EGL-38 Pax regulates the *ovo*-related gene *lin-48* during *Caenorhabditis elegans* organ development

Andrew D. Johnson[1], Daniel Fitzsimmons[2], James Hagman[2,3] and Helen M. Chamberlin[1,*]

[1]Department of Molecular Genetics, Ohio State University, Columbus, OH 43210, USA
[2]Department of Immunology, University of Colorado Health Sciences Center, Denver, CO 80262, USA
[3]Department of Immunology, National Jewish Medical and Research Center, Denver, CO 80206, USA
*Author for correspondence (e-mail: chamberlin.27@osu.edu)

SUMMARY

The Pax gene *egl-38* plays an important role in the development of several organs in *C. elegans*. To understand how a Pax transcription factor influences distinct developmental choices in different cells and tissue types, we have characterized a second gene, *lin-48*. *lin-48* functions with *egl-38* in the development of one structure, the hindgut, but not in other tissues such as the egg-laying system. We show that *lin-48* encodes a C2H2 zinc-finger protein that is similar to the product of the *Drosophila* gene *ovo* and is expressed in the hindgut cells that develop abnormally in *lin-48* mutants. We present evidence that *lin-48* is a target for EGL-38 in hindgut cells. We show that *lin-48* requires *egl-38* for its expression in the hindgut.

Using deletion analysis, we have identified two elements in the *lin-48* promoter that are necessary for *lin-48* expression. We demonstrate that EGL-38 binds with high affinity to one of these elements. In addition, we have observed genetic interactions between mutations in the *lin-48* promoter and specific alleles of *egl-38*. These experiments demonstrate a functional link between Pax and Ovo transcription factors, and provide a model for how Pax transcription factors can regulate different target genes in different cells.

Key words: Organogenesis, Pax2, Paired domain, *C. elegans*

INTRODUCTION

Pax transcription factors play an important role in organogenesis during animal development. In mammals, genes of the *Pax2*, *Pax5* and *Pax8* (*Pax2/5/8*) subclass play a role in the development of cell types and organs including B lymphocytes, the kidney, the thyroid, the eye and the brain (Urbánek et al., 1994; Torres et al., 1995; Schwartz et al., 1997; Mansouri et al., 1998). Pax factors influence development by binding DNA in a sequence-specific manner and affecting the expression of target genes. The DNA-binding domain of Pax transcription factors (the paired domain) is composed of two helix-turn-helix domains that can interact with DNA across 20 base pairs (Czerny et al., 1993; Xu et al., 1999). Pax proteins can bind DNA as a monomer, but they can also act in combination with other transcription factors (Fitzsimmons et al, 1996). Owing to the relatively large size of the DNA-binding domain and the possibility for altered binding properties when complexed with other proteins, the consensus DNA sequence for Pax-responsive elements is notably degenerate (Czerny et al., 1993). This suggests that a wide range of sites within a genome are capable of binding Pax proteins, but that cellular context may influence Pax DNA binding and activity in vivo. Differences in cellular context may include other proteins that interact with Pax proteins to restrict or enhance their DNA binding, or which modulate their

functional properties when bound to DNA (Eberhard et al., 2000). To better understand how a single factor recognizes different targets in different cells, it is essential first to identify tissue-specific targets and then to characterize the regulation of these target genes.

The *C. elegans* gene *egl-38* encodes a Pax transcription factor that is most similar to the mammalian Pax2/5/8 subclass of factors (Chamberlin et al., 1997; Czerny et al., 1997). *egl-38* is an essential gene, and mutants that bear a strong reduction-of-function allele die as embryos or soon after hatching (Chamberlin et al., 1997). Analysis of three non-null alleles has permitted characterization of additional *egl-38* functions in patterning of cell types during development of the hindgut (rectal epithelium), the egg-laying system and the spicules of the male tail. Genetically, these three alleles preferentially disrupt different functions of *egl-38*. Each allele corresponds to a different missense mutation that affects the DNA binding domain of EGL-38. The localization of these tissue-preferential mutations to the DNA binding domain suggests a model in which alterations of the DNA-binding properties of EGL-38 have different consequences in different tissues, i.e. these mutations preferentially affect the ability of EGL-38 to bind to and regulate certain targets and not others.

To better understand how a Pax transcription factor might affect the expression of different genes in different tissues, we have characterized a second gene that functions with *egl-38* in

Table 1. *lin-48* shares a subset of *egl-38* functions

Gene function	Gene		Refs
	egl-38	*lin-48*	
Hindgut development	U, F, K′ specification	U, F, K′ specification	Chamberlin et al., 1997; 1999
Egg-laying system	Vulval expression of *lin-3* (and uv1 specification)	None	Chamberlin et al., 1997; Chang et al., 1999
Viability	Essential	Not essential	Chamberlin et al., 1997; 1999
Male spicule development	Yes*	Yes*	Chamberlin et al., 1999; Jiang and Sternberg, 1999; H. M. C., unpublished

*egl-38 has additional functions in male spicule development not shared with *lin-48*.

the development of the hindgut: *lin-48*. 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 (Fig. 1). However, *egl-38* and *lin-48* are functionally distinct (Table 1). This analysis suggest *lin-48* function is associated with *egl-38* in the development of the hindgut, but not in other cell types. To investigate the functional relationship between *egl-38* and *lin-48*, we initiated a molecular analysis of *lin-48*. We report that *lin-48* encodes a C2H2 zinc-finger protein similar to the product of the *Drosophila ovo* gene. Our results indicate LIN-48 is localized to nuclei, and required for the specification of specific cell types as are *Drosophila* and mammalian OVO. We show that there are at least two important regulatory elements in the *lin-48* promoter, and that EGL-38 can bind specifically to one of these elements in vitro. In addition, the different mutant alleles of *egl-38* exhibit allele-preferential sensitivity to mutations in the *lin-48* promoter. Taken together, these results identify *lin-48* as a tissue-restricted target for EGL-38, and provide the first evidence for a direct relationship between Pax factors and *ovo* genes.

MATERIALS AND METHODS

Nematode strains

Nematode strains were cultured according to standard techniques (Sulston and Hodgkin, 1988). Mutations used are described by Hodgkin (Hodgkin, 1997), and are as noted.

Linkage group (LG) III: *egl-5(sy279)*; *lin-48(sa469), lin-48(sy234), lin-48(sy548)* (Chamberlin et al., 1999; Jiang and Sternberg, 1999); *unc-119(e2498)*.

LG IV: *egl-38(n578), egl-38(sy287), egl-38(sy294), egl-38(s1775)* (Chamberlin et al., 1999).

LG V: *him-5(e1490)*.

Molecular cloning of *lin-48*

lin-48 was mapped to LG III between *unc-93* and *dpy-17* (Chamberlin et al., 1999). Cosmids and DNA sequence from this genomic region were provided by Alan Coulson and the *C. elegans* sequencing consortium. DNA was microinjected into the mitotic germline of hermaphrodites according to the method of Mello et al. (Mello et al., 1991). 100 ng/μl of plasmid containing the *rol-6(su1006)* allele (pRF4) was co-injected as a marker with 1-10 ng/μl of test DNA into *lin-48(sa469); him-5(e1490)* animals. Heritable lines were tested by assaying Rol males for rescue to wild-type tail morphology. Transgenes containing the cosmid F34D10 rescued *lin-48* in two out of two heritable lines. Subclones of the cosmid were prepared using standard methods (Ausubel et al., 2000). pTJ972 (Fig. 2A) is a 10.5 kb *Bam*HI subclone from F34D10 into pBluescript (Stratagene) that rescued *lin-48(sa469)* in three out of three heritable lines. We sequenced the DNA of *lin-48* mutants as described previously (Chamberlin et al., 1997). We used BLAST 2.0 (http://www.ncbi.nlm.nih.gov/BLAST/) to identify and evaluate the molecular homologs of LIN-48 and ClustalW 1.7 (http://dot.imgen.bcm.tmc.edu:9331/) and Boxshade 3.21 (http://www.isrec.isb-sib.ch:8080/software/BOX_form.html) to align and display the zinc-finger domain in Fig. 2B. We used RT-PCR to characterize *lin-48* cDNA. Our cDNA sequence results match those of Schonbaum, Fantes and Mahowald (GenBank Accession Number, AF134806), and indicate that *lin-48* is *trans*-spliced to SL1.

Construction and analysis of *lin-48::gfp* transgenes

We made GFP reporter constructs for *lin-48* expression using GFP (pPD) vectors provided by Andy Fire. Upstream deletions were generated by designing forward PCR primers corresponding to different positions in the *lin-48* promoter region. Point mutations were generated using a two-step, PCR-based method (Ausubel et al., 2000). The cloned PCR fragments were sequenced to verify. Transgenes for all GFP reporter constructs were produced by microinjection of plasmid DNA into animals, as described above. For all experiments that test expression of *lin-48* promoter constructs, 40-50 ng/μl reporter clone was co-injected with 15 ng/μl pDP#MM016 (*unc-119(+)* plasmid; Maduro and Pilgrim, 1995) into *unc-119(e2498); him-5(e1490)* animals.

The activity of each transgene was assessed in animals from heritable transgenic lines. Larvae were anesthetized on pads of 5% agar containing 5 mM sodium azide, and scored for sex, larval stage and GFP expression at 1000×. For almost all experimental conditions, the expression of at least two independently derived transgenes was tested. For critical transgenes, two independently isolated DNA clones were also tested. For the data in Figs 4 and 6, L1 and L2 animals of both sexes were scored. Each hindgut cell was scored for expression, resulting in four hindgut cells scored for each animal. Transgenic animals were verified by confirming expression of GFP in at least one cell in the animal before scoring.

EGL-38 protein expression and EMSA

For production of recombinant EGL-38 DNA-binding domain (EGL-38 DBD) in *Escherichia coli*, nucleotide sequences comprising amino acids 22-156 (Chamberlin et al., 1997) were amplified using PCR and cloned into pET11a vector. Expression of EGL-38 DBD was induced in *E. coli* strain BL21lysSDE3 cells. Bacterial lysate proteins were prepared as described previously for Pax5 (Wheat et al., 1999). The integrity of the EGL-38 protein was confirmed with SDS-PAGE.

EMSA was performed essentially as described in Wheat et al. (Wheat et al., 1999). For DNA probes, oligonucleotides (Integrated DNA Technologies, Coralville, IA) 5′TCGACGGTGCATTTATGAA-GCGTGACGGTAAGC and 5′TCGAGCTTACCGTCACGCTT-CATAAATGCACCG, or 5′TCGAGCAGACACCCATGGTTGA-GTGCCCTCCAGG and 5′TCGACCTGGAGGGCACTCAACC-ATGGGTGTCTGC were annealed to make the lre2 or CD19 probes, respectively. Labeling of double stranded oligonucleotide probes with ^{32}P, probe purification and preparation of competitor oligonucleotides have been reported previously (Fitzsimmons et al., 1996). Competitor oligonucleotides included 5′TCGAGAAAGGCGCAAGTTTGCGG-TGCGCGATTG and 5′TCGACAATCGCGCACCGCAAACTT-GCGCCTTTC (lre1), 5′CGGTGCATTTATGAAGCGTGACGG-

TAAG and 5′CTTACCGTCACGCTTCATAAATGCACCG (lre2), and 5′TCGAGATCCTTCTGGGAATTCCTAGATC and 5′TCGA-GATCTAGGAATTCCCAGAAGGATC (STAT3). CD19 competitor was identical to CD19 probe oligonucleotides.

RESULTS

Summary of *egl-38* and *lin-48* function in *C. elegans* hindgut development

We identified alleles of both *lin-48* and *egl-38* in a genetic screen for genes important in the development of the *C. elegans* hindgut (Chamberlin et al., 1999). The *C. elegans* hindgut can serve as a simple model for animal organ development. It is composed of only 11 cells, representing eight distinct cell types (Fig. 1A; Sulston et al., 1983). Development of the *C. elegans* digestive system and differentiation of hindgut cell types occurs during embryogenesis. However, the hindgut cells also contribute to larval development as one (in hermaphrodite animals) or five (in male animals) of the cells undergo further stereotypic postembryonic cell divisions (Sulston and Horvitz, 1977; Sulston et al., 1980). We have used a combination of postembryonic cell lineage analysis and analysis of the production of differentiated cell types to interpret the patterning defects in the hindgut of *egl-38* and *lin-48* mutant animals (Chamberlin et al., 1997; Chamberlin et al., 1999). This analysis shows that *egl-38* and *lin-48* are each required for the normal specification of cell type for three cells in the middle of the hindgut (U,F and K′), and to make these cells different from other cells in the hindgut (Fig. 1B, C).
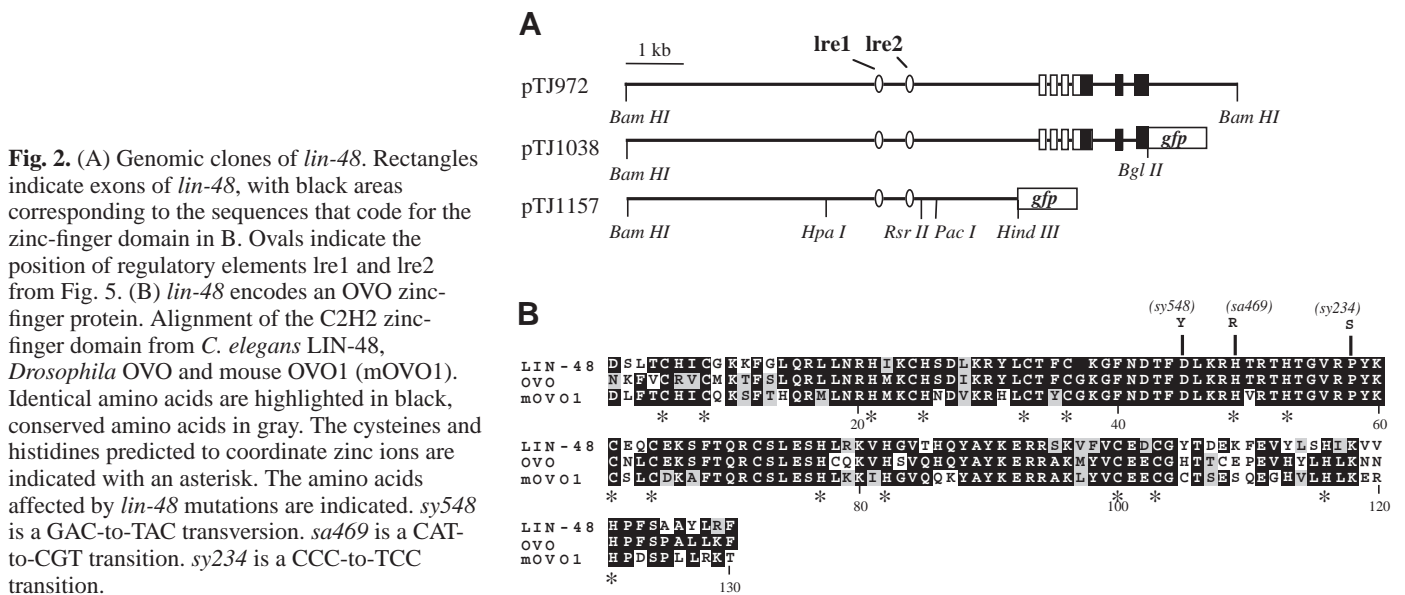
lin-48 encodes a zinc finger protein similar to Drosophila OVO

We used DNA transformation rescue and the sequence of DNA from mutant animals to identify the gene F34D10.5 as *lin-48* (see Materials and Methods). *lin-48* encodes a protein that contains four C2H2 zinc-finger repeats (Fig. 2B). Within the 130 amino acids of the zinc-finger domain, LIN-48 is 73% identical to *Drosophila* OVO and 66% identical to mouse

A. wild type

B. *egl-38*

C. *lin-48*

Fig. 1. *egl-38* and *lin-48* affect the development of cells in the middle region of the *C. elegans* hindgut (indicated with a shaded box). (A) Wild-type pattern of hindgut cell types. (B) In *egl-38* mutants, the presumptive F and U cells develop like their posterior neighbors, B and Y (Chamberlin et al., 1997). The presumptive K′ cell can develop like its sibling and neighbor, K. (C) In *lin-48* mutants, the presumptive U cell can develop like its neighbor and lineal homolog, B. The presumptive F cell develops abnormally (asterisk; Chamberlin et al., 1999). The presumptive K′ cell can develop like its sibling and neighbor, K. Int, intestine; irv, intestinal rectal valve cells (also called vir); rep, rectal epithelial cells. Anterior is towards the left, dorsal is upwards.
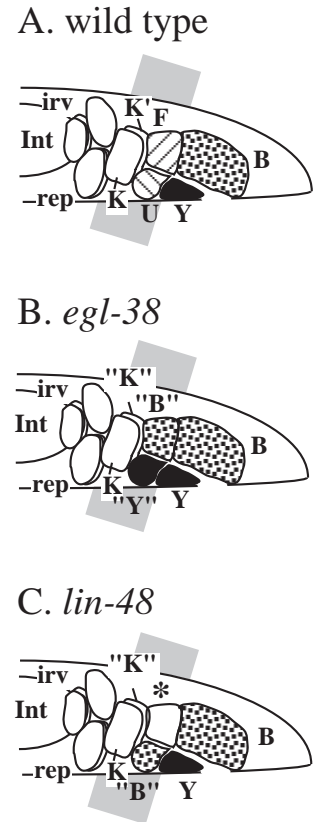
mOVO1 (Mevel-Ninio et al., 1991; Dai et al., 1998). OVO proteins in *Drosophila* and mouse act as transcription factors and play an important role in the development of several distinct cell types (Oliver et al., 1987; Payre et al., 1999; Dai et al., 1998). *Drosophila* OVO binds DNA in a sequence-specific 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). We characterized the lesions associated with three *lin-48* mutations. Although all are

Fig. 2. (A) Genomic clones of *lin-48*. Rectangles indicate exons of *lin-48*, with black areas corresponding to the sequences that code for the zinc-finger domain in B. Ovals indicate the position of regulatory elements lre1 and lre2 from Fig. 5. (B) *lin-48* encodes an OVO zinc-finger protein. Alignment of the C2H2 zinc-finger domain from *C. elegans* LIN-48, *Drosophila* OVO and mouse OVO1 (mOVO1). Identical amino acids are highlighted in black, conserved amino acids in gray. The cysteines and histidines predicted to coordinate zinc ions are indicated with an asterisk. The amino acids affected by *lin-48* mutations are indicated. *sy548* is a GAC-to-TAC transversion. *sa469* is a CAT-to-CGT transition. *sy234* is a CCC-to-TCC transition.

**Fig. 3.** Expression of *lin-48*. (A,C,E,G) Nomarski DIC micrographs of L1 larvae. (B,D,F,H) Epi-fluorescent micrographs of the same animals. The animal in A and B bears a transgene containing pTJ1038, whereas the animals in C-H bear transgenes containing pTJ1157 or a mutant derivative (see Fig. 2A). (A,B) LIN-48 tagged with GFP is localized to cell nuclei. Nuclei of F and U are indicated. The location of K and K′ is marked; the nuclei of these cells are out of the plane of focus. Expression of LIN-48::GFP transgenes is at low levels, and the F cell is expressing at a slightly lower level than the other cells in this particular male larva. (C,D) Expression pattern of *lin-48::gfp* in wild type. The location of the hindgut and the excretory duct cell are indicated with an arrow and an arrowhead, respectively. Expression is observed in the U, F, K and K′ hindgut cells, the excretory duct cell and several cells in the head in the larva. Phasmid cell expression does not initiate until late L1 stage. (E,F) Expression pattern of *lin-48::gfp* transgenes that include mutations in both lre1 and lre2. These mutations disrupt expression in the hindgut cells, and excretory duct expression can be reduced. (G,H) Expression pattern of intact *lin-48::gfp* transgenes in an *egl-38(sy294)* mutant animal. The mutant animal fails to express *lin-48::gfp* in hindgut cells. Scale bars: in A, 20 μm for A,B; in C, 20 μm for C-H.
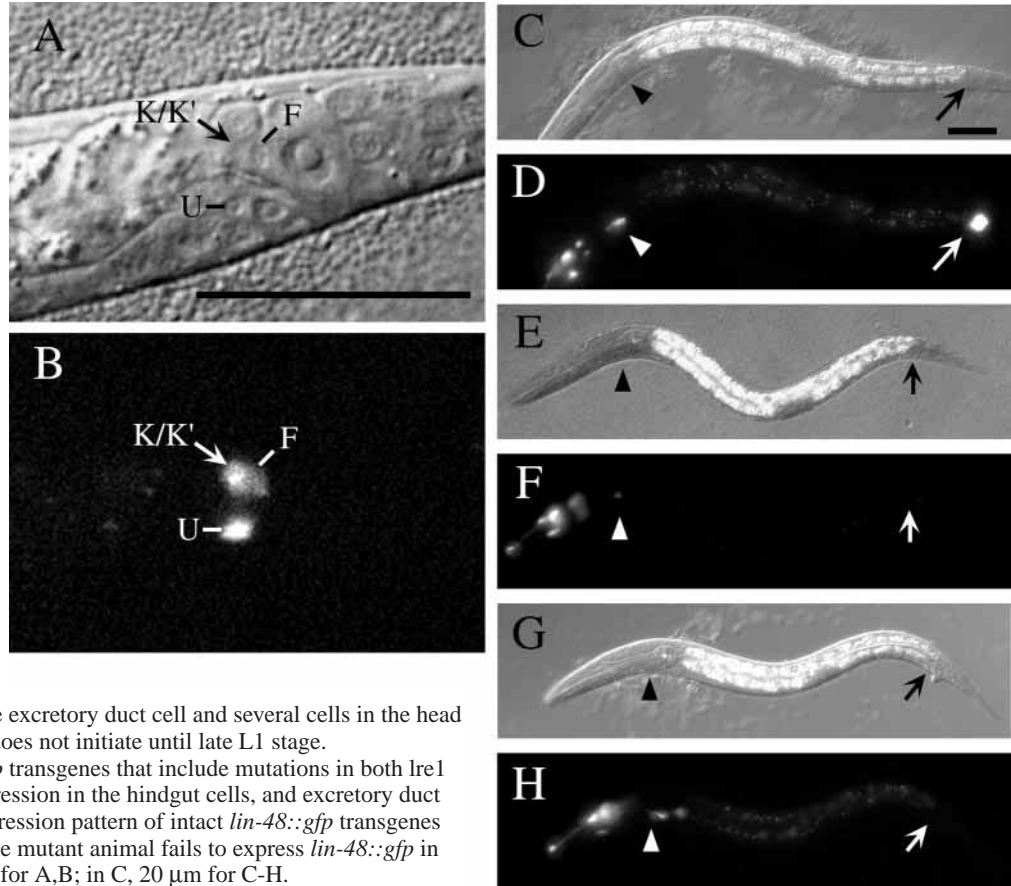
missense mutations, each mutation is recessive and the mutant phenotypes are not enhanced when the alleles are tested in *trans* to a deficiency (Chamberlin et al., 1999). Thus, we believe these mutations are reduction- or loss-of-function alleles.

lin-48 is expressed in hindgut cells

To investigate the expression pattern of *lin-48*, we created a reporter in which the last two codons of *lin-48* were replaced with sequences encoding the green fluorescent protein (GFP; pTJ1038; Fig. 2A). When expressed in mutant animals, these transgenes are capable of rescuing *lin-48*. In the hindgut, these transgenes are expressed in U, F, K′ and K cells (Fig. 3A,B). As the development of U, F and K′ is affected in *lin-48* mutants, this expression pattern is consistent with *lin-48* acting directly within the cells that express the gene. The significance of *lin-48::gfp* expression in K is not clear, although this expression is affected in the same manner as the other hindgut expression in our experimental analysis (see below). In addition to hindgut cells, *lin-48::gfp* is expressed in the excretory duct cell, neuronal support cells of the phasmid and labial sensory structures and a small number of additional unidentified cells in the head. Male animals exhibit additional expression in the developing tail structures (data not shown). *lin-48::gfp* expression is initiated in late embryogenesis and persists into adulthood. The chimeric LIN-48::GFP protein is localized to the nuclei of expressing cells, consistent with the idea that OVO-related proteins like LIN-48 function as transcription factors. A second reporter construct that includes only the *lin-48* upstream regions (pTJ1157; Fig. 2A) expresses in the same pattern as the full-length transgene, but is expressed at much higher levels (Fig. 3C, D).

The lin-48 promoter contains DNA elements necessary for its expression

To investigate the regulation of *lin-48*, we performed a deletion analysis of the *lin-48* promoter. We generated a series of clones that progressively deleted more upstream regions of the *lin-48* promoter between positions 0 and 5205 (Fig. 4B-F; we defined the *Bam*HI site upstream of *lin-48* as 0 following the standard of Okkema et al. (Okkema et al., 1993)). When tested for expression in transgenic animals, these transcriptional reporters identified a domain between 4697 and 4892 required for normal expression of *lin-48* in all four hindgut cells and the excretory duct cell. Reporters that included sequences 4697 and upstream were expressed in hindgut cells, the excretory duct and other cells (head and phasmid), whereas clones including 4892 or less were expressed only in head and phasmid. These experiments also identified a region upstream of 2846 necessary for expression in many of the head cells, as the full-length reporter is expressed in up to ten head cells, whereas the truncated reporters generally express in only two cells. As no mutant reporter transgenes in either wild-type or mutant animals failed to express in the two head cells and the two phasmid cells, we used expression in these cells as an internal control for expression level and presence of the transgene in all subsequent experiments.

To identify potential regulatory elements in the *lin-48* promoter, we inspected the sequence between 4697 and 4892,

and identified a domain with similarity to mammalian DNA elements that bind Pax proteins. As our genetic results indicated *lin-48* requires the Pax gene *egl-38* for its expression in hindgut cells (see below), we mutated this site in a reporter that included sequences to 4697, and found that the mutant transgenes fail to express in hindgut and excretory duct cells (Fig. 4G). We define this site as lre2 (*lin-48* regulatory element 2; Fig. 5B), as we ultimately identified two elements in the *lin-48* promoter. Although truncated reporters are sensitive to mutations in lre2, we found that expression is restored when the site is mutated in a full-length reporter (Fig. 4M). This result suggests an additional element(s) is present upstream of 4697 that is capable of promoting expression of *lin-48* in the hindgut. To identify this element, we performed a second deletion analysis of the *lin-48* promoter, starting with DNA that included the mutant form of lre2. These reporters identified a second region between positions 4118 and 4191 required for expression of *lin-48* in hindgut cells (Fig. 4H-K). We inspected the sequence between 4118 and 4191, and identified another domain with moderate similarity to mammalian DNA elements that bind Pax proteins. We define this site as lre1. We mutated this site in a reporter including sequences to 3985 and containing the mutant form of lre2, and found that the mutant transgenes fail to express in hindgut cells (Fig. 4L). As recombinant EGL-38 does not bind this site with high affinity (see below), it is possible that the similarity of this element to Pax-binding sites is coincidental. Nevertheless, the mutation analysis identifies it as an important regulatory element. To confirm the importance of the two sites, we created full-length reporter transgenes bearing mutations in both lre2 and lre1, and found that these two mutations effectively eliminated expression of *lin-48::gfp* in hindgut cells (Figs 3E,F, 4O).

lin-48 requires *egl-38* for its expression in hindgut cells

To investigate the relationship between *lin-48* and *egl-38*, we characterized *lin-48* expression in *egl-38* mutant backgrounds. Expression of *lin-48::gfp* in the hindgut cells of *egl-38(s1775)* and *egl-38(sy294)* mutants is eliminated, although expression in other cell types is maintained (Figs 3G,H, 6B,C). This result indicates that *lin-48* requires *egl-38* for its expression in the hindgut. However, hindgut expression of *lin-48::gfp* is still observed in *egl-38(sy287)* and *egl-38(n578)* mutants (Fig. 6D,E). Taken together, these results provide a link between the activity associated with different *egl-38* alleles in different developmental processes and their ability to activate *lin-48* expression, as genetic tests indicate *s1775* and *sy294* strongly disrupt *egl-38* activity in

hindgut development, whereas *sy287* and *n578* disrupt this function to a lesser extent (Table 2; Chamberlin et al., 1997). To further investigate the relationship between *egl-38* genotype and *lin-48* expression, we tested the expression of transgenes bearing mutations in lre1 or lre2 in *egl-38(sy287)* and *egl-38(n578)* mutants (Fig. 6F-K). *egl-38(sy287)* mutants exhibit moderate defects in hindgut development, these animals exhibit reduced expression of *lin-48::gfp* in hindgut cells, and this expression is sensitive to a single mutation in lre1 or lre2. In contrast, *egl-38(n578)* mutants exhibit only minor defects in hindgut development, and they still express mutant transgenes, although at reduced levels. This sensitivity of mutant *lin-48* transgenes to different *egl-38* genotypes provides further support for an in vivo relationship between EGL-38 and the promoter of *lin-48*.

To investigate the specificity of the relationship between *egl-38* and *lin-48*, we characterized *lin-48::gfp* expression in *egl-5* mutants (Fig. 6L). *egl-5* is the *C. elegans* posterior HOM-C gene required for regional specification of cells in the hindgut (Chisholm, 1991). We find that *lin-48::gfp* expression is normal in *egl-5* mutants. This indicates that not all genes required for normal development of hindgut cells affect *lin-48* expression. Finally, we tested *lin-48::gfp* expression in *lin-48* mutants, and found it to be unaltered (Fig. 6M). This result indicates there is no positive autoregulation of *lin-48*, and
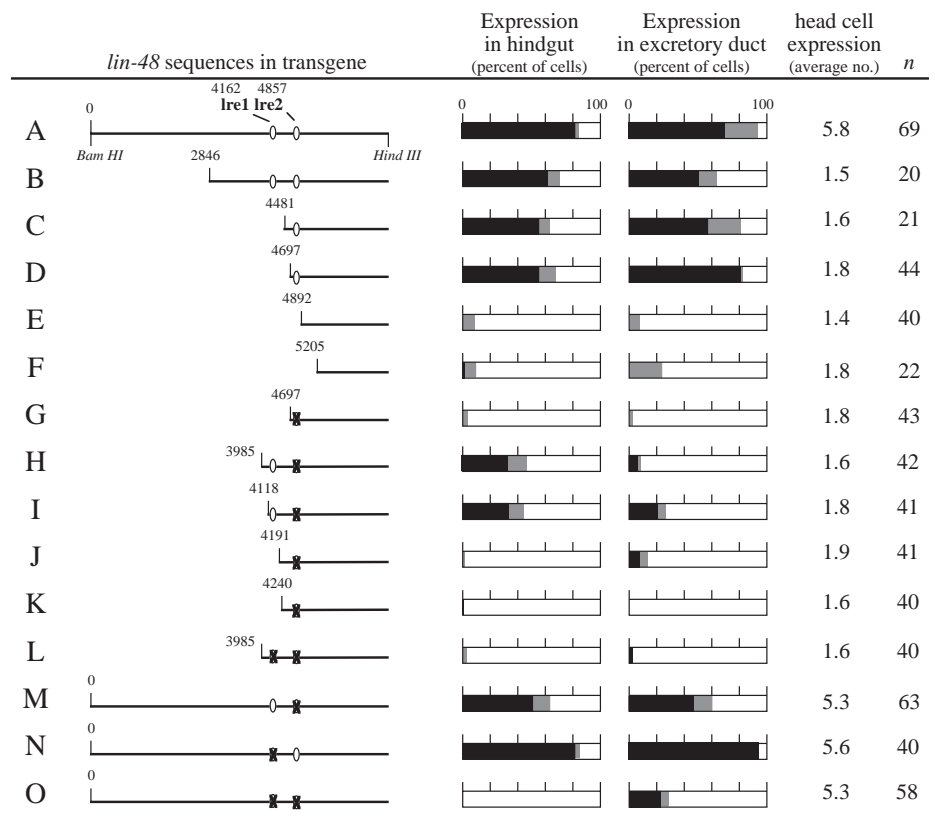
Fig. 4. The *lin-48* promoter contains two redundant elements important for *lin-48* expression in hindgut cells. *lin-48* sequences were tested for their ability to drive expression of GFP, and are diagrammed using the conventions of Fig. 2A. Transgenes containing the mutant sequences illustrated in Fig. 5 are indicated with X. The percentages of cells expressing GFP (black bar), expressing very low but detectable levels of GFP (gray bar) or not expressing GFP (white bar) are indicated for each construct. *n*, number of animals scored for expression.

A.

```
A.  H2B-2.2     TCAATACGCAACGAAGCGAAGACGGACC
    PRS-5       AGCGAGATATCTAGAGCGGAACGGTTCC
    5'Sγ2a      GATCAGAATTGTGAAGCGTGACCATAGA
    CD19-1      CAGACACCCATGGTTGAGTGCCCTCCAG
    CD19-2      AGAATGGGGCCTGAGGCGTGACCACCGC

                A                 GA  GG
    consensus   .....G..CA.TG..GCGTGACCA....
```

B.

```
B.  lin-48 lre2      CGGTGCATTTATGAAGCGTGACGGTAAG
    lin-48 lre2 mut  CGGTGCATTTATTCTAGATAGTGGTAAG
```

C.

```
C.  lin-48 lre1      GAAAGGCGCAAGTTTGCGGTGCGCGATT
    lin-48 lre1 mut  GAAAGGCGCAAGTTATCTAGAGTGAATT
```

Fig. 5. The *lin-48* promoter elements share similarity with Pax-binding sites. (A) Some representative binding sites for mammalian PAX5, and a consensus sequence (Czerny et al, 1993). (B,C) Sequence of the two elements in the *lin-48* promoter, and the mutant elements used to test the function of each site in vivo. Sequences that correspond to the consensus are underlined. Nucleic acids changed in the mutant promoters are in italics.

provides indirect evidence that LIN-48 does not reciprocally affect the expression of *egl-38*.

EGL-38 binds to the *lin-48* promoter

Our genetic studies suggest that EGL-38 may regulate *lin-48* transcription directly by binding to regulatory elements that may include lre1 and/or lre2. To assess whether these sites include recognition sequences for EGL-38, we expressed the DNA-binding domain (DBD) of the protein in *E. coli* and

Table 2. Summary of the functions associated with different *egl-38* alleles

Function‡	Allele*				
	+	*n578*	*sy287*	*sy294*	*s1775*
Hindgut development	+++	+++	++	−	−
Egg-laying	+++	−	+++	++	−
Viability	+++	+++	++	+	−
Male spicule development	+++	++	−	−	−
Activation of *lin-48::gfp*	+++	+++	+	−	−
Activation of lre1 mutant *lin-48::gfp*	+++	+	−	n.d.	n.d.
Activation of lre2 mutant *lin-48::gfp*	++	+	−	n.d.	n.d.

s1775 is a lethal allele that behaves as a genetic null. *n578*, *sy287*, and *sy294* are all non-null, reduction of function alleles that preferentially disrupt different functions of *egl-38*.

‡Data for developmental functions of *egl-38* alleles are summarized (Chamberlin et al., 1997; Chamberlin et al., 1999; H. M. C., unpublished). In these experiments each allele was tested for each function in *trans* to the other alleles and in *trans* to a deficiency that deletes the gene.

n.d., not determined.

tested its DNA-binding abilities in vitro in an electrophoretic mobility shift assay (EMSA; see Materials and Methods). It was shown previously that promoter sequences of the murine CD19 gene include a site that binds proteins of the Pax2/5/8 family, including EGL-38, with high affinity (Czerny et al., 1997). In initial experiments, EGL-38 DBD binding was evidenced by detection of a single band with CD19 control or lre2 probe DNAs at very great (up to 100,000-fold) dilutions of bacterial lysate, but was detected only weakly using the lre1 probe (undetectable at greater than 1:15 dilution; data not shown). Binding was not detected using control lysate from *E.*
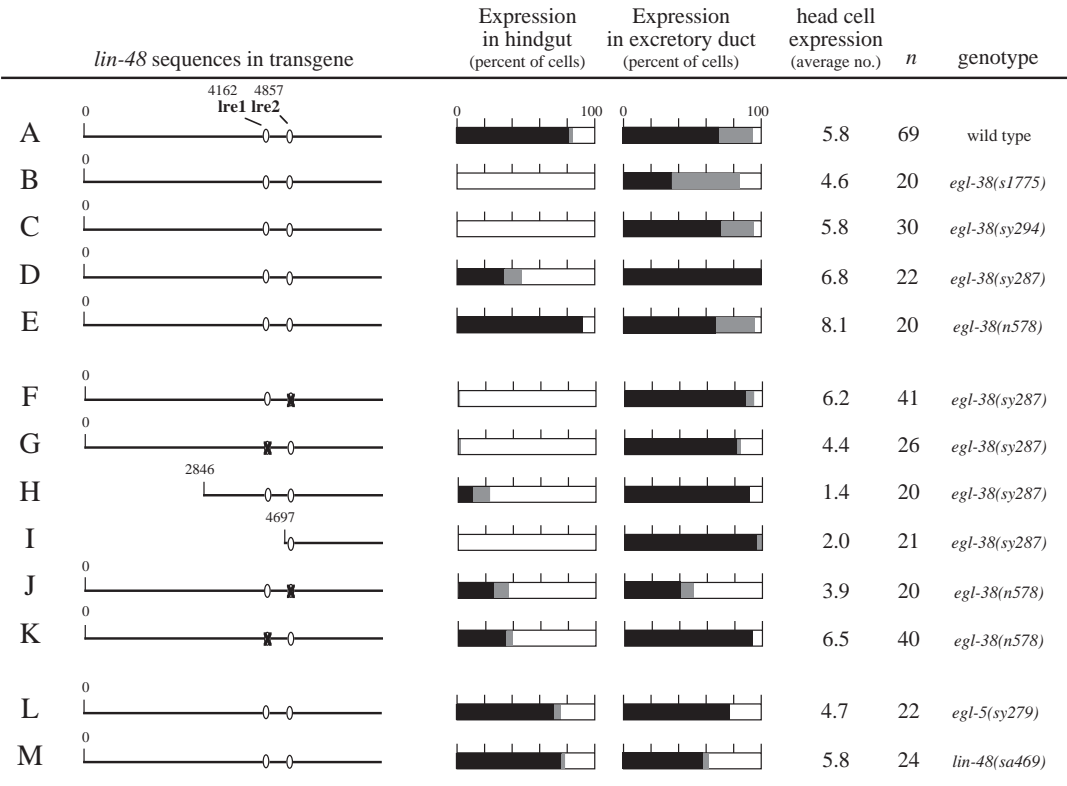
Fig. 6. Mutations in *egl-38* affect the expression of *lin-48* in hindgut cells. Data presented as in Fig. 4.
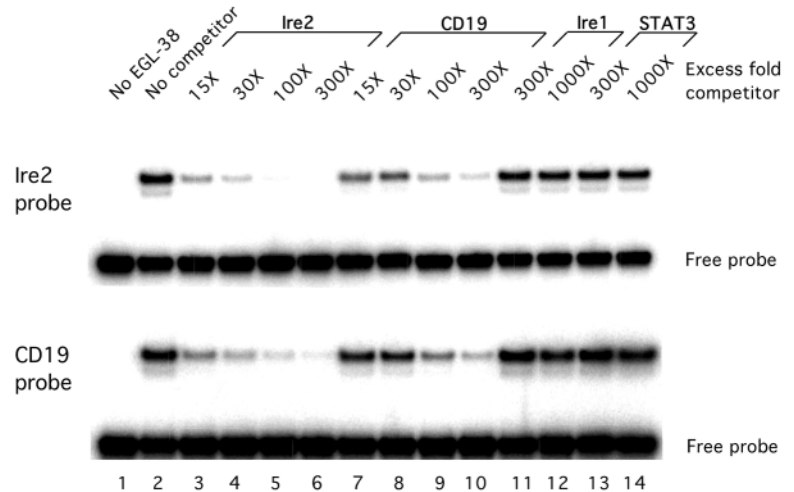
Fig. 7. The DNA-binding domain of EGL-38 specifically binds the *lin-48* lre2 site with high affinity. *E. coli* lysate containing EGL-38 DBD was incubated with the lre2 probe at a final dilution of 1:51,200 (top), or with the CD19 control probe at 1:6400 (bottom). Competitor sequences are described in Materials and Methods. The CD19 probe corresponds to the CD19-1 element of Fig. 5.

coli containing empty expression vector. To demonstrate the specificity of these interactions, we incubated binding reactions in the absence or presence of excess double-stranded competitor oligonucleotides (Fig. 7). EGL-38 DBD binds the lre2 probe with high affinity, as evidenced by the efficient competition of binding to the lre2 by low levels of unlabeled competitor oligonucleotides (Fig. 7, lanes 3-6, top). Similarly, binding of EGL-38 DBD to the CD19 probe (lower panel) was efficiently competed by low levels of lre2 competitor (Fig. 7, lanes 3-6, bottom). As expected, binding to each probe was competed, although less efficiently, by excess CD19 competitor (Fig. 7, lanes 7-10). Competition was not detected using lre1 DNA, or by control STAT3-binding sites, even when added at 1000-fold molar excess (Fig. 7, lanes 11-14). Together, these data show that EGL-38 specifically binds the lre2 site in vitro with relatively high affinity.

DISCUSSION

The role of *lin-48 ovo* in *C. elegans* development

In this paper, we report the characterization of the *C. elegans ovo*-related gene *lin-48* and provide evidence that it is a direct target for the EGL-38 Pax transcription factor in hindgut cells. *lin-48* and *egl-38* are functionally similar in that they affect the development of the same subset of hindgut cells (Fig. 1). However, their specific hindgut functions differ. For example, the presumptive U cell behaves like its posterior neighbor Y in *egl-38* mutants, whereas it behaves like its right side lineal homolog B in *lin-48* mutants. Anterior/posterior, dorsal/ventral, and left/right patterning information plays a role in hindgut development (Chamberlin et al., 1999; Wollard and Hodgkin, 2000). Our experiments provide a link between genes important in anterior/posterior (*egl-38*) and left/right (*lin-48*) patterning. Furthermore, these experiments suggest that *egl-38* and *lin-48* are not part of a strictly linear pathway, as the function(s) of *egl-38* and *lin-48* in the hindgut are not identical. We speculate that *lin-48* is only one of the hindgut targets of EGL-38, and it mediates a subset of functions.

   lin-48 is expressed in a small number of cells in addition to the hindgut cells. Further work will be necessary to clarify the function of *lin-48* in these cells. In particular, it will be

interesting to investigate the potential role of *lin-48* in the development of the excretory duct cell. The excretory system is proposed to mediate osmotic regulation, and the excretory duct cell is essential for viability (Nelson and Riddle, 1984). As *lin-48* mutants are viable, *lin-48* can not be essential for excretory duct cell development or differentiation. However, *lin-48* mutant stocks exhibit a low but reproducible level of lethality, and the inviable animals die around hatching with the characteristics of animals that lack a functional excretory system (Chamberlin et al., 1999). Thus *lin-48* may play a role in excretory duct development, but its function may be compensated by another gene in most animals. 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 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, and *lin-48* plays a role in development of the hindgut (which develops into the adult male cloaca) and potentially the excretory system.

The functional relationship between Pax factors and *ovo* genes

Our experiments indicate *lin-48* is a direct target for EGL-38 in *C. elegans*. A direct link between Pax factors and *ovo* genes has not been previously reported. However, genetic parallels in mammals indicate the potential for a conserved functional relationship between these classes of genes. In vertebrates, the *Pax2* gene is essential for development of kidney, brain and ear (Torres et al., 1995; Torres et al., 1996; Schwartz et al., 1997), and the *Pax8* gene plays a role in thyroid and kidney development (Mansouri et al., 1998; Carroll and Vize, 1999). Mouse *Ovo1* is expressed abundantly in the kidney, and is required for its normal differentiation (Dai et al., 1998). Thus, as in *C. elegans*, *Ovo1* acts in a subset of the cells that require Pax2/5/8 factors. Future experiments will be required to test whether *Ovo1* is a target for Pax2 or Pax8 during kidney development. As all of the functions of the *Drosophila*

Pax2/5/8 gene *sparkling* (*shaven*) have not been characterized (Fu et al., 1998), it is not known whether there are developmental functions shared by *ovo* and *sparkling*.

Tissue-restricted activity of EGL-38

An interesting feature of the genetics of *egl-38* is that mutations that preferentially affect a subset of *egl-38* functions correspond to mutations in the DNA-binding domain. This contrasts with the tissue-preferential alleles of the *Drosophila* Pax gene *sparkling*, which affect non-coding regulatory parts of the gene (Fu et al., 1998). We have shown the tissue-preferential activity of each allele also correlates with ability to promote *lin-48* gene expression in hindgut cells. For example, the *sy294* allele preferentially disrupts development of hindgut cells and mutants fail to express *lin-48*. In contrast, the *n578* allele preferentially disrupts development of the egg-laying system, and disrupts hindgut development to a minimal extent. Correspondingly, *egl-38(n578)* mutants can express *lin-48* even when the *lin-48* promoter is compromised by mutations. These results suggest the tissue-preferential alleles affect the ability of EGL-38 to regulate certain target genes and not others. The mutations may affect the ability of EGL-38 protein to bind particular DNA targets (Czerny et al., 1993; Czerny and Busslinger, 1995), or to interact with protein partners (Fitzsimmons et al., 1996).

Our characterization of *lin-48* indicates that EGL-38 has tissue-restricted targets that are expressed in only a subset of EGL-38-expressing cells. We have identified two promoter elements important for *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 *lin-48* expression pattern observed in *egl-38* mutants, and single lre1 or lre2 mutant transgenes are equally sensitive to the *egl-38(sy287)* and *egl-38(n578)* mutant backgrounds. As EGL-38 does not specifically bind lre1 in vitro, it is possible that it acts indirectly through lre1, or that in vivo EGL-38 can bind lre1, but it requires another protein or proteins to bind with high affinity. Alternatively, as *lin-48::gfp* is in multiple copies and overexpressed from the transgenes, it is possible that lre2 alone mediates the in vivo response, but lre1 is capable of functioning when multiple copies of the gene are present. Further experiments will be required to distinguish among these possibilities.

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*. Our analysis of the *lin-48* promoter, however, identified only elements that mediate the response to EGL-38. 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 between lre2 and the downstream *Hin*dIII site, as we have systematically analyzed only the region containing lre2 and upstream. It could act through a DNA element immediately adjacent to lre1 or lre2, which would have been deleted at the same time as deleting these EGL-38-sensitive sites. 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. Future work to identify additional genes important for *lin-48* expression should clarify how the EGL-38 Pax protein mediates tissue-restricted gene expression.

REFERENCES

Andrews, J., Garcia-Estefania, D., Delon, I., Lu, J., Mevel-Ninio, M., Spierer, A., Payre, F., Pauli, D. and Oliver, B. (2000). OVO transcription factors function antagonistically in the Drosophila female germline. *Development* **127**, 881-892.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (2000). Current protocols in molecular biology. New York, New York: John Wiley & Sons.

Carrol, T. J. and Vize, P. D. (1999). Synergism between *Pax-8* and *lim-1* in embryonic kidney development. *Dev. Biol.* **214**, 46-59.

Chamberlin, H. M., Palmer, R. E., Newman, A. P., Sternberg, P. W., Baillie, D. L. and Thomas, J. H. (1997). The *PAX* gene *egl-38* mediates developmental patterning in *Caenorhabditis elegans*. *Development* **124**, 3919-3928.

Chamberlin, H. M., Brown, K. B., Sternberg, P. W. and Thomas, J. H. (1999). Characterization of seven genes affecting *Caenorhabditis elegans* hindgut development. *Genetics* **153**, 731-742.

Chisholm, A. (1991). Control of cell fate in the tail region of C. elegans by the gene egl-5. *Development* **111**, 921-932.

Czerny, T., Schaffner, G. and Busslinger, M. (1993). DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* **7**, 2048-2061.

Czerny, T. and Busslinger, M. (1995). DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Mol. Cell. Biol.* **15**, 2858-2871.

Czerny, T., Bouchard, M., Kozmik, Z. and Busslinger, M. (1997). The characterization of novel Pax genes of the sea urchin and Drosophila reveal an ancient evolutionary origin of the Pax2/5/8 subfamily. *Mech. Dev.* **67**, 179-192.

Dai, X., Schonbaum, C., Degenstein, L., Bai, W., Mahowald, A. and Fuchs, E. (1998). The *ovo* gene required for cuticle formation and oogenesis in flies is involved in hair formation and spermatogenesis in mice. *Genes Dev.* **12**, 3452-3463.

Eberhard, D., Jimenez, G., Heavey, B. and Busslinger, M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J.* **19**, 2292-2303.

Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S.-M., Wasylyk, B. and Hagman, J. (1996). Pax-5(BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* **10**, 2198-2211.

Fu, W., Duan, H., Frei, E. and Noll, M. (1998). *shaven* and *sparkling* are mutations in separate enhancers of the Drosophila Pax2 homolog. *Development* **125**, 2943-2950.

Hodgkin, J. (1997). Genetics. In *C. elegans II* (ed. D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess), pp. 881-1047. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Jiang, L. I. and Sternberg, P. W. (1999). Socket cells mediate spicule morphogenesis in *Caenorhabditis elegans* males. *Dev. Biol.* **211**, 88-99.

Lee, S. and Garfinkel, M. D. (2000). Characterization of Drosophila OVO protein DNA binding specificity using random DNA oligomer selection suggests zinc finger degeneration. *Nucleic Acids Res.* **28**, 826-834.

Lu, J., Andrews, J., Pauli, D. and Oliver, B. (1998). Drosophila OVO zinc-finger protein regulates *ovo* and *ovarian tumor* target promoters. *Dev. Genes Evol.* **208**, 213-222.

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

Mansouri, A., Chowdhury, K. and Gruss, P. (1998). Follicular cells of the thyroid gland require *Pax8* gene function. *Nat. Genet.* **19**, 87-90.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959-3970.

Mevel-Ninio, M., Terracol, R. and Kafatos, F. C. (1991). The *ovo* gene of Drosophila encodes a zinc finger protein required for female germ line development. *EMBO J.* **10**, 2259-2266.

Nelson, F. K. and Riddle, D. L. (1984). Functional study of the *Caenorhabditis elegans* secretory-excretory system. *J. Exp. Zool.* **231**, 45-56.

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A. and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* **135**, 385-404.

Oliver, B., Perrimon, N. and Mahowald, A. P. (1987). The *ovo* locus is required for sex-specific germ line maintenance in Drosophila. *Genes Dev.* **1**, 913-923.

Payre, F., Vincent, A. and Carreno, S. (1999). *ovo/svb* integrates Wingless and DER pathways to control epidermis differentiation. *Nature* **400**, 271-275.

Schwartz, M., Cecconi, F., Bernier, G., Andrejewski, N., Kammandel, B., Wagner, M. and Gruss, P. (1997). Conserved biological function between *Pax-2* and *Pax-5* in midbrain and cerebellum development: evidence from targeted mutations. *Proc. Natl. Acad. Sci. USA* **94**, 14518-14523.

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

Sulston, J. E., Albertson, D. G. and Thomson, J. N. (1980). The *C. elegans* male: postembryonic development of nongonadal structures. *Dev. Biol.* **78**, 542-576.

Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64-119.

Sulston, J. E. and Hodgkin, J. (1988). Methods. In *The nematode* Caenorhabditis elegans (ed. W. Wood), pp. 587-606. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Torres, M., Gomez-Pardo, E., Dressler, G.R. and Gruss, P. (1995). Pax-2 controls multiple steps of urogenital development. *Development* **121**, 4057-4065.

Torres, M., Gomez-Pardo, E. and Gruss, P. (1996). Pax2 contributes to inner ear patterning and optic nerve trajectory. *Development* **122**, 3381-3391.

Urbánek, P., Wang, Z.-Q., Fetka, I., Wagner, E. F. and Busslinger, M. (1994). Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* **79**, 901-912.

Wheat, W., Fitzsimmons, D., Lennox, H., Krautkramer, S. R., Gentile, L. N., McIntosh, L. P. and Hagman, J. (1999). The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes. *Mol. Cell. Biol.* **19**, 2231-2241.

Wollard, A. and Hodgkin, J. (2000). The *Caenorhabditis elegans* fate-determining gene *mab-9* encodes a T-box protein required to pattern the posterior hindgut. *Genes Dev.* **14**, 596-603.

Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L. and Pabo, C. O. (1999). Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev.* **13**, 1263-1275.