<tr>
<td><em>tam-1</em>(cc567) ; <em>lin-15A</em>(n767)</td>
<td>3.11 (n=96)</td>
<td>8</td>
</tr>
<tr>
<td><em>tam-1</em>(gu44) ; <em>lin-15A</em>(n767)</td>
<td>3.27 (n=56)</td>
<td>23</td>
</tr>
<tr>
<td><em>lin-15B</em>(n744)</td>
<td>3.02 (n=58)</td>
<td>2</td>
</tr>
<tr>
<td><em>lex-1</em>(gu24) ; <em>lin-15B</em>(n744)</td>
<td>3.00 (n=36)</td>
<td>0</td>
</tr>
<tr>
<td><em>lex-1</em>(gu47) ; <em>lin-15B</em>(n744)</td>
<td>3.00 (n=35)</td>
<td>0</td>
</tr>
<tr>
<td><em>lex-1</em>(gu48) ; <em>lin-15B</em>(n744)</td>
<td>3.06 (n=69)</td>
<td>4</td>
</tr>
<tr>
<td><em>tam-1</em>(cc567) ; <em>lin-15B</em>(n744)</td>
<td>3.00 (n=31)</td>
<td>0</td>
</tr>
<tr>
<td><em>tam-1</em>(gu44) ; <em>lin-15B</em>(n744)</td>
<td>3.00 (n=33)</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 4.3:** *lex-1* is a class B synMuv gene
Table 4.4: Alleles of *lex-1* exhibit intragenic complementation and allele-specific intergenic interference with *tam-1*. The table shows percentage of phenotypically wild type animals in the indicated transheterozygous animals; number of animals scored is indicated in parentheses.
Figure 4.1: *lin-48::gfp* expression intensity is altered in *tam-1* and *lex-1* mutants. Nomarski (A, B), and epifluorescence (C, D) photomicrographs of L4 animals. A wild-type animal (A, C) expresses *lin-48::gfp* abundantly in the hindgut cells U, F, K, and K’ (not in plane), whereas expression in a *lex-1* mutant (B, D) is reduced or absent. The image in C corresponds to a 0.1 second exposure, and the image in D a 5.0 second exposure, under otherwise identical camera settings.
**Figure 4.2: nac-2::gfp expression intensity is reduced in tam-1 and lex-1 mutants.**

Nomarski (A, C, E) and epifluorescence (B, D, F) photographs of larva carrying a 3 kB nac-2::gfp reporter transgene. A wild-type larva (A, B) expresses GFP in the pharynx (arrow), excretory cell, and hindgut cells. A lex-1(gu47) larva (C, D) expresses normal levels of GFP in the hindgut but GFP expression is greatly reduced in the pharynx (arrow) and excretory cell. A tam-1(cc567) larva (E, F) expresses normal levels of GFP in the pharynx (arrow) and hindgut but expression is lost in the excretory cell. Exposure conditions and time are the same for all epifluorescence images. nac-2::gfp expression is reduced to varying degrees in each tissue type of lex-1(gu47) and tam-1(cc567) larva (G). For each genotype, 30 larva were scored for normal, reduced, or absent GFP expression in the pharynx, excretory cell, and hindgut cells. The percent of larva with normal levels of GFP expression is represented in black, reduced expression in gray, and no expression in white.
Figure 4.2
**Figure 4.3:** *tam-1* alleles isolated in the genetic screen.  
(A) Gene organization for *tam-1*. Each box represents an exon. The regions coding for the RING finger is in black and the B-box domains is shaded. The untranslated regions are dotted. The location of each allele is indicated with an asterisk (*).  
(B) Amino acid sequence alignment of the RING finger domain of *C. elegans* TAM-1 and other RING finger proteins (alignment is that of Hsieh et al. 1999). The amino acid substitution associated with each mutation affecting this domain is shown with an arrow. The specific alterations associated with each allele are as follows.  
*gu41* is a tGc (cys) to tAc (tyr) transition affecting codon 21.  
*gu43* is a tGc (cys) to tAc (tyr) transition affecting codon 36.  
*gu44* is a Ccc (pro) to Tcc (ser) transition affecting codon 70.  
*gu49* is a cCa (pro) to cTa (leu) transition affecting codon 131.  
*gu20* is a aGa (arg) to aAa (lys) transition affecting codon 367.  
*gu25* is a Cca (pro) to Tca (ser) transition affecting codon 407.  
*gu50* is a Ccc (pro) to Tcc (ser) transition affecting codon 708.  
*gu45* is a Gag (glu) to Aag (lys) transition affecting codon 762.
Figure 4.4: Molecular cloning of lex-1. (A) A genetic map of LG IV is shown above, including the unc-22-deficiencies tested against lex-1. Below, a diagram of clones tested for rescue of lex-1 is aligned with the predicted genes in the lex-1 genomic region (from http://wormbase.org/). lex-1 corresponds to F11A10.1. (B) Gene organization for lex-1. Each box represents an exon. The regions coding for the ATPase domain is in black and the Bromodomain is shaded. The untranslated region is dotted. The location of each allele is indicated with an asterisk (*). gu24 is a gCt (ala) to gTt (val) transition affecting codon 548. gu48 is a gGa (gly) to gAa (glu) transition affecting codon 1001. gu47 is a G to A transition affecting the splice acceptor site for the last intron (tattcaG to tattcaA). (C) Amino acid sequence alignment for the ATPase domain from C. elegans LEX-1 (Cel, accession no. DQ140399), mouse ATAD2 (Mmu, accession no. Q8CDM1), human (Hsa, accession no. XP039676), S. pombe (Spo, accession no. NP594020), S. cerevisae Yta7 (Sce, accession no CAA56963), and A. thaliana (Ath, AAF29398). The amino acid substitution associated with gu24 is indicated. The sequences included in the alignment include those from Pfam domain PF00004 (Finn et al. 2006). (D) Amino acid sequence alignment for the LEX-1 bromodomain and that from other species, as done for the ATPase domain (C). The amino acid substitution associated with gu48 is indicated. The sequences included in the alignment include those from Pfam domain PF00439 (Finn et al. 2006). An alignment of LEX-1 and other family members across the full protein is included as Figure 4.6.
Figure 4.4

A

LG IV

unc-24  3.51

3.5  4.0  4.5  5.0  5.5

dpy-20  5.22

unc-22  5.47

sDf69 deletes lex-1
sDf65 deletes lex-1
sDf9  deletes lex-1
sDf82 does not delete lex-1

rescue

10/10  

pRT1

10/10  

pRT4

9/9    

pRT5

1/6    

pRT6

15/15  

pRT

12075k  12080k  12085k

K08E4.6

K08E4.7

F11A10.1

F11A10.2

F11A10.3

F11A10.9

F11A10.6

B

gu24

gu48

gu47

AAA ATPase

Bromodomain

Continued
Figure 4.5: A–D Anti-PGL-1 staining detects P granule localization in wild type and mutants. *lex-1 (C) and *tam-1 (D) mutant L1 larvae exhibit P granule localization in the germline (labeled with asterisk), similar to wild type (A), and distinct from *lin-15B(n744) (B), in which P granules are mislocalized to somatic cells (arrowheads).
**Figure 4.6:** Amino acid sequence alignment of LEX-1-related proteins from *C. elegans* LEX-1 (Cel, accession no. DQ140399), mouse ATAD2 (Mmu, accession no. Q8CDM1), human (Hsa, accession no. XP039676), *S. pombe* (Spo, accession no. NP594020), *S. cerevisae* Yta7 (Sce, accession no. CAA56963), and *A. thaliana* (Ath, AAF29398). The ATPase domain of Figure 4.4 C is boxed with a dotted line. The bromodomain of Figure 4.4 D is boxed with a solid line.
<table>
<thead>
<tr>
<th>Cell</th>
<th>Mau</th>
<th>Hsa</th>
<th>Spo</th>
<th>Ath</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MPRSDGFSPRENLRRSARHDHRSYAGCQNEFSDMYPASRSRSRSGQVDENG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mau</td>
<td>VYNTKSSRLTGSGKEFGPAGGAEFAPATGSSHPSSTRCTXAACPAAAKGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5l</td>
<td></td>
<td>5 YTRSEKTHNHNYYEQEEHHAISSEDBERYTRRSESNSMTYRQQVMAIDESERNKHQVYP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td>6 SGAGATLDEKRYEDSLSDHVSPFAKRTLQEPSVCDSEEKSESTGQKEWNN - L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6l</td>
<td></td>
<td>11 PAKKEIKYLSSEKFDASEAYRNVTYPERRRTRSTRRRSSMRHELQGSEKSPVEEFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7l</td>
<td></td>
<td>17 KAAAKRLGSSQFPHLNAADDLPMLR - GGEGRIVLPFAITBDMAPCAEEDLDILEKIQSN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C8l</td>
<td></td>
<td>2 EEE - EGAEEDEQGEKUPEHEHESSNAHSSSSAPROYSBQQPYQFPNA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9l</td>
<td></td>
<td>27 SAREHRRABLEHHEVAN - QMRHHRNENSGPBRSSDDSDDDDVMPDRDPKEQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C10l</td>
<td></td>
<td>33 PPMMHRGERSRPFRFINTEDELGA - QRLIKMLMYEPTDAQQG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td>3 PFRMRK - MRRAPIS - FKEKRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td>34 RSKSM - ACARHDDHRFRADLA - VILRYHNGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td>35 VPGPSADFHNHKLTTBEDSSDIDTGLG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td>34 AIPWARGQSSGAPPGFLDTYUSBGSLNVQAGSWQHGSDDGAAALTSQVYAPAESKG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C11l</td>
<td></td>
<td>37 ADDDMGSDSCDVGCOGHIQSTESEPFFPHNYFKEFNNPFGFTDF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mau</td>
<td></td>
<td>38 LADNMNIDRDNIGKGNHHLYSLSTETVDAQH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsa</td>
<td></td>
<td>39 LCDLLOVDLSTLLGILKNHFLKLYQTPFNPAGYHPFG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spo</td>
<td></td>
<td>40 GDECOQIQGNNEDDEIGDQTYGFEIY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
Continued
Figure 4.6 continued
5.1 INTRODUCTION

In this study, I used \textit{C. elegans} as a model organism to study how a given Pax transcription factor participates in the development of different organs through regulating target genes in different cell types. By utilizing a genetic screen, mutants were identified and further characterized. The two major categories of the mutant genes and encoded proteins represent potential positive and negative regulatory factors, respectively. In addition, the new isolated \textit{egl-38(gu22)} allele was compared to previously described alleles to extend the analysis of each allele for a range of functions. Each \textit{egl-38} allele disrupts certain functions of \textit{egl-38} but they cannot be ordered into an allelic series. Instead, the alleles exhibit cell or tissue preferential defects compared to each other. Furthermore, this work identified a new class B synthetic multivulva (synMuv) gene, \textit{lex-1}. Data suggest that the bromodomain protein LEX-1 and the RING finger protein TAM-1 function together to influence chromatin structure and to promote expression from repetitive sequences.
5.2 A genetic screen for factors affecting \textit{lin-48} expression

Our lab’s characterization of \textit{lin-48} indicates that EGL-38 has tissue-restricted targets that are expressed in only a subset of EGL-38-expressing cells (Johnson et al, 2001). One way EGL-38 may have different targets in different tissues is to act in a combinatorial manner with one or more additional transcription factors. In this model, both EGL-38 and the second factor would be necessary for the hindgut expression of \textit{lin-48}. Consequently, if both EGL-38 and an additional factor are required, then the second factor must meet one of the following criteria. It could act through a DNA element. This raises the possibility that EGL-38 and the second factor would physically interact. Alternatively, the second factor may not act through a discrete site, but act in a manner different from EGL-38. For example, it might influence accessibility of the \textit{lin-48} regulatory regions.

To identify additional genes important for \textit{lin-48} expression, I have performed a genetic screen for altered \textit{lin-48::gfp} expression pattern and identified mutant candidates that might represent genes that mediate the response to \textit{egl-38}, or that might function in parallel to \textit{egl-38}. This genetic screen was designed to recover mutations in genes that influence \textit{lin-48} expression in a cell-specific manner, and it did indeed recover a new allele of \textit{egl-38}, \textit{gu22}, confirming that the screen can identify mutants with Egl-38 phenotype.

In addition to the new \textit{egl-38} allele, two major categories of mutant candidates were isolated. One category of mutants has reduced \textit{lin-48::gfp} expression in the hindgut, which is a phenocopy of \textit{egl-38}. The other category showed enhance and ectopic GFP
expression. The two categories represent candidates for positive and negative regulatory factors, respectively. Most mutants in both categories have pleiotropic mutant phenotypes but exhibit moderate to low penetrance for all defects. Based on the mutant phenotypes, I interpret that the affected genes for both categories are likely modulators of lin-48 expression rather than genes that have an ‘on/off’ effect of lin-48 expression. Since EGL-38 is important for lin-48 expression in a tissue-restricted manner, the responsible genes might include cofactors that function with EGL-38, genes important for EGL-38 expression or activity. However, genes regulating lin-48 expression through other mechanisms are also possible, such as chromatin modifying enzymes or general transcriptional machinery.

The genes lex-1 and tam-1 were shown to be affected in mutants with reduced GFP expression, and were further characterized as described in Chapter 4. Mutants of category 2 exhibit pleiotropic defects, including ectopic and altered lin-48::gfp expression patterns, and morphological and physiological defects. Since category 2 mutants exhibit broadly enhanced lin-48::gfp expression which is not necessarily limited to egl-38 functioning tissues, the responsible genes of category 2 might be involved in pathways which are not egl-38-specific or are not only involved in the egl-38 pathway. Moreover, the broad range of pleiotropic defects suggests that the affected genes identified as category 2 mutants act in processes in addition to lin-48 regulation. Because of mutant strains exhibit some growth retardation and lethality, genes in category 2 might be essential.
5.3 Alterations to a Pax protein DNA binding domain can affect its function in a tissue-preferential manner

DNA binding affinity is a key component of transcription factor function. Our results with EGL-38 indicate that mutations in the DNA binding domain of Pax proteins can alter their DNA binding and their in vivo activity (Zhang et al., 2005). Many mutations that affect the DNA binding domain of Pax proteins have been characterized previously (Tang et al., 1997; DeStefano et al., 1998; Vilain et al., 2001; Jumlongras et al., 2004). However, since we have characterized several alleles and functions of egl-38, we have been able to demonstrate that the alleles that affect the DNA binding domain of EGL-38 can preferentially affect its activity in certain functions, but not others. Our results with EGL-38 suggest one of two possible models to explain the tissue-preferential effects. One possibility is that the different mutant alleles affect the ability of EGL-38 to bind to certain sequences, but not others. This model predicts that either there are one or a few key targets in each cell type, or that there are specific response element sequences that are associated with all the targets for EGL-38 in a particular cell type. An alternative model is that, although the mutations affect DNA binding, their tissue preferential effects result from their impact on EGL-38 function in combination with other proteins that influence target sequence recognition in vivo. This model predicts that there would be cell-specific co-factors that function with EGL-38 and participate in identifying which sequences correspond to an appropriate DNA binding site or that enhance EGL-38 transactivation activity. It also suggests that the interaction between the co-factor and EGL-38 would be mediated, or otherwise affected, by the EGL-38 DNA binding domain. Although we favor the second model, clearly distinguishing between the two will require
identification of proteins that act with EGL-38 and a large collection of cell-specific EGL-38 target genes.

5.4 The bromodomain protein LEX-1 acts with TAM-1 to modulate gene expression in C. elegans

In this work, I have genetically and molecularly characterized the *lex-1* gene in *C. elegans*. *lex-1* encodes a protein with a bromodomain and an ATPase domain similar to yeast Yta7, a protein important for maintaining chromatin states and repressing the inappropriate spreading of heterochromatin silencing. I found that *lex-1* shares phenotypic features with a previously described gene, *tam-1*, in that both genes affect expression from multicopy transgenes, and act as synMuv B genes. *lex-1* and *tam-1* are functionally different from other synMuv B genes, as the mutants do not exhibit germline features in somatic cells. I speculate LEX-1 and TAM-1, a RING-finger/B-box protein, are part of a complex that enhances transcription from repetitive DNA sequences by modulating chromatin structure.

5.4.1 Molecular and genetic characterization of the *lex-1* gene

Three alleles of *lex-1* were identified in a genetic screen for altered expression of a *lin-48::gfp* transgene. The *lex-1* gene encodes a protein with a bromodomain and an ATPase domain. LEX-1 shares sequence similarity with proteins found throughout
eukaryotes (Figure 4.4), including the Yta7 protein from yeast. Yta7 is part of a protein complex that associates with chromatin (Tackett et al. 2005). It co-purifies with histones that are enriched for distinct patterns of posttranslational modification, and functions to maintain the boundary between transcriptionally active and transcriptionally silenced regions of the chromosome (Jambunathan et al. 2005; Tackett et al. 2005). Altogether, the data suggest that Yta7-related proteins associate with particular chromosomal regions, and act to modulate or stabilize chromatin states. The data with LEX-1 are consistent with it functioning in a similar manner in *C. elegans*. For example, LEX-1 complexes may associate with transgene sequences and stabilize expression, counterbalancing the factors that act to repress transcription from repetitive sequences. According to this model, I propose that LEX-1 would normally act to enhance expression from (unknown) endogenous loci that are proximal to repetitive or heterochromatic sequences, and that the repetitive transgenes serve as a reporter for LEX-1 activity. However, it cannot rule out the alternative possibility that LEX-1 influences expression from repetitive transgenes through an indirect mechanism.

Other AAA+ ATPase proteins like LEX-1 multimerize into large complexes or arrays to direct translocation or remodeling of target substrates (reviewed by Erzberger and Berger 2006). I observed intragenic complementation between *lex-1(gu48)* and *lex-1(gu24)* alleles, a result consistent with the idea that LEX-1 proteins multimerize. However, it is interesting that these alleles are recessive to wild type (rather than antimorphic and dominant), and can compensate for each other (rather than interfere). This suggests that if LEX-1 is indeed part of a multimer complex, it is not necessary for all molecules in the complex to be fully functional.
5.4.2 LEX-1 and TAM-1 function together to modulate gene expression

Both *lex-1* and *tam-1* act as weak synMuv B genes, but do not exhibit ectopic P granules as do some other synMuv B mutants. The genetic interaction between *lex-1*(*gu48*) and *tam-1* mutants is also consistent with the idea that they function in a similar process in vivo. One hypothesis that accounts for the observed genetic interaction is that the two proteins interact with each other or are part of a complex, with LEX-1 responsible for recruiting TAM-1 to appropriate locations on the chromatin through its bromodomain. This hypothesis is favored over others (such as one in which maternal LEX-1 participates in the zygotic transcription of *tam-1*), as it accounts for the observed genetic interaction with the allele that affects the bromodomain (*gu48*), yet the absence of a similar interaction with the alleles that affect the ATPase domain (*gu24*) of the protein product, or splicing of the last exon (*gu47*).

The presence of a bromodomain and an ATPase domain in LEX-1 suggests it may recognize acetylated histones, and use the hydrolysis of ATP to alter chromatin structure. In other words, it may act as an interpreter of the histone code. I also speculate that LEX-1 is important for the recruitment of TAM-1 to chromatin. TAM-1 contains a RING finger and a B-box, domains that suggest it functions as an E3 ubiquitin protein ligase (reviewed by Joazeiro and Weissman 2000). Although a common role for E3 ligases is to attach the poly-ubiquitin chains that target proteins for proteasomal degradation, they can also mediate reversible mono- and poly-ubiquitinations that alter protein activity, subcellular localization, or interactions (reviewed by Schnell and Hicke 2003). For example, dynamic ubiquitination and de-ubiquitination of histone H2B is important for
gene transcription activation, and ubiquitination of histone H2A is associated with chromatin silencing (de Napoles et al. 2004; Henry et al. 2003). In the context of the genetic results, the favored model is that TAM-1 is an ubiquitin E3 ligase that modifies histones, and thereby influences chromatin structure and gene expression.

Although further experiments will be required to determine how TAM-1 and LEX-1 influence vulval development, results of this work predict that their synMuv effect would also be mes-4-independent. Sequence features and phenotypic analysis suggest TAM-1 and LEX-1 influence gene transcription. However, current data do not point to a clear transcriptional activation or repression mechanism. Indeed, in contrast to the germline/soma synMuv B proteins, LEX-1 and TAM-1 may act to regionally maintain transcriptional states or inhibit the spread of silencing, rather than to repress transcription. Thus LEX-1 and TAM-1 could act to promote the transcription from repetitive DNA in a direct, rather than indirect, manner.

5.5 CONCLUSION

In summary, results from this work demonstrate that gene transcription in development can be regulated through either sequence-specific DNA binding proteins that mediate gene-selective transcriptional repression or activation, or protein complexes that act to modulate the chromatin environment. In agreement of such a view, I provided additional evidence that egl-38 alleles with mutations that affect the DNA binding domain of the protein can preferentially affect egl-38 functions in different tissues. In
addition, the genetic and molecular analysis suggests that LEX-1 and TAM-1 function together to influence chromatin structure and to promote expression from repetitive sequences. Future genetic, biochemical and protein localization experiments should clarify the molecular relationship between LEX-1 and TAM-1, and their role in modulating gene expression. Molecular approaches, such as immunoprecipitation, would help to identify DNA-protein complexes composed of LEX-1 and TAM-1.

In future, with regard to the category 2 mutants, molecular cloning of these mutant alleles with enhanced and ectopic lin-48::gfp expression could provide a molecular understanding for how lin-48 is regulated in C. elegans. Since these mutant animals exhibit morphological and physiological defects, further study would reveal the roles of the responsible genes in animal development. One category 2 mutant, gu36, was identified recently as an allele of cul-2 (Xiaodong Wang, unpublished). The cullin CUL-2 is a crucial component of a subclass of multisubunit cullin-RING ubiquitin-ligases (E3s). The specificity of CUL-2-based complexes is provided by variable substrate-recognition subunits that bind to specific substrates (Petroski and Deshaies, 2005). In Caenorhabditis elegans, CUL-2 regulates several key processes in cell division and embryonic development, including meiotic progression, anterior–posterior polarity and mitotic chromatin condensation (Petroski and Deshaies, 2005). Moreover, a recent study showed that CUL-2 ubiquitin ligase degrades TRA-1 to regulate C. elegans sex determination (Starostina et al., 2007). Further experiments should clarify the whether the gu36 mutation in CUL-2 disrupts the ubiquitin-dependent degradation pathway and in turn causes pleiotropic defects in gu36 mutant animals.
LIST OF REFERENCES


