DOI: 10.1534/genetics.104.029017

Identification of Residues of the *Caenorhabditis elegans* LIN-1 ETS Domain That Are Necessary for DNA Binding and Regulation of Vulval Cell Fates

Ginger R. Miley,* Douglas Fantz,*,[1] Danielle Glossip,* Xiaowei Lu,[†,2] R. Mako Saito,[‡] Robert E. Palmer,[§,3] Takao Inoue,[§] Sander van den Heuvel,[‡] Paul W. Sternberg[§] and Kerry Kornfeld*,[4]

**Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, Missouri 63110,* [†]*Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139,* [‡]*Massachusetts General Hospital Cancer Center, Charlestown, Massachusetts 02129 and* [§]*Howard Hughes Medical Institute, Division of Biology, California Institute of Technology, Pasadena, California 91125*

Manuscript received April 7, 2004
Accepted for publication May 6, 2004

ABSTRACT

LIN-1 is an ETS domain protein. A receptor tyrosine kinase/Ras/mitogen-activated protein kinase signaling pathway regulates LIN-1 in the P6.p cell to induce the primary vulval cell fate during *Caenorhabditis elegans* development. We identified 23 *lin-1* loss-of-function mutations by conducting several genetic screens. We characterized the molecular lesions in these *lin-1* alleles and in several previously identified *lin-1* alleles. Nine missense mutations and 10 nonsense mutations were identified. All of these *lin-1* missense mutations affect highly conserved residues in the ETS domain. These missense mutations can be arranged in an allelic series; the strongest mutations eliminate most or all *lin-1* functions, and the weakest mutation partially reduces *lin-1* function. An electrophoretic mobility shift assay was used to demonstrate that purified LIN-1 protein has sequence-specific DNA-binding activity that required the core sequence GGAA. LIN-1 mutant proteins containing the missense substitutions had dramatically reduced DNA binding. These experiments identify eight highly conserved residues of the ETS domain that are necessary for DNA binding. The identification of multiple mutations that reduce the function of *lin-1* as an inhibitor of the primary vulval cell fate and also reduce DNA binding suggest that DNA binding is essential for LIN-1 function in an animal.

INTRACELLULAR signaling specifies many cell fates during development. The *Caenorhabditis elegans* vulva is a useful model system for understanding how signal transduction cascades regulate cell fates. The vulva is a specialized epidermal structure used for egg laying and sperm entry that is formed by the descendants of three ectodermal blast cells, P5.p, P6.p, and P7.p (Horvitz and Sternberg 1991). In wild-type hermaphrodites, the anchor cell of the somatic gonad signals to P6.p using the LIN-3 epidermal growth factor-like ligand (reviewed by Greenwald 1997; Kornfeld 1997; Sternberg and Han 1998). LIN-3 presumably binds to the LET-23 receptor tyrosine kinase (RTK). This is likely to stimulate receptor autophosphorylation and create docking sites for the SEM-5 adaptor protein and the LET-341 Ras guanine nucleotide exchange factor. LET-341 is likely to cause LET-60 Ras to release GDP, resulting in GTP binding and LET-60 activation. Activated LET-60 Ras can bind and activate the serine/threonine kinase LIN-45 Raf. Activated LIN-45 phosphorylates and thereby activates the MEK-2 mitogen-activated protein (MAP) kinase kinase. MEK-2 phosphorylates and thereby activates the MPK-1 extracellular signal-regulated kinase (ERK) MAP kinase. MPK-1 appears to phosphorylate multiple target proteins, including the LIN-1 ETS transcription factor, and these modifications cause P6.p to adopt the 1° vulval cell fate (eight descendants). When P6.p is induced to adopt the 1° vulval cell fate, it signals to P5.p and P7.p through the LIN-12 Notch receptor, causing these cells to adopt the 2° vulval cell fate (seven descendants). The anchor cell signal can also promote the 2° cell fate in P5.p and P7.p. While P3.p, P4.p, and P8.p are capable of adopting vulval fates, they normally receive neither of these signals and thus adopt the 3° nonvulval cell fate (two descendants). In hermaphrodites with a loss-of-function (lf) mutation in any of the core signaling genes, P5.p, P6.p, and P7.p adopt 3° nonvulval cell fates, resulting in a worm with a vulvaless (Vul) phenotype. By contrast, in hermaphrodites with a gain-of-function (gf) mutation that constitutively acti-

[1]*Present address:* Department of Biology, Wake Forest University, Winston-Salem, NC 27109.

[2]*Present address:* Stanford University, Biological Sciences Department, Stanford, CA 94305.

[3]Deceased.

[4]*Corresponding author:* Department of Molecular Biology and Pharmacology, Washington University School of Medicine, Campus Box 8103, 660 S. Euclid Ave., St. Louis, MO 63110.
E-mail: kornfeld@molecool.wustl.edu

vates a core signaling gene such as *let-60 ras*, P3.p, P4.p, and P8.p inappropriately adopt vulval cell fates; the resulting ectopic tissue forms a series of ventral protrusions, the multivulva (Muv) phenotype.

Genetic studies indicate that the *lin-1* gene is a critical target of the RTK/Ras/ERK signaling pathway. *lin-1(lf)* mutations cause a strong Muv phenotype, indicating that *lin-1* activity inhibits the 1° vulval cell fate and/or promotes the 3° nonvulval cell fate (Sulston and Horvitz 1981; Ferguson and Horvitz 1985). The Muv phenotype caused by a *lin-1(lf)* mutation is epistatic to the Vul phenotype caused by a loss-of-function mutation in *mpk-1* or other core signaling genes, indicating that *lin-1* functions downstream of *mpk-1* if these genes act in a linear signaling pathway (Ferguson *et al.* 1987; Lackner *et al.* 1994; Wu and Han 1994). *lin-1(gf)* mutations cause a Vul phenotype, reinforcing the conclusion that *lin-1* activity inhibits the 1° vulval cell fate. *lin-1(gf)* alleles appear to be constitutively active and unresponsive to negative regulation by the RTK/Ras/ERK signaling pathway (Jacobs *et al.* 1998). *lin-1* also appears to be a critical target of the RTK/Ras/ERK signaling pathway during the development of the excretory duct cell in the first larval stage (Yochem *et al.* 1997). Loss-of-function mutations in core signaling genes such as *mek-2* cause a larval lethal phenotype characterized by a rigid, rod-like morphology because the excretory duct cell fate is not specified correctly. This lethal phenotype is suppressed by *lin-1(lf)* mutations, indicating that *lin-1* is negatively regulated by RTK/Ras/ERK signaling during excretory duct cell development (Kornfeld *et al.* 1995). Consistent with this model, *lin-1(gf)* mutations cause a larval lethal phenotype characterized by a rigid rod-like morphology (Jacobs *et al.* 1998).

The *lin-1* gene encodes a 441-amino-acid protein that contains a highly conserved ETS domain (Beitel *et al.* 1995). Many ETS domains display DNA-binding activity, and the minimal recognition sequence is typically GGAA/T (Karim *et al.* 1990; Shore *et al.* 1996; Remy and Baltzinger 2000). The structures of the ETS domains of PU.1, SAP-1, and Elk-1 bound to DNA were determined using X-ray crystallography (Kodandapani *et al.* 1996; Mo *et al.* 1998, 2000). These ETS domains have a winged helix-loop-helix structure; the α3 helix lies in the major groove and contacts the GGA motif and the flanking amino acids contact the phosphate backbone of the DNA. Several ETS proteins have been demonstrated to regulate transcription positively or negatively, and these proteins contain additional domains that are necessary for activation or repression (Laudet *et al.* 1999; Mavrothalassitis and Ghysdael 2000).

Beitel *et al.* (1995) characterized the molecular lesions of 16 *lin-1(lf)* alleles. Two alleles have missense changes that affect the ETS domain, 6 alleles have nonsense changes, and 8 alleles have gross rearrangements. These studies indicate that the ETS domain is necessary for LIN-1 function. However, LIN-1 protein has not been demonstrated to bind DNA or regulate transcription. LIN-1 was reported to interact with LIN-31, a putative transcription factor of the HNF-3/forkhead family (Tan *et al.* 1998). Jacobs *et al.* (1998) characterized the molecular lesions of 6 *lin-1(gf)* alleles. Three alleles have missense changes, 2 alleles have nonsense changes, and one mutation affects mRNA splicing. These six mutations affect the C terminus of LIN-1 to alter or eliminate a conserved FQFP motif. The C terminus of LIN-1 is phosphorylated by ERK MAP kinase, and the FQFP motif is a docking site that mediates the high-affinity interactions with ERK (Jacobs *et al.* 1998, 1999; Tan *et al.* 1998). LIN-1 also contains a D-domain docking site that mediates high-affinity interactions with ERK (Yang *et al.* 1998; Jacobs *et al.* 1999). The FQFP motif and the D domain function independently and additively to mediate high-affinity interactions with ERK and direct phosphorylation of specific S/TP residues (Fantz *et al.* 2001). These studies indicate that phosphorylation of LIN-1 by ERK reduces or eliminates the ability of LIN-1 to inhibit the 1° vulval cell fate.

To investigate the role of specific residues and domains of LIN-1, we identified 23 loss-of-function *lin-1* mutations by conducting a variety of genetic screens. We identified the molecular lesions in these alleles and in previously described *lin-1* alleles that had not been characterized molecularly. Ten nonsense and 9 missense mutations that affect conserved amino acids in the ETS domain were identified. To elucidate how these missense mutations affect the function of the ETS domain, we established an assay for DNA binding. The LIN-1 ETS domain displayed sequence-specific DNA-binding activity, and each of these missense substitutions reduced or eliminated LIN-1 DNA binding. These findings contribute to the understanding of ETS domain function by identifying residues that are necessary for DNA binding. The combination of genetic and biochemical studies demonstrates that DNA binding is necessary for LIN-1 function in a worm.

MATERIALS AND METHODS

General methods and strains: *C. elegans* strains were cultured as described by Brenner (1974) and grown at 20° unless otherwise noted. The wild-type strain and parent of all mutant strains was *C. elegans* variety Bristol strain N2. Mutations used in this study are described by Riddle *et al.* (1997) and are as follows: linkage group (LG)III—*lin-12(n137n460ts)*, *lin-12(n950)*; LGIV—*unc-5(e53)*, *bli-6(sc16)*, *dpy-20(e1282)*, *lin-1(e1026*, *e1275*, *n176)* (Horvitz and Sulston 1980), *lin-1(e1777*, *n303*, *n304*, *n383*, *n431*, *n746*, *n753*, *n757*, *n1047*, *n1054*, *n1140)* (Ferguson and Horvitz 1985), *lin-1(ar147*, *m546*, *n381)* (Beitel *et al.* 1995), *lin-1(n2374)* (Thomas *et al.* 2003), *lin-1(ga56*, *ga68)* (Eisenmann and Kim 2000), *lin-1(he117*, *he119*, *n1814*, *n1815*, *n1816*, *n1817*, *n1848*, *n2692*, *n2693*, *n2694*, *n2696*, *n2701*, *n2704*, *n2705*, *n2714*, *n2750*, *n3000*, *n3443*, *n3522*, *sy254*, *sy289*, *sy321*, *sy618)* (this study); LGV—*let-341(n1613ts)*, *him-5(e1490)*; LGX—*lin-15(n767)*, *lon-2(e678)*. We used standard genetic techniques to separate the *lin-1(lf)* mutations from *lin-12(n137n460ts)*, *lin-*

12(n950), *let-341(n1613ts)*, or *lin-15(n767)*, to backcross these mutations to N2, to position these mutations on the genetic map, and to perform complementation tests (Brenner 1974).

Identification of *lin-1(lf)* mutations: In general, genetic screens were conducted as described by Brenner (1974) using ethyl methanesulfonate (EMS) as a mutagen. To identify mutations that suppress the rod-like larval lethality caused by *let-341(n1613ts)*, we mutagenized *let-341(n1613ts)* hermaphrodites with EMS, raised F_1 hermaphrodites at the permissive temperature of 20° until the L4 larval stage, and then transferred the animals to the nonpermissive temperature of 25°. The F_2 self-progeny were screened for survivors. We screened 40,000 haploid genomes and identified five mutations that caused a Muv phenotype and failed to complement the Muv phenotype of *lin-1(e1275)*, indicating that these are alleles of *lin-1: n1814*, *n1815*, *n1816*, *n1817*, and *n1848*. We also isolated one *let-60(n1849gf)* Muv mutation, five linked mutations that are likely to be intragenic revertants, and several unlinked suppressors that did not cause a Muv phenotype. In a related screen, we mutagenized *let-341(n1613ts)* hermaphrodites with EMS and screened for mutations that suppressed both the larval lethal and vulvaless phenotypes at 25°. An F_1 screen of 1.3×10^6 haploid genomes identified 21 Muv mutations and 60 non-Muv mutations. An F_2 screen of 1.2×10^5 haploid genomes identified 33 Muv mutations and 50 non-Muv mutations. These Muv mutations included the following 10 *lin-1* alleles: *n2692*, *n2693*, *n2694*, *n2696*, *n2701*, *n2704*, *n2705*, *n2714*, *n2750*, and *n3000*.

To identify class B synthetic Muv mutations, we mutagenized a strain containing the class A mutation *lin-15(n767)* with EMS and conducted a clonal screen for F_2 progeny that displayed a Muv phenotype. We screened 6760 haploid genomes and identified two mutations, *n3443* and *n3522*, that caused independent Muv phenotypes. *n3443* and *n3522* displayed linkage to *dpy-20(IV)* and failed to complement the Muv phenotype of *lin-1(e1275)*.

To identify enhancers of the *lin-12(gf)* Muv phenotype, we conducted two screens. First, we mutagenized *lin-12(n137n460)* hermaphrodites using trimethylpsoralen (TMP) with ultraviolet (UV) light (Yandell *et al.* 1994) and screened F_2 self-progeny for extra vulval precursor cell (VPC) divisions. We screened 10,480 haploid genomes and identified the Muv mutation *he117*. Second, we mutagenized *lin-12(n950); him-5(e1490)* hermaphrodites with EMS and screened F_2 self-progeny for extra VPC divisions. We screened 2450 haploid genomes and identified the Muv mutation *he119*. *he117* and *he119* animals were backcrossed four times to N2. Each mutation displayed linkage to *unc-5(IV)*. *he117* and *he119* failed to complement the Muv phenotype of *lin-1(e1777)* and *lin-1(n304)*.

To identify mutations that cause a protruding vulva, we used EMS to mutagenize wild-type hermaphrodites and screened F_2 self-progeny for animals with a protruding vulva. We screened ~7000 haploid genomes and identified 11 mutations, one of which, *sy618*, was Muv and mapped to chromosome IV. The *sy618* strain was backcrossed twice to N2, and *sy618* failed to complement the Muv phenotype of *lin-1(e1777)*. To identify mutations that cause vulval defects that result in an extruded gonad at the L4 molt, we used EMS to mutagenize wild-type hermaphrodites and conducted an F_1 clonal screen. We identified the mutation *sy289* and backcrossed *sy289* twice to N2. To identify mutations that cause a Muv phenotype, we mutagenized wild-type hermaphrodites with EMS and isolated the Muv mutation *sy321*. *sy254* was isolated as an F_2 Muv after TMP and UV mutagenesis of N2.

To confirm that the eight alleles that contained the wild-type sequence in the *lin-1* coding region are indeed alleles of *lin-1*, we performed an additional complementation test. Males containing the mutant alleles *ga68*, *he119*, *n1054*, *n1814*, *n1817*, *n2693*, *n2705*, and *n3443* were mated to *bli-6(sc16) lin-1(e1275)* hermaphrodites. Non-Bli, Muv hermaphrodites were observed in each case, indicating that these eight mutations failed to complement the *lin-1(e1275)* Muv phenotype.

Determination of DNA sequences of *lin-1* alleles: We determined the DNA sequence of 28 *lin-1* alleles (Table 1). Genomic DNA was derived from homozygous mutant adult hermaphrodites and amplified by polymerase chain reaction (PCR) according to Williams *et al.* (1992). *lin-1* contains six exons (Beitel *et al.* 1995). We PCR amplified fragments containing each exon and adjacent introns as shown in Figure 1A and as described by Jacobs *et al.* (1998) with the following modifications: GBO6f (5′-CCTTTCGAAATCGCTTCAAATC) and GBO14r were used to amplify a 514-bp fragment containing 109 bp of intron 2, exon 3, and 69 bp of intron 3. GBO15f (5′-GTGATAACAAACATTTTGTCAGTTG) and GBO17r were used to amplify a 538-bp fragment containing 121 bp of intron 3, exon 4, and 137 bp of intron 4. We purified the PCR-amplified DNA fragments and determined the complete sequences of both strands using an automated ABI Prism 377 DNA sequencer or an ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, CA).

Phenotypic analysis: The 10 *lin-1* alleles described in Table 2 were backcrossed twice to N2. Phenotypes were analyzed by placing one egg per petri dish for strains that did not display a highly penetrant egg-laying (Egl) defect or one L1 larva per petri dish for strains that displayed a highly penetrant Egl defect. Animals were allowed to develop at 20° and inspected daily using a dissecting microscope to identify: (1) animals that died during larval development or at the transition from L4 to young adult, (2) the number of pseudovulval protrusions on adults, and (3) the ability to generate live progeny during the first 2 days of adulthood. To analyze temperature sensitivity, we placed 15 young adult hermaphrodites on petri dishes, allowed the F_1 progeny to develop at 15°, 20°, and 25°, and scored the percentage of the F_1 population that displayed the Muv phenotype. To further analyze *lin-1(ga56)*, we placed one egg per petri dish at 15°, 20°, and 25° and scored the Muv phenotype of adults.

Expression and purification of LIN-1 proteins: A fragment of a *lin-1* cDNA encoding LIN-1 residues 1–278 (Beitel *et al.* 1995) was ligated to pGEX-2T or pGEX-KG (Pharmacia LKB Biotechnologies, Piscataway, NJ) to generate plasmids encoding the fusion protein glutathione-*S*-transferase (GST):LIN-1(1–278). Standard *in vitro* mutagenesis techniques (Sambrook *et al.* 1989) were used to modify the *lin-1* coding region to create the following plasmids: pDG91 encodes GST:LIN-1 (1–278 M114K), pDG92 encodes GST:LIN-1(1–278 Y127F), pDG93 encodes GST:LIN-1(1–278 Y126F), pDG94 encodes GST:LIN-1(1–278 R124Q), pDG95 encodes GST:LIN-1(1–278 R121K), pDG96 encodes GST:LIN-1(1–278 E92K), pDG97 encodes GST:LIN-1(1–278 G106R), and pDG112 encodes GST:LIN-1(1–278 L70F). We verified the entire LIN-1 coding region of each plasmid by DNA sequencing. The GST:LIN-1(1–278 M114K), GST:LIN-1(1–278 L70F), and GST:LIN-1(1–278 E92K) proteins did not express efficiently in bacteria (data not shown). To express these proteins in S*f*9 insect cells, we ligated the coding regions for GST and LIN-1 into the vector pFastBac1 (Invitrogen, San Diego) to create the following plasmids: pDG104 encodes GST:LIN-1(1–278 WT), pDG105 encodes GST:LIN-1(1–278 M114K), pDG106 encodes GST:LIN-1(1–278 E92K), and pDG122 encodes GST:LIN-1(1–278 L70F).

For bacterial expression we transformed plasmids into *Escherichia coli* strain BL21, induced expression with 0.05 mM isopropyl thiogalactoside, disrupted the cells with sonication in 1× PBS buffer containing 1× complete protease inhibitor cocktail (PIC by Roche) and 0.2 M dithiothreitol (DTT), and prepared

protein extracts. GST:LIN-1 fusion proteins were either soluble in 1× phosphate buffered saline (PBS) or solubilized in 1× PBS containing 7 M urea and then dialyzed overnight against PBS. S*f*9 insect cells were prepared according to Ausubel (1987) and the manufacturer's instructions (Invitrogen). We transformed plasmids into *E. coli* DH10Bac cells (Invitrogen), isolated the recombinant bacmid DNA, transfected S*f*9 cells, harvested the virus, and infected S*f*9 cells. Infected S*f*9 cells were scraped from the dish, washed with 1× PBS, pelleted, frozen, and treated with insect cell lysis buffer [50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1 mM DTT, 10 mM PMSF, 1× complete PIC (Roche)]. To purify GST:LIN-1 fusion proteins, we incubated protein extracts with glutathione-sepharose 4B, extensively washed, eluted bound proteins using 10 mM glutathione (in 50 mM Tris-HCl, pH 8.0) or 100 mM triethylamine, dialyzed the eluted proteins against kinase assay buffer (36 mM Tris, 0.1 mM EGTA, pH 7.08), and then dialyzed overnight against 1× gel shift buffer (200 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 10% glycerol). Protein samples were stored at 4° and used within 24 hr or stored at −80° after addition of 50% glycerol (Tanaka *et al.* 2002). To determine the amount of intact fusion protein in a partially purified sample, we separated the protein sample by SDS-PAGE with a 4% stacking/10% resolution gel, stained with Coomassie blue, and estimated the amount of fusion protein by comparison to the staining of known amounts of bovine serum albumin in adjacent lanes.

Electrophoretic mobility shift assay: Double-stranded DNA was created by annealing two single-stranded oligonucleotides. Wild-type oligonucleotides were 5′-CTAGAGCTGAATAACC GGAAGTAACTAT and 5′-CATGATAGTTACTTCCGGTTAT TCAGCT. Mutant oligonucleotides were 5′-CTAGAGCTGAA TAAGCTACTGTAACTAT and 5′-CATGATAGTTACAGTAG CTTATTCAGCT. DNA was radioactively labeled by using DNA polymerase I Klenow (New England Biolabs, Beverly, MA) to fill the 3′ single-stranded overhang with [α-^{32}P]dATP (Ausubel 1987). Radiolabeled DNA was purified by phenol:chloroform extraction and a Sephadex G-50 column (Ausubel 1987; Sambrook *et al.* 1989). We incubated ~30,000 cpm of radiolabeled DNA, various amounts of protein, and 1× reaction buffer (5 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 5% glycerol) in a total volume of 12.5 μl for 20 min on ice or at room temperature. Samples were fractionated using a 6% polyacrylamide gel (30% acrylamide/0.8% bis-acrylamide) with 0.5× TBE (Ausubel 1987; Sambrook *et al.* 1989). Gels were dried and exposed to film or a Kodak PhosphorScreen (Molecular Dynamics, Sunnyvale, CA). The PhosphorScreen was scanned using a Molecular Imager FX (Bio-Rad, Richmond, CA) and analyzed with Quantity One 4.1.0 (Bio-Rad) software.

RESULTS

Identification of 23 *lin-1* alleles: To identify mutations in genes that regulate vulval development, we conducted several different genetic screens. Here we focus on 23 *lin-1* alleles that were identified in these screens. The *let-341* gene encodes a guanine nucleotide exchange factor that is necessary for Ras-mediated signaling; *let-341(lf)* mutations cause rod-like, larval lethality and a vulvaless defect (Johnsen and Baillie 1991; Clark *et al.* 1992; Chang *et al.* 2000). *lin-1(lf)* mutations suppress the *let-341(lf)* lethal phenotype, and the *lin-1(lf)* Muv phenotype is epistatic to the *let-341(lf)* Vul phenotype, indicating that *let-341* functions upstream of *lin-1* (Clark *et al.* 1992). We identified 15 alleles of *lin-1* as suppressors of the *let-341(n1613ts)* lethal phenotype, including *n1814*, *n1815*, *n1816*, *n1817*, *n1848*, *n2692*, *n2693*, *n2694*, *n2696*, *n2701*, *n2704*, *n2705*, *n2714*, *n2750*, and *n3000*. *lin-15A* is a synthetic multivulva gene, and *lin-15A(n767)* hermaphrodites display normal vulval development (Ferguson and Horvitz 1989; Sternberg and Horvitz 1989). We identified the *lin-1* alleles *n3443* and *n3522* in a screen for multivulva hermaphrodites after mutagenizing *lin-15A(n767)* hermaphrodites. *lin-12* encodes a receptor similar to Notch (Yochem *et al.* 1988). *lin-12* is activated in P5.p and P7.p to promote the 2° cell fate, and *lin-12(gf)* mutations cause a Muv phenotype characterized by ectopic 2° cell fates. We identified the *lin-1* alleles *he117* and *he119* as enhancers of the *lin-12(gf)* Muv phenotype. We identified four *lin-1* alleles by mutagenizing wild-type animals and screening for vulval defects. The allele *sy618* was identified by screening for animals with a protruding vulva, the allele *sy289* was identified by screening for animals that died at the L4 molt because of an extruded gonad, and the alleles *sy254* and *sy321* were identified by screening for animals with a multivulval phenotype.

Each of these 23 newly identified mutations caused a multivulval phenotype that was similar in appearance to the multivulval phenotype caused by *lin-1(lf)* mutations. Furthermore, these mutations were positioned on chromosome IV like the *lin-1* gene (see MATERIALS AND METHODS). To determine if these were alleles of *lin-1*, we performed complementation tests. Each mutation failed to complement the Muv phenotype caused by a *lin-1(lf)* allele (see MATERIALS AND METHODS). These results suggest that these are alleles of *lin-1*.

Previous studies have identified 20 loss-of-function *lin-1* mutations and 7 gain-of-function *lin-1* mutations. These *lin-1(lf)* mutations include *e1026, e1275, n176* (Horvitz and Sulston 1980), *e1777, n303, n304, n383, n431, n746, n753, n757, n1047, n1054, n1140* (Ferguson and Horvitz 1985), *ar147, m546, n381* (Beitel *et al.* 1995), *ga56*, *ga68* (Eisenmann and Kim 2000), and *n2374* (Thomas *et al.* 2003). Therefore, the 23 mutations described above bring the total number of *lin-1(lf)* mutations to 43. The 7 *lin-1(gf)* mutations include *ky54*, *n1761*, *n1790*, *n1855*, *n2515*, *n2525* (Jacobs *et al.* 1998), and *cs50* (Rocheleau *et al.* 2002).

Molecular characterization of *lin-1* alleles: The 7 *lin-1(gf)* mutations have been characterized molecularly (Jacobs *et al.* 1998; Rocheleau *et al.* 2002), and 16 *lin-1(lf)* mutations have been characterized molecularly by Beitel *et al.* (1995). To identify residues and domains that are necessary for the function of the *lin-1* gene, we investigated the molecular lesions of the 27 *lin-1(lf)* mutations that had not been characterized molecularly. These include 20 of the newly identified mutations and 7 previously identified mutations. The *lin-1* gene contains six exons (Beitel *et al.* 1995). We determined

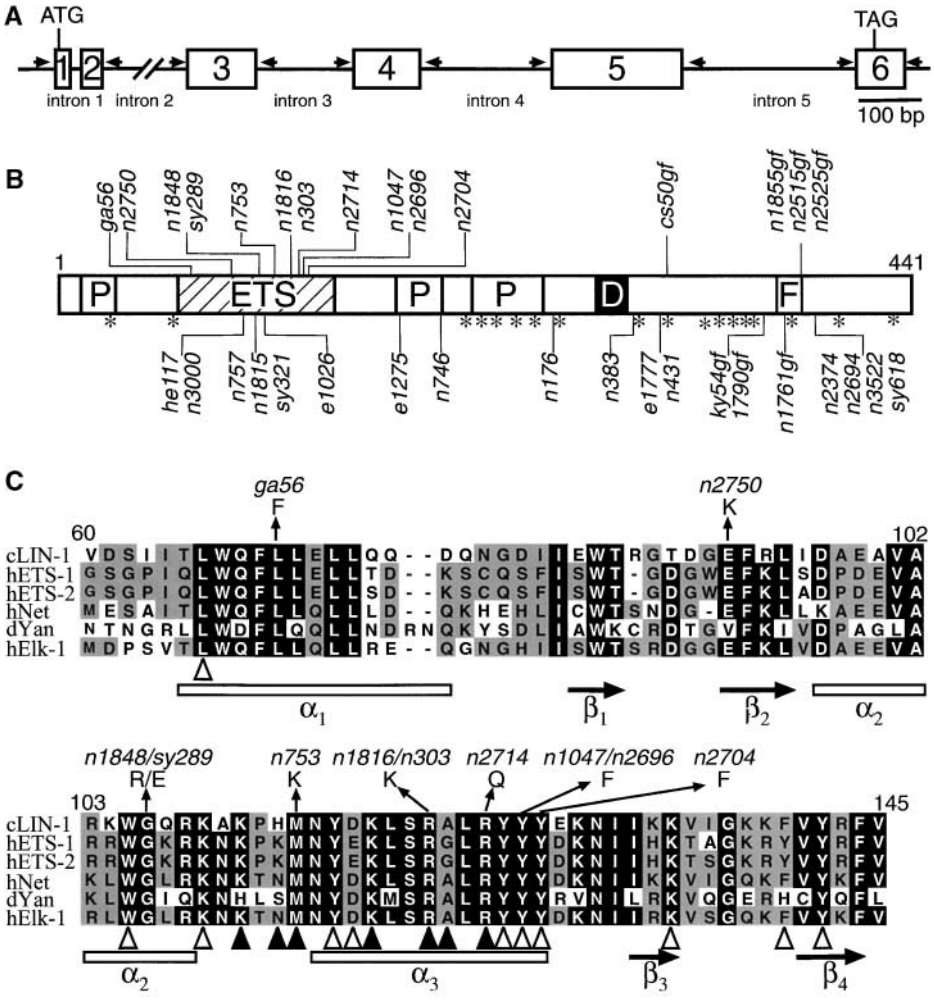

FIGURE 1.—LIN-1 gene and protein. (A) The *lin-1* locus is diagrammed with exons (numbered boxes), introns (lines), and the start (ATG) and stop (TAG) codons of the open reading frame (BEITEL *et al.* 1995). The beginning of exon 1 is defined by the position of an SL1 *trans*-spliced leader in *lin-1* mRNA, and the end of exon 6 is defined by the position of a poly(A) tail in *lin-1* mRNA. Arrows show the positions of pairs of oligonucleotides that were used to amplify exons and splice sites for DNA sequencing (see MATERIALS AND METHODS; JACOBS *et al.* 1998). (B) Diagram of the LIN-1 protein with predicted PEST domains (P; amino acids 11–27, 176–199, and 216–252); the ETS domain (amino acids 60–145); the D domain (D; amino acids 277–291), and the FXFP motif (F; amino acids 382–385) that are ERK MAP kinase docking sites; and serine/threonine proline motifs that are potential ERK phosphorylation sites (asterisks). Lines indicate the positions of *lin-1* missense mutations (above) and nonsense and splice site mutations (below). (C) ClustalW alignment of the ETS domains of *C. elegans* LIN-1 (amino acids 60–145; BEITEL *et al.* 1995), human ETS-1 (amino acids 331–415; WATSON *et al.* 1988), human ETS-2 (amino acids 359–443; WATSON *et al.* 1988), human Net (amino acids 1–85; GIOVANE *et al.* 1994), Drosophila Yan (amino acids 392–479; LAI and RUBIN 1992), and human Elk-1 (amino acids 1–86; RAO *et al.* 1989). Identities are shown between two to four proteins (shaded background) and five to six proteins (solid background). Substitutions caused by *lin-1(lf)* missense mutations are shown above. Solid and open triangles indicate residues of human Elk-1 that interact with nucleotides in the core binding motif (GGAA) and nucleotides outside the core binding motif, respectively, on the basis of the crystal structure of Elk-1 bound to DNA (MO *et al.* 2000). α-helices (boxes) and β-sheets (arrows) of human Elk-1 are indicated.

the DNA sequence of these six exons and each splice junction, a region that includes the entire *lin-1* coding region, using DNA from each of these 27 mutants (Figure 1A). A single base pair change compared to the wild-type sequence was detected in 19 alleles. Table 1 shows these nucleotide substitutions and the predicted amino acid substitutions. Seven newly identified alleles and two previously identified but molecularly uncharacterized alleles have missense changes: *n1816*, *n1848*, *n2696*, *n2704*, *n2714*, *n2750*, *sy289*, *ga56*, and *n753* (Table 1). BEITEL *et al.* (1995) previously identified missense changes in *n303* and *n1047*. Thus, a total of 11 *lin-1(lf)* alleles have missense changes (Figure 1, B and C). The alleles *n303* and *n1816* have the identical base pair change and result in the substitution R121K; only the allele *n1816* was further characterized. The alleles *n1047* and *n2696* have the identical base pair change and result in the substitution Y126F; only the allele *n1047* was further characterized. The mutations *n1848* and *sy289* affect adjacent base pairs and change the same codon resulting in the substitutions G106R and G106E, respectively. Thus, these 11 independently derived mutations affect eight different amino acids.

Seven newly identified mutations and three previously identified but molecularly uncharacterized mutations are nonsense changes: *he117*, *n3000*, *n1815*, *sy321*, *e1026*, *n746*, *n2374*, *n2694*, *n3522*, and *sy618* (Table 1). BEITEL *et al.* (1995) showed that six mutations are nonsense changes: *n757*, *e1275*, *n176*, *n383*, *n431*, and *e1777* (Table 1). Therefore, a total of 16 *lin-1(lf)* alleles contain nonsense changes. These mutations cause premature stop codons at nine different amino acid residues (Figure 1B). The alleles *n3000* and *he117* contain the identical base pair change and result in the substitution E99Ochre. Codon 105 is mutated in three alleles: *n757* and *n1815* have the identical base pair change that results in an amber stop codon whereas *sy321* affects the adjacent base pair and results in an opal stop codon.

TABLE 1

Identification and molecular characterization of *lin-1* alleles

Allele[a]	Nucleotide change[b]	Codon change[b]	Allele identification[c]	Molecular characterization[c]
		Missense		
ga56	CTT → TTT	L70F	Eisenmann and Kim (2000)	This study
n2750	GAA → AAA	E92K	This study	This study
n1848	GGA → AGA	G106R	This study	This study
sy289	GGA → GAA	G106E	This study	This study
n753	ATG → AAG	M114K	Ferguson and Horvitz (1985)	This study
n303	AGA → AAA	R121K	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
n1816	AGA → AAA	R121K	This study	This study
n2714	CGA → CAA	R124Q	This study	This study
n1047	TAT → TTT	Y126F	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
n2696	TAT → TTT	Y126F	This study	This study
n2704	TAT → TTT	Y127F	This study	This study
cs50gf	CCG → CTG	P315L	Rocheleau *et al.* (2002)	Rocheleau *et al.* (2002)
n2515gf	CCG → CTG	P384L	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
n1855gf	CCG → TCG	P384S	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
n2525gf	CCG → TCG	P384S	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
		Nonsense		
he117	GAA → TAA	E99Ochre	This study	This study
n3000	GAA → TAA	E99Ochre	This study	This study
n757	TGG → TAG	W105Amber	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
n1815	TGG → TAG	W105Amber	This study	This study
sy321	TGG → TGA	W105Opal	This study	This study
e1026	CAA → TAA	Q107Ochre	Horvitz and Sulston (1980)	This study
e1275	CGA → TGA	R175Opal	Horvitz and Sulston (1980)	Beitel *et al.* (1995)
n746	CAA → TAA	Q196Ochre	Ferguson and Horvitz (1985)	This study
n176	CGA → TGA	R255Opal	Horvitz and Sulston (1980)	Beitel *et al.* (1995)
n383	CAA → TAA	Q298Ochre	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
e1777	CAG → TAG	Q309Amber	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
n431	CAG → TAG	Q309Amber	Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
ky54gf	CGA → TGA	R352Opal	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
n1790gf	CGA → TGA	R352Opal	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
n2374	CAG → TAG	Q390Amber	Thomas *et al.* (2003)	This study
n2694	CAG → TAG	Q390Amber	This study	This study
n3522	CAG → TAG	Q390Amber	This study	This study
sy618	CAG → TAG	Q390Amber	This study	This study
		Splice site		
n1761gf	CAGgt → CAGat	Splice defect after Q379	Jacobs *et al.* (1998)	Jacobs *et al.* (1998)
		Rearrangments/deletion		
ar147	Exons 4, 5, and 6 rearranged/duplicated		Beitel *et al.* (1995)	Beitel *et al.* (1995)
m546	Exons 3 and 4 rearranged		Beitel *et al.* (1995)	Beitel *et al.* (1995)
n304	Exons 3 and 4 deleted; 5 and 6 rearranged		Ferguson and Horvitz (1985)	Beitel *et al.* (1995)
n381	Exon 5 rearranged/duplicated		Beitel *et al.* (1995)	Beitel *et al.* (1995)
n1140	Intron 3 and exon 4, 61 bp deletion		Ferguson and Horvitz (1985)	Beitel *et al.* (1995); this study
n2692	Exon 5 rearranged/duplicated		This study	Beitel *et al.* (1995)
n2701	Exons 5 and 6 rearranged/duplicated		This study	Beitel *et al.* (1995)
sy254	Exons 3 and 6 deleted; exons 4 and 5 rearranged		This study	Beitel *et al.* (1995)
		Wild-type ORF		
ga68			Eisenmann and Kim (2000)	This study
he119			This study	This study
n1054			Ferguson and Horvitz (1985)	This study
n1814			This study	This study
n1817			This study	This study
n2693			This study	This study
n2705			This study	This study
n3443			This study	This study

[a] *lin-1* alleles are classified on the basis of the presence of a missense mutation, a nonsense mutation, a splice site mutation, a rearrangement and/or deletion, or a wild-type open reading frame (ORF).

[b] Wild-type (left) and mutant (right) DNA sequences are shown with the corresponding codon change in missense and nonsense alleles.

[c] The identification of the *lin-1* alleles and the analysis of the molecular lesion are described in this study or in the cited references.

TABLE 2

Phenotypes of *lin-1* missense alleles and LIN-1 mutant proteins

		Lethality (%)[a]				Pseudovulva[d]					
Allele	Lesion	L1–L3	L4/Adult	N[b]	Muv (%)[c]	Average	Range	Egl[e]	Sterility (%)[f]	N[g]	DNA binding[h]
N2	WT	1	0	186	0	0	0	No	0	184	Strong
sy254	Deletion	6	54	182	100	3.6	3–6	Yes	16	73	ND
n1047	Y126F	7	41	148	100	3.2	2–4	Yes	21	76	Undetectable
sy289	G106E	5	35	195	100	3.6	2–6	Yes	34	117	ND
n1848	G106R	7	33	112	100	3.5	3–5	Yes	28	67	Undetectable
n1816	R121K	5	27	190	100	3.9	2–5	Yes	20	129	Undetectable
n2704	Y127F	1	30	172	99	3.2	1–5	No	10	118	Very weak
n2750	E92K	6	13	168	98	3.5	0–5	No	11	137	Undetectable
n2714	R124Q	5	9	174	99	3.3	1–5	No	8	151	Undetectable
n753	M114K	5	3	187	99	3.2	1–5	Yes	2	173	Undetectable
ga56	L70F	1	1	142	86	2.5	0–4	No	2	139	Weak

[a] Percentage of all animals that died during larval stages L1, L2, or L3 (L1–L3) or during the L4 to adult transition (L4/Adult).
[b] N is the number of larvae analyzed.
[c] Percentage of adult hermaphrodites that displayed one or more protrusions that were not at the position of the normal vulva.
[d] The average number of pseudovulval protrusions, and the largest and smallest number observed in the populations (range).
[e] "Yes" means eggs were never or rarely laid on the petri dish, and "no" means eggs were frequently laid.
[f] Percentage of adult hermaphrodites that did not produce live progeny.
[g] N is the number of adult hermaphrodites analyzed.
[h] A summary of the DNA-binding affinity displayed by mutant LIN-1(1–278) proteins based on the EMSA experiments shown in Figure 4. ND, not determined.

The alleles *n431* and *e1777* contain the identical base pair change and result in the substitution Q309Amber. The alleles *n2374*, *n2694*, *n3522*, and *sy618* contain the identical base pair change and result in the substitution Q390Amber.

Eight *lin-1(lf)* alleles contain rearrangements and/or deletions of the *lin-1* locus (Beitel *et al.* 1995; Table 1). This group includes the alleles *n2692*, *n2701*, and *sy254* that were identified in the genetic screens described above. Beitel *et al.* (1995) used Southern blots to show that *lin-1(n1140)* contains a deletion that was estimated to be ~100 bp. We determined the sequence of the *n1140* allele; *n1140* contains a 61-bp deletion that includes 9 bp of intron 3 and 52 bp of exon 4 (Table 1).

Six newly identified alleles and two previously identified alleles contained the wild-type sequence of DNA in the *lin-1* coding region and the portions of the introns that were analyzed: *ga68*, *he119*, *n1054*, *n1814*, *n1817*, *n2693*, *n2705*, and *n3443* (Table 1). To confirm that these mutations affect the *lin-1* locus, we performed an additional complementation test. Each allele failed to complement the Muv phenotype of *lin-1(e1275)* (see materials and methods). These findings suggest that the molecular lesions in these *lin-1* alleles are outside the coding regions, and these mutations are likely to affect the expression or processing of the *lin-1* mRNA.

The *lin-1(lf)* missense mutations form an allelic series: To determine how the amino acids affected by the *lin-1(lf)* missense mutations contribute to the activity of LIN-1, we analyzed the phenotypes caused by these mutations. To analyze larval development, we placed one egg or L1 larva on a petri dish and used a dissecting microscope to analyze viability. *lin-1* mutants occasionally died during the L1, L2, and L3 larval stages. *lin-1* mutants often died during the transition from the L4 larval stage to the young adult stage, and these dead hermaphrodites frequently displayed an extruded gonad, suggesting that the lethality is related to defective vulval morphogenesis. Table 2 shows the phenotypes of nine *lin-1* missense mutants, wild type, and *lin-1(sy254)*, a putative null mutation that causes a rearrangement and deletion of the *lin-1* locus. The *lin-1(sy254)* strain displayed 6% lethality during the L1 to L3 larval stages and 54% lethality during the L4 to young adult transition for a total of 60% lethality (Table 2, line 2). The nine *lin-1* missense mutant strains displayed a wide range of lethality. The penetrance of all larval lethality ranged from 48% for *lin-1(n1047)* to 2% for *lin-1(ga56)* (Table 2, lines 3–11). These results indicate that these nine missense mutations reduce *lin-1* activity to different extents and can be ordered in an allelic series.

To analyze adult phenotypes, we used a dissecting microscope to determine the number of pseudovulval protrusions of animals that survived to the adult stage. The *lin-1(sy254)* strain displayed a Muv phenotype that was 100% penetrant, and these mutants had an average of 3.6 pseudovulvae (range of 3–6; Table 2, line 2). Eight strains containing *lin-1* missense mutations displayed a Muv phenotype that was 98–100% penetrant and had an average of 3.2–3.9 pseudovulvae (Table 2, lines 3–10). The *lin-1(ga56)* mutant strain displayed a Muv phenotype with a penetrance of 86%, and these mutants had

a smaller average number of 2.5 pseudovulvae (range of 0–4; Table 2, line 11). We analyzed the overall function of the egg-laying system. *lin-1(sy254)* and five *lin-1* missense mutants displayed a highly penetrant egg-laying defect and died from internally hatched progeny without depositing eggs on the petri dish. By contrast, four *lin-1* missense mutants laid abundant eggs on the petri dishes, demonstrating the presence of a functional vulval passageway (Table 2).

We analyzed reproductive function by determining the number of animals that survived to adulthood and did not produce live progeny. The cellular basis for this *lin-1* sterile phenotype has not been established. The *lin-1(sy254)* strain displayed 16% sterility; seven of the *lin-1(lf)* missense mutants displayed a similar value (Table 2, lines 2–9). By contrast, the *lin-1(n753)* and *lin-1(ga56)* mutant strains had a lower value of 2% sterility (Table 2, lines 10–11). The analysis of the adult phenotypes, like the analysis of the larval phenotypes, indicated that the missense mutations reduce *lin-1* activity to different extents and clearly identified *lin-1(ga56)* as the weakest loss-of-function mutation.

***lin-1(ga56)* causes a temperature-sensitive phenotype:** The *lin-1(e1275 R175Opal)* mutation causes a Muv phenotype that is temperature sensitive (Beitel *et al.* 1995). To determine if any of the *lin-1(lf)* missense mutations cause a Muv phenotype that is temperature sensitive, we analyzed strains containing the missense alleles at 15°, 20°, and 25°. The Muv phenotype caused by the seven missense mutations *n753*, *n1047*, *n1848*, *n1816*, *n2704*, *n2714*, and *n2750* displayed a similar penetrance at these three temperatures (data not shown). By contrast, the Muv phenotype caused by *lin-1(ga56)* displayed a penetrance of 53% at 15° ($N = 167$), 83% at 20° ($N = 171$), and 87% at 25° ($N = 166$), demonstrating that the *lin-1(ga56)* Muv phenotype is heat sensitive. For comparison, the penetrance of the Muv phenotype caused by *lin-1(e1275 R175Opal)* was 65% at 15° ($N = 972$), 91% at 20° ($N = 911$), and 100% at 25° ($N = 863$).

Wild-type LIN-1 binds DNA: The predicted LIN-1 protein contains a highly conserved ETS domain (Figure 1C). The ETS domains of a variety of proteins have been demonstrated to bind DNA that contains the core motif GGAA/T (Karim *et al.* 1990). To determine if the LIN-1 ETS domain binds DNA, we developed an electrophoretic mobility shift assay (EMSA). Purified LIN-1 protein was obtained by generating a plasmid that encodes a fusion protein containing GST and the N-terminal 278 residues of LIN-1, GST:LIN-1(1–278) (Figure 2A). The protein was expressed in *E. coli* and partially purified by GST-affinity chromatography. The concentration of purified protein was estimated by resolving the partially purified extract by SDS-PAGE, staining with Coomassie blue, and comparing the intensity of GST:LIN-1(1–278) to known concentrations of bovine serum albumin.

To determine if the purified LIN-1 protein can bind DNA, we chose to use a DNA fragment from the Drosophila E74 gene that contains the core motif GGAA and displays high-affinity binding to several ETS proteins (Figure 2B; Shore and Sharrocks 1995; Shore *et al.* 1996). A radiolabeled, 32-bp DNA fragment was incubated with purified GST:LIN-1(1–278) protein, resolved on a nondenaturing polyacrylamide gel, and visualized. The DNA displayed similar mobility with no added protein and GST

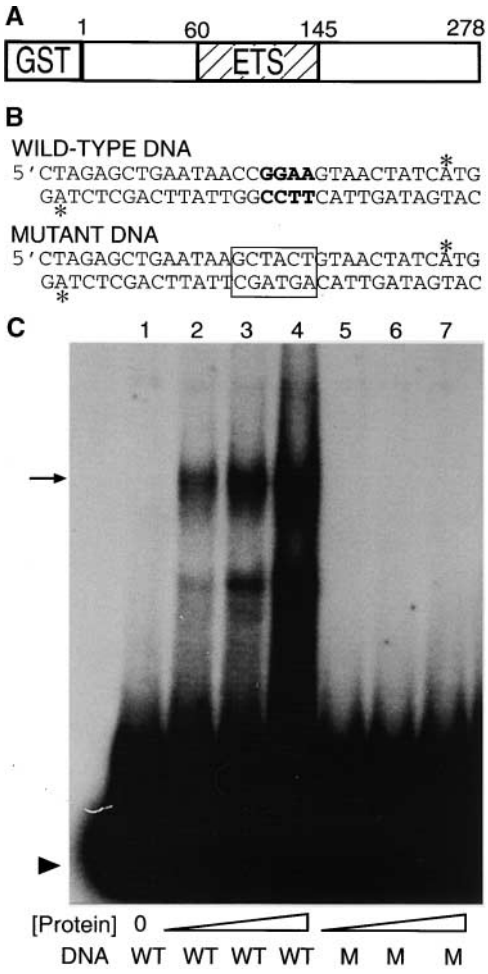

Figure 2.—LIN-1 has sequence-specific DNA-binding activity. (A) Diagram of the fusion protein containing GST (amino acids 1–222; not drawn to scale) and LIN-1 amino acids 1–278 (numbered above). (B) Sequences of the wild-type DNA derived from the Drosophila E74 gene and the mutant DNA. The core GGAA motif necessary for ETS domain binding is in boldface type, and the mutated nucleotides are boxed. (*) denotes ^{32}P-labeled adenine. (C) Phosphorimage of an EMSA. Lanes are numbered above. The concentration of GST:LIN-1(1–278) protein is indicated below: lane 1, no protein; lanes 2 and 5, 8 ng of protein; lanes 3 and 6, 16 ng of protein; lanes 4 and 7, 32 ng of protein. Wild-type (WT) DNA was included in lanes 1–4 and mutant (M) DNA in lanes 5–7. The arrowhead indicates unbound DNA; the arrow indicates DNA that is bound to GST:LIN-1(1–278) and displays retarded mobility. The lower, less prominent band is also likely a complex of protein and DNA. To quantify binding, we measured the band indicated by the arrow.

alone, indicating that GST does not bind the DNA (Figure 2C, lane 1, and data not shown). The addition of GST:LIN-1(1–278) caused the appearance of a slowly migrating band (Figure 2C, lane 2). The intensity of this band was proportional to the amount of GST:LIN-1(1–278) added to the reaction (Figure 2C, lanes 2–4). The band was eliminated by a mutation of the core GGAA motif (Figure 2C, lanes 5–7). These results indicate that the band is a complex between GST:LIN-1(1–278) and wild-type DNA and suggest that LIN-1 protein has a high-affinity, sequence-specific DNA-binding activity.

LIN-1 mutant proteins displayed reduced DNA binding: The *lin-1(lf)* missense mutations affect conserved residues in the ETS domain (Figure 1C). To investigate how these amino acid substitutions affect DNA binding, we used site-directed mutagenesis to engineer eight *lin-1(lf)* missense mutations into plasmids encoding GST:LIN-1(1–278). Mutant proteins were expressed in bacteria or insect cells and partially purified by affinity chromatography. Binding experiments with mutant GST:LIN-1(1–278) were conducted in parallel with wild-type GST:LIN-1(1–278) for comparison. To determine how the protein concentration affects DNA binding, we used six different protein amounts in the binding reactions: 0, 1.56, 3.1, 6.25, 12.5, 25, and 50 ng (Figure 3, left). The results of the EMSA were quantified by using a phosphorimager to measure the radioactivity in the band corresponding to the protein:DNA complex (Figure 3, right).

Six mutant proteins displayed no detectable DNA binding at any of the protein concentrations analyzed (Figure 3, B–G). The results suggest that these six mutant proteins have a dramatic reduction in DNA-binding affinity. The GST:LIN-1(1–278 Y127F) protein displayed no DNA binding in reactions containing 1.56, 3.13, 6.25, and 12.5 ng of protein and significantly diminished DNA binding in reactions containing 25 and 50 ng of protein (Figure 3H). The GST:LIN-1(1–278 L70F) protein displayed reduced but detectable DNA binding in reactions containing 6.25, 12.5, 25, and 50 ng of protein (Figure 3A). Therefore, these two amino acid substitutions reduce but do not eliminate

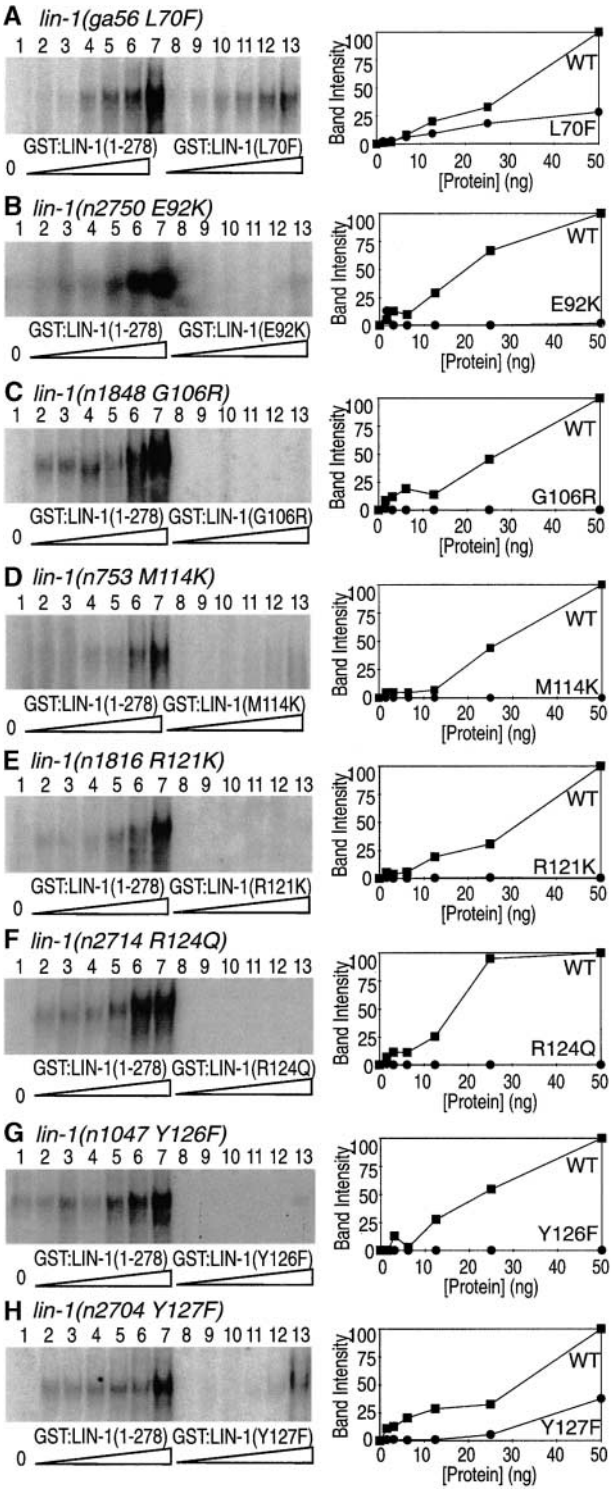

FIGURE 3.—Missense changes reduce or eliminate LIN-1 DNA binding. A–H have a similar organization. We performed an EMSA analysis using wild-type and mutant proteins in parallel. (Left) A phosphorimage of the gel with lanes numbered above. Each reaction contained radiolabeled wild-type DNA. Lane 1 was a control reaction containing no protein (0). Lanes 2–7 contained 1.56, 3.13, 6.25, 12.5, 25, and 50 ng of wild-type GST:LIN-1(1–278), respectively. Lanes 8–13 contained 1.56, 3.13, 6.25, 12.5, 25, and 50 ng of the indicated GST:LIN-1 mutant protein, respectively. Each LIN-1 mutant protein contained LIN-1 residues 1–278. The band in lanes 2–7 is the complex of GST:LIN-1(1–278) and the radiolabeled DNA and corresponds to the band indicated by an arrow in Figure 2C. We quantified the intensity of the band in each lane using a phosphorimager. The signals were adjusted by subtracting the background intensity of lane 1 (0 ng of protein) and normalized by setting the signal in lane 7 equal to 100. These data are graphed on the right side. Squares show wild-type GST:LIN-1(1–278) from lanes 2–7, and circles show mutant GST:LIN-1 from lanes 8–13. The wild-type and mutant LIN-1 proteins in A, B, and D were expressed in insect cells. The wild-type and mutant LIN-1 proteins in C and E–H were expressed in *E. coli*. Wild-type LIN-1(1–278) derived from *E. coli* and insect cells displayed similar DNA binding.

the DNA-binding activity of LIN-1. The *lin-1(ga56L70F)* mutation results in a mutant protein that retains the most DNA-binding activity in this group of eight mutant LIN-1 proteins.

DISCUSSION

Identification of *lin-1* alleles: A large collection of *lin-1* alleles has been generated by conducting many different genetic screens in many different laboratories. Because these alleles were created by random chemical mutagenesis and isolated by screening for altered nematode morphology, this collection represents a large scale and relatively unbiased investigation of the residues and domains of LIN-1 that are important for function in an animal. Here we describe the isolation of 23 *lin-1* alleles in genetic screens for the Muv phenotype, the protruding vulva phenotype, the extruded gonad phenotype, and the suppression of the *let-341* larval lethal phenotype. Twenty *lin-1(lf)* mutations were previously identified in screens for the Muv and protruding vulva phenotypes (Horvitz and Sulston 1980; Ferguson and Horvitz 1985; Beitel *et al.* 1995; Eisenmann and Kim 2000; Thomas *et al.* 2003). Thus, a total of 43 *lin-1(lf)* mutations have now been described. Seven *lin-1* gain-of-function mutations were previously identified in screens for suppression of Muv, enhancement of Vul, or larval lethal phenotypes (Jacobs *et al.* 1998; Rocheleau *et al.* 2002).

To exploit these *lin-1* alleles for a structure/function study, we analyzed the molecular lesions. Here we describe the molecular analysis of 28 *lin-1* alleles. As a result of these studies and previous investigations, each of these 50 loss-of-function and gain-of-function *lin-1* alleles has now been characterized molecularly. All 11 missense mutations that cause a loss of function cluster in the ETS domain, and all four missense mutations that cause a gain of function cluster in the C-terminal region. The 18 nonsense mutations are distributed from codon 99 to codon 390. One mutation affects a splice site. Eight alleles have deletions or rearrangements of the locus. Eight alleles have the wild-type sequence in the coding region, suggesting that these mutations affect the regulation of *lin-1* expression.

The analysis of these mutations indicates that many of the chemically inducible single base pair mutations that can significantly affect the function of *lin-1* have been identified. In seven cases, the identical mutation was isolated twice (R121K, Y126F, P384L, E99Ochre, W105Amber, Q309Amber, R352Opal), and in one case, the identical mutation was isolated four times (Q390Amber). In three cases, two different mutations that affect the same codon were isolated (G106R/E, P384S/L, W105Amber/Opal). Overall, 9 of the 20 codons affected by nonsense or missense mutations were mutated in two or more independently identified strains. The identification of the same molecular lesion in two independently derived *lin-1* alleles strongly supports the conclusion that the phenotype is caused by the molecular lesion and does not require a second mutation in the genome. Furthermore, the finding that nearly half of the codons affected by single base changes were affected in two or more strains indicates that molecular lesions of *lin-1* have been heavily sampled during these different genetic screens. EMS, the mutagen used in most of these experiments, causes primarily GC-to-AT changes and therefore primarily affects certain codons (Coulondre and Miller 1977). It is likely that most EMS-induced mutations of *lin-1* that cause a strong Muv or Vul phenotype have been identified.

***lin-1* missense mutations identify two important functions of LIN-1—regulation by ERK and DNA binding—and identify eight residues of the LIN-1 ETS domain that are necessary for DNA binding:** The *lin-1* missense mutations form two clusters. The four gain-of-function mutations affect the C terminus of *lin-1*. Three mutations affect proline 384, which is part of the FXFP motif. Biochemical studies demonstrated that the FXFP motif functions as a docking site for ERK MAP kinase (Jacobs *et al.* 1998, 1999). The proline 384 mutations diminish the binding affinity of LIN-1 and ERK MAP kinase, resulting in LIN-1 protein that is not effectively regulated and is constitutively active as an inhibitor of the 1° vulval cell fate. The *lin-1(cs50gf)* mutation affects proline 315 and may affect phosphorylation of serine 314 by ERK (Rocheleau *et al.* 2002). Together, the *lin-1(gf)* mutations identify a C-terminal domain of LIN-1 that binds ERK MAP kinase and is necessary for regulation of LIN-1 and induction of the 1° vulval cell fate in P6.p.

The 11 loss-of-function *lin-1* missense mutations cluster in the ETS domain. Several ETS domains have been shown to bind DNA with the core motif GGAA/T (Karim *et al.* 1990; Shore and Sharrocks 1995; Shore *et al.* 1996). The structure of ETS domains bound to DNA has been characterized using X-ray crystallography (Kodandapani *et al.* 1996; Mo *et al.* 1998, 2000). ETS domains exhibit a winged helix-loop-helix topology. The α3 DNA recognition helix is embedded in the major groove of the DNA and has base-specific interactions with the GGA core motif (Figure 1C). The β3-turn-β4 region, termed the winged segment, and the turn between the α2 and α3 helices make extensive contacts to the phosphate backbone of the DNA. The ETS domain of LIN-1 has been highly conserved during evolution and displays extensive similarity to the ETS domain of human Elk-1. Here we demonstrate that the LIN-1 ETS domain has sequence-specific DNA-binding activity that requires the GGAA core motif. Therefore, LIN-1 and vertebrate Elk-1 share a conserved sequence and a conserved function of DNA binding.

We analyzed how the *lin-1(lf)* missense substitutions affect DNA binding; eight missense substitutions eliminate or significantly reduce DNA binding. The *lin-1(n303 R121K)* and *lin-1(n2714 R124Q)* mutations affect

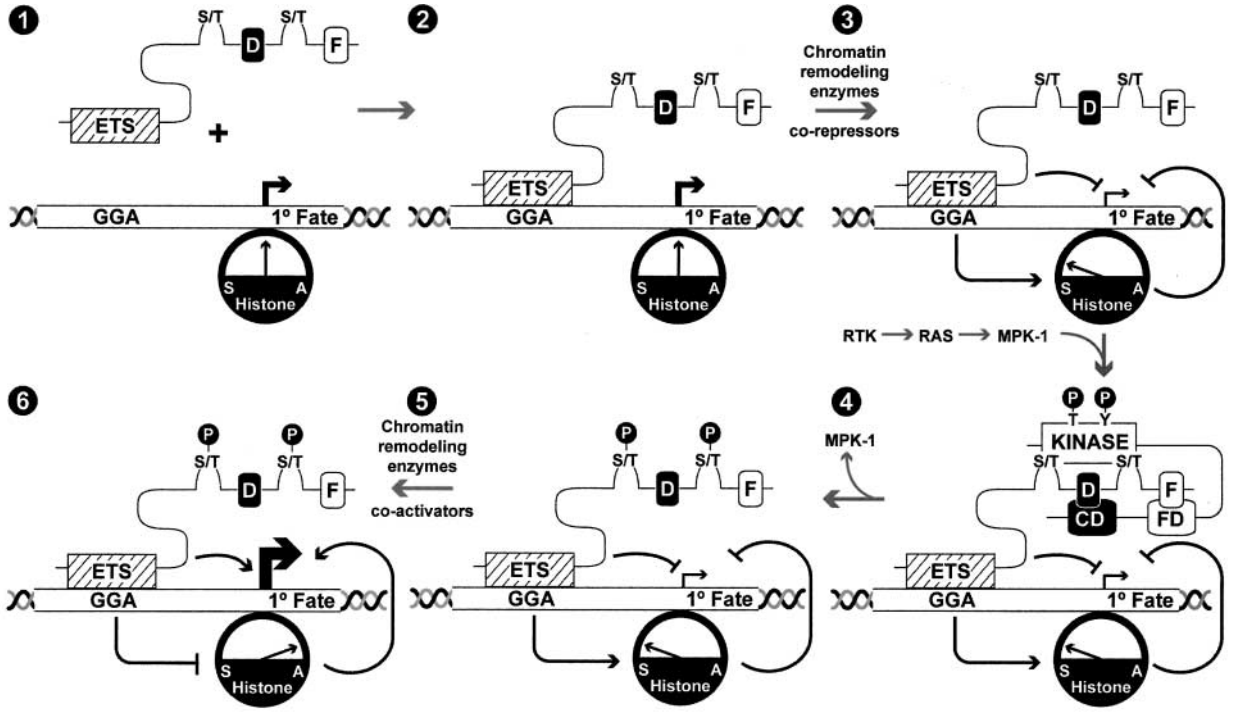

FIGURE 4.—A model of LIN-1 regulation of transcription. LIN-1 is diagrammed with the ETS domain (diagonal lines), D domain (solid), FXFP motif (open), and serine/threonine residues that can be phosphorylated by ERK (S/T). A LIN-1 target gene that promotes the 1° vulval cell fate (1° fate) contains the nucleotide sequence GGA. An interacting histone can vary between silencing (S) and activating (A). The size of the arrow indicates the level of target gene transcription. MPK-1 ERK MAP kinase is diagrammed with the kinase domain, a CD domain that interacts with the D domain (TANOUE *et al.* 2001), and a FD domain that interacts with the FXFP motif. The position of the FD domain on ERK has yet to be defined.

two arginine residues that are in the α3 recognition helix of Elk-1 and make extensive contacts with the GGA core motif (Figure 1C, solid arrowheads). The mutant proteins displayed no detectable DNA binding. The *lin-1(n1047 Y126F)* and *lin-1(n2704 Y127F)* mutations affect two tyrosine residues that are in the α3 recognition helix of Elk-1 and make contacts with nucleotides outside the GGA core (Figure 1C, open arrowheads). The LIN-1(Y126F) protein displayed no detectable DNA binding, and the LIN-1(Y127F) protein displayed very weak DNA binding. The *lin-1(n753 M114K)* mutation affects a methionine in the turn region of Elk-1 that makes contacts with the GGA core (Figure 1C, solid arrowheads). This mutant protein displayed no detectable DNA binding. These results are consistent with the model that the ETS domain of LIN-1 adopts a structure that is similar to the structure of Elk-1. These studies contribute to an understanding of the ETS domain by demonstrating that residues predicted to contact the DNA are essential for DNA binding.

Three missense substitutions affect residues that do not directly contact DNA but are highly conserved nonetheless. *lin-1(n2750 E92K)* affects a glutamic acid in the β2 turn and *lin-1(n1848 G106R)* affects a glycine in the α2 helix (Figure 1C). Both of these mutant LIN-1 proteins had no detectable DNA binding. The *lin-1(ga56 L70F)* mutation affects a leucine in the α1 helix. This mutant LIN-1 protein had reduced but still detectable DNA binding. These data indicate that these residues are important for the conformation of the ETS domain and may be necessary to correctly position the residues that directly contact the DNA.

DNA binding is necessary for LIN-1 function in animals: In addition to determining how the *lin-1(lf)* substitutions affect DNA binding, we determined how these mutations affect *lin-1* function in animals. These mutations reduce the activity of *lin-1* at multiple developmental stages and can be arranged in an allelic series. The strongest mutations cause a phenotype that is similar to the phenotype caused by a *lin-1* null mutation, whereas the *lin-1(ga56 L70F)* mutation causes a milder phenotype, indicating that it partially reduces *lin-1* activity. An analysis of the DNA-binding activity of the mutant proteins demonstrated that each mutant protein had a profound defect in DNA binding and that the LIN-1(L70F) mutant retained the most DNA-binding activity. These results establish a correlation between the DNA-binding activity of the mutant LIN-1 protein and the function of the mutant *lin-1* allele in animals, indicating that the defects in DNA binding are the cause of mutant phenotypes. The expression level of the mutant LIN-1 proteins was not examined; some of these mutations may also affect LIN-1 protein stability. The similarity between the strong missense mutations and the *lin-1* null

mutation suggests that DNA binding is necessary for all of the functions of *lin-1*.

Although the structures of ETS domains have been analyzed extensively, few studies have investigated the function of specific residues for DNA binding and activity in animals. Golay *et al.* (1988) identified one missense mutation in the *v-ets* oncogene that reduces the ability of the virus to transform cells. The analogous substitution in Elk-1 (R74D) reduced DNA affinity ~10-fold (Janknecht and Nordheim 1992). These findings indicate that DNA binding is important for transformation activity of the viral ETS gene. Residues of PU.1 that are necessary for DNA binding were identified by making glycine substitutions (Kodandapani *et al.* 1996). Tootle *et al.* (2003) introduced two analogous substitutions in Drosophila Yan (W439G and K443G) and demonstrated that the mutant Yan protein was mislocalized to the cytoplasm of cultured cells. These findings indicate that DNA binding is important for nuclear localization of Yan. The results presented here show that eight different *lin-1* missense mutations that result in reduced DNA binding also reduce or eliminate function in animals. Thus, DNA binding is necessary for the function of LIN-1, and this may be the conserved property of ETS proteins.

On the basis of these genetic and biochemical studies, we propose a model of LIN-1 function shown in Figure 4. Newly synthesized LIN-1 protein enters the nucleus, and the ETS domain binds to GGA motifs in the promoters of target genes (Figure 4, parts 1 and 2). These target genes are likely to promote the 1° vulval cell fate, and LIN-1 may repress the transcription of these target genes by interacting with corepressors and chromatin-remodeling enzymes that affect histones bound to the promoter and condense the chromatin (Figure 4, part 3). When the anchor cell signals P6.p, resulting in the activation of LET-23 RTK, LET-60 Ras, and MPK-1 ERK, phosphorylated MPK-1 enters the nucleus. MPK-1 interacts with the D domain and the FXFP motif of LIN-1 and phosphorylates one or more serine/threonine residues (Figure 4, parts 4 and 5). Phosphorylation may cause LIN-1 to activate transcription of target genes, and phosphorylated LIN-1 might recruit coactivators and/or chromatin-remodeling enzymes that decondense the chromatin (Figure 4, part 6). It is also possible that phosphorylation prevents LIN-1-mediated repression but does not cause LIN-1 to activate transcription. The genetic and biochemical studies of *lin-1(lf)* missense mutations support the importance of the ETS domain interaction with DNA shown in parts 1 and 2 of Figure 4. The genetic and biochemical analyses of *lin-1(gf)* missense mutations support the importance of interaction with MPK-1 (Figure 4, parts 4 and 5). It is notable that domains of LIN-1 that interact with corepressors, coactivators, and chromatin-remodeling enzymes were not defined by the collection of LIN-1 missense mutations. Thus, the mechanisms of LIN-1 transcriptional regulation shown in Figure 4, parts 3 and 6, have yet to be well defined. Biochemical screens for proteins that interact with LIN-1 are a promising approach that may complement the genetic screens described here and illuminate these aspects of LIN-1 function.

We are grateful to the following individuals for providing *lin-1* alleles: Bob Horvitz and members of his lab including Greg Beitel, Jeff Thomas, Chip Ferguson, Erik Jorgenson, Scott Clark, Craig Ceol, Frank Stegmeier, Melissa Harrison, and Na An as well as Joseph Lee, Dave Eisenmann, and Stuart Kim. We thank Parie Garg and Andrew Turk for assistance with experiments. Some strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the National Center for Research Resources of the National Institutes of Health (NIH). This work was supported by grants from the NIH to K.K. (CA-84271). K.K. is a scholar of the Leukemia and Lymphoma Society. P.W.S is an Investigator and R.E.P. was an associate with the Howard Hughes Medical Institute. D.F. was supported by an NIH grant (GM-20450-01). R.M.S. has a postdoctoral fellowship from the American Cancer Society and the Massachusetts General Hospital Fund for Medical Discovery.

LITERATURE CITED

Ausubel, F. M., 1987 *Current Protocols in Molecular Biology*. John Wiley & Sons, New York.

Beitel, G. J., S. Tuck, I. Greenwald and H. R. Horvitz, 1995 The *Caenorhabditis elegans* gene *lin-1* encodes an ETS-domain protein and defines a branch of the vulval induction pathway. Genes Dev. **9:** 3149–3162.

Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94.

Chang, C., N. A. Hopper and P. W. Sternberg, 2000 *Caenorhabditis elegans* SOS-1 is necessary for multiple RAS-mediated developmental signals. EMBO J. **19:** 3283–3294.

Clark, S. G., M. J. Stern and H. R. Horvitz, 1992 Genes involved in two *Caenorhabditis elegans* cell-signaling pathways. Cold Spring Harbor Symp. Quant. Biol. **57:** 363–373.

Coulondre, C., and J. H. Miller, 1977 Genetic studies of the lac repressor. IV. Mutagenic specificity in the lacI gene of *Escherichia coli*. J. Mol. Biol. **117:** 577–606.

Eisenmann, D. M., and S. K. Kim, 2000 Protruding vulva mutants identify novel loci and Wnt signaling factors that function during *Caenorhabditis elegans* vulva development. Genetics **156:** 1097–1116.

Fantz, D. A., D. Jacobs, D. Glossip and K. Kornfeld, 2001 Docking sites on substrate proteins direct extracellular signal-regulated kinase to phosphorylate specific residues. J. Biol. Chem. **276:** 27256–27265.

Ferguson, E. L., and H. R. Horvitz, 1985 Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. Genetics **110:** 17–72.

Ferguson, E. L., and H. R. Horvitz, 1989 The multivulva phenotype of certain *Caenorhabditis elegans* mutants results from defects in two functionally redundant pathways. Genetics **123:** 109–121.

Ferguson, E. L., P. W. Sternberg and H. R. Horvitz, 1987 A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*. Nature **326:** 259–267.

Giovane, A., A. Pintzas, S. M. Maira, P. Sobieszczuk and B. Wasylyk, 1994 Net, a new ets transcription factor that is activated by Ras. Genes Dev. **8:** 1502–1513.

Golay, J., M. Introna and T. Graf, 1988 A single point mutation in the *v-ets* oncogene affects both erythroid and myelomonocytic cell differentiation. Cell **55:** 1147–1158.

Greenwald, I. S., 1997 Development of the vulva, pp. 519–541 in *C. elegans II*, edited by D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Horvitz, H. R., and P. W. Sternberg, 1991 Multiple intercellular signalling systems control the development of the *Caenorhabditis elegans* vulva. Nature **351:** 535–541.

Horvitz, H. R., and J. E. Sulston, 1980 Isolation and genetic characterization of cell-lineage mutants of the nematode *Caenorhabditis elegans*. Genetics **96:** 435–454.

Jacobs, D., G. J. Beitel, S. G. Clark, H. R. Horvitz and K. Kornfeld, 1998 Gain-of-function mutations in the *Caenorhabditis elegans lin-1* ETS gene identify a C-terminal regulatory domain phosphorylated by ERK MAP kinase. Genetics **149:** 1809–1822.

Jacobs, D., D. Glossip, H. Xing, A. J. Muslin and K. Kornfeld, 1999 Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase. Genes Dev. **13:** 163–175.

Janknecht, R., and A. Nordheim, 1992 Elk-1 protein domains required for direct and SRF-assisted DNA-binding. Nucleic Acids Res. **20:** 3317–3324.

Johnsen, R. C., and D. L. Baillie, 1991 Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans*. Genetics **129:** 735–752.

Karim, F. D., L. D. Urness, C. S. Thummel, M. J. Klemsz, S. R. McKercher *et al.*, 1990 The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes Dev. **4:** 1451–1453.

Kodandapani, R., F. Pio, C. Z. Ni, G. Piccialli, M. Klemsz *et al.*, 1996 A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. Nature **380:** 456–460.

Kornfeld, K., 1997 Vulval development in *Caenorhabditis elegans*. Trends Genet. **13:** 55–61.

Kornfeld, K., K. L. Guan and H. R. Horvitz, 1995 The *Caenorhabditis elegans* gene *mek-2* is required for vulval induction and encodes a protein similar to the protein kinase MEK. Genes Dev. **9:** 756–768.

Lackner, M. R., K. Kornfeld, L. M. Miller, H. R. Horvitz and S. K. Kim, 1994 A MAP kinase homolog, *mpk-1*, is involved in ras-mediated induction of vulval cell fates in *Caenorhabditis elegans*. Genes Dev. **8:** 160–173.

Lai, Z. C., and G. M. Rubin, 1992 Negative control of photoreceptor development in *Drosophila* by the product of the *yan* gene, an ETS domain protein. Cell **70:** 609–620.

Laudet, V., C. Hanni, D. Stehelin and M. Duterque-Coquillaud, 1999 Molecular phylogeny of the ETS gene family. Oncogene **18:** 1351–1359.

Mavrothalassitis, G., and J. Ghysdael, 2000 Proteins of the ETS family with transcriptional repressor activity. Oncogene **19:** 6524–6532.

Mo, Y., B. Vaessen, K. Johnston and R. Marmorstein, 1998 Structures of SAP-1 bound to DNA targets from the *E74* and *c-fos* promoters: insights into DNA sequence discrimination by Ets proteins. Mol. Cell **2:** 201–212.

Mo, Y., B. Vaessen, K. Johnston and R. Marmorstein, 2000 Structure of the Elk-1-DNA complex reveals how DNA-distal residues affect ETS domain recognition of DNA. Nat. Struct. Biol. **7:** 292–297.

Rao, V. N., K. Huebner, M. Isobe, A. ar-Rushdi, C. M. Croce *et al.*, 1989 Elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation breakpoints. Science **244:** 66–70.

Remy, P., and M. Baltzinger, 2000 The Ets-transcription factor family in embryonic development: lessons from the amphibian and bird. Oncogene **19:** 6417–6431.

Riddle, D. L., T. Blumenthal, B. J. Meyer and J. R. Priess, 1997 *C. elegans II*. Cold Spring Harbor Laboratory Press, Plainview, NY.

Rocheleau, C. E., R. M. Howard, A. P. Goldman, M. L. Volk, L. J. Girard *et al.*, 2002 A *lin-45* raf enhancer screen identifies *eor-1, eor-2* and unusual alleles of Ras pathway genes in *Caenorhabditis elegans*. Genetics **161:** 121–131.

Sambrook, J., T. Maniatis and E. F. Fritsch, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shore, P., and A. D. Sharrocks, 1995 The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. Nucleic Acids Res. **23:** 4698–4706.

Shore, P., A. J. Whitmarsh, R. Bhaskaran, R. J. Davis, J. P. Waltho *et al.*, 1996 Determinants of DNA-binding specificity of ETS-domain transcription factors. Mol. Cell. Biol. **16:** 3338–3349.

Sternberg, P. W., and M. Han, 1998 Genetics of RAS signaling in *C. elegans*. Trends Genet. **14:** 466–472.

Sternberg, P. W., and H. R. Horvitz, 1989 The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. Cell **58:** 679–693.

Sulston, J. E., and H. R. Horvitz, 1981 Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. Dev. Biol. **82:** 41–55.

Tan, P. B., M. R. Lackner and S. K. Kim, 1998 MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31 WH transcription factor complex during *C. elegans* vulval induction. Cell **93:** 569–580.

Tanaka, M., A. Ueda, H. Kanamori, H. Ideguchi, J. Yang *et al.*, 2002 Cell-cycle-dependent regulation of human aurora A transcription is mediated by periodic repression of E4TF1. J. Biol. Chem. **277:** 10719–10726.

Tanoue, T., R. Maeda, M. Adachi and E. Nishida, 2001 Identification of a docking groove on ERK and p38 MAP kinases that regulates the specificity of docking interactions. EMBO J. **20:** 466–479.

Thomas, J. H., C. J. Ceol, H. T. Schwartz and H. R. Horvitz, 2003 New genes that interact with *lin-35* Rb to negatively regulate the *let-60* ras pathway in *Caenorhabditis elegans*. Genetics **164:** 135–151.

Tootle, T. L., P. S. Lee and I. Rebay, 2003 CRM1-mediated nuclear export and regulated activity of the receptor tyrosine kinase antagonist YAN require specific interactions with MAE. Development **130:** 845–857.

Watson, D. K., M. J. McWilliams, P. Lapis, J. A. Lautenberger, C. W. Schweinfest *et al.*, 1988 Mammalian *ets-1* and *ets-2* genes encode highly conserved proteins. Proc. Natl. Acad. Sci. USA **85:** 7862–7866.

Williams, B. D., B. Schrank, C. Huynh, R. Shownkeen and R. H. Waterston, 1992 A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. Genetics **131:** 609–624.

Wu, Y., and M. Han, 1994 Suppression of activated Let-60 ras protein defines a role of *Caenorhabditis elegans* Sur-1 MAP kinase in vulval differentiation. Genes Dev. **8:** 147–159.

Yandell, M. D., L. G. Edgar and W. B. Wood, 1994 Trimethylpsoralen induces small deletion mutations in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **91:** 1381–1385.

Yang, S. H., P. R. Yates, A. J. Whitmarsh, R. J. Davis and A. D. Sharrocks, 1998 The Elk-1 ETS-domain transcription factor contains a mitogen-activated protein kinase targeting motif. Mol. Cell. Biol. **18:** 710–720.

Yochem, J., M. Sundaram and M. Han, 1997 Ras is required for a limited number of cell fates and not for general proliferation in *Caenorhabditis elegans*. Mol. Cell. Biol. **17:** 2716–2722.

Yochem, J., K. Weston and I. Greenwald, 1988 The *Caenorhabditis elegans lin-12* gene encodes a transmembrane protein with overall similarity to *Drosophila* Notch. Nature **335:** 547–550.

Communicating editor: K. Kemphues