ELSEVIER

Gene 254 (2000) 253–263

GENE
AN INTERNATIONAL JOURNAL ON GENES, GENOMES AND EVOLUTION

www.elsevier.com/locate/gene

The *C. elegans* gene *lin-9*,which acts in an Rb-related pathway, is required for gonadal sheath cell development and encodes a novel protein

Greg J. Beitel [a,1], Eric J. Lambie [b], H. Robert Horvitz [a,*]

[a] *Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA*

Received 23 May 2000; accepted 29 June 2000
Received by J. Widom

Abstract

The *Caenorhabditis elegans* gene *lin-9* functions in an Rb-related pathway that acts antagonistically to a receptor tyrosine kinase/Ras signal transduction pathway controlling vulval induction. We show that *lin-9* is also required for the development of the sheath cells in the hermaphrodite gonad and for the development of the male spicule, rays and gonad. *lin-9* is transcribed as two alternatively spliced 2.4 kb messages, which use two distinct polyadenylation sites and are SL1 *trans*-spliced. The conceptual translation of *lin-9* cDNA sequences predicts proteins of 642 and 644 amino acids with a significant similarity to predicted *Drosophila* and vertebrate proteins. We suggest that *lin-9* is the founding member of a new protein family that functions in Rb-related pathways in many species.

Keywords: Alternative splicing; ras; Rb; Cell fate; Signal transduction; Vulval induction

1. Introduction

To understand cell–cell interactions and signal transduction pathways, we and others have studied the development of the vulva of the *Caenorhabditis elegans* hermaphrodite (reviewed by Kornfeld, 1997). During vulval development, at least three distinct sets of cell–cell interactions specify which three of six hypodermal blast cells, known as P3.p, P4.p, P5.p, P6.p, P7.p and P8.p (collectively, the P(3–8).p cells), are induced to produce descendants that form the vulva. The P(3–8).p cells are multipotent and can adopt either vulval or non-vulval fates. The cells P(5–7).p normally adopt vulval fates in response to an inductive signal from the anchor cell of the gonad. Vulval induction is mediated by the *let-23* receptor tyrosine kinase (RTK)/*let-60* Ras signaling pathway; this RTK/Ras pathway is conserved among *C. elegans*, *Drosophila* and mammals (reviewed by Kornfeld, 1997). The inductive signal appears to act by overcoming signals that prevent the expression of vulval fates and that the P(3–8).p cells receive from the surrounding syncytial hypoderm (Herman and Hedgecock, 1990). The inhibitory hypodermal-to-P(3–8).p signaling seems to be mediated by two redundant pathways such that mutations disrupting only one of the pathways do not cause abnormal vulval development (Ferguson and Horvitz, 1985, 1989). Only if at least one gene in each pathway is mutated do the cells P3.p, P4.p and P8.p inappropriately adopt vulval fates, resulting in a Multivulva (Muv) phenotype. This phenotype is termed 'synthetic Multivulva' (synMuv), and the two sets of synMuv genes have been termed 'A' and 'B' (Ferguson and Horvitz, 1989).

Whereas loss-of-function mutations in synMuv genes result in excessive vulval development, loss-of-function mutations in the RTK/Ras pathway genes prevent vulval development, indicating that the synMuv genes antagonize the *let-23* RTK/*let-60* Ras pathway activity during vulval development. That vulval development is blocked

Abbreviations: bp, base pair(s); kb, kilobase(s); EST, expressed sequence tag; Muv, Multivulva; ORF, open reading frame; polyA, polyadenosine; RTK, receptor tyrosine kinase; synMuv, synthetic Multivulva; ts, temperature-sensitive; UTR. untranslated region.

* Corresponding author. Tel.: +1-617-253-4671; fax: +1-617-253-8126.
E-mail address: horvitz@mit.edu (H.R. Horvitz)
[1] Present address: Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208, USA.

0378-1119/00/$ - see front matter
PII: S0378-1119(00)00296-1

in mutants defective in both synMuv genes and the RTK/Ras pathway suggests that at least some synMuv genes act upstream of, or parallel to, the Ras pathway (Ferguson et al., 1987; Clark et al., 1994; Huang et al., 1994; Thomas and Horvitz, 1999). At least 22 loci can be mutated to cause a synMuv phenotype (Ferguson et al., 1987; Ferguson and Horvitz, 1989; Lu and Horvitz, 1998; Hsieh et al., 1999; Solari and Ahringer, 2000; J. Thomas and H.R. Horvitz, in preparation). Two class B synMuv genes encode proteins similar to the retinoblastoma (Rb) protein and the Rb-binding protein RbAp48 (Lu and Horvitz, 1998). Rb and Rb-interacting proteins are important cellular regulators and potent oncogenes. To further analyze synMuv gene function, we cloned another class B synMuv gene, *lin-9*, and characterized several phenotypic effects of *lin-9* mutations. Our results show that *lin-9* defines a new gene family that includes the recently identified *Drosophila* protein Aly, the activity of which appears to be regulated by subcellular localization (White-Cooper et al., 2000).

2. Materials and methods

2.1. Strains and genetic nomenclature

C. elegans strain variety Bristol N2 was the wild-type parent of all mutant strains used. Strains were handled and maintained using standard methods. Genetic markers used were described by Hodgkin et al. (1988) and were as follows: LGII: *lin-8*(*n111*); LGIII: *lin-36*(*n747*), *unc-36*(*e251*), *lin-9*(*n112, n942, n943*), *unc-32*(*e189*); LGX: *lin-15*(*n433*).

2.2. Germline transformation and subclone construction

We followed the protocol of Mello et al. (1991) for germline transformation experiments using the *rol-6* pRF4 plasmid as a co-injection marker. We subcloned fragments of the cosmid C07D12 into the pBlueScript SK(+) plasmid (Stratagene Cloning Systems, La Jolla, CA). We constructed the *Xho*I–*Xho*I fragment containing the four amino acid insertion at the *lin-9* *Bam*HI site by subcloning the C07D12 7.1 kb *Xho*I–*Xho*I fragment into the *Xho*I and *Sal*I sites of pBlueScript SK(+). This plasmid was digested with *Eag*I and then briefly treated with ExonucleaseIII to remove the *Bam*HI and *Sac*I sites in the pBlueScript polylinker. The resulting plasmid was digested with *Bam*HI, treated with T4 DNA polymerase, and then religated in the presence of a 5′GACTAGTC linker sequence. One of the resulting clones had the sequence 5′ …GATCGACTAGTCGATCC… 3′ at the former *Bam*HI site, which results in an in-frame insertion of Arg Leu Val Asp. We constructed the *Xho*I–*Xho*I fragment containing the in-frame amber stop codon in the putative *lin-9* ORF using a clone obtained in the T4 DNA polymerase and linker ligation step described above with the sequence 5′ …GATCACTAGTCGATCC… 3′ at the former *Bam*HI site. We digested this plasmid with *Spe*I, filled in the ends using Klenow DNA polymerase I, and religated. This procedure produced a plasmid having the sequence 5′ …GATCACTAGCTAGTCGATCC… 3′, where the underlined TAG is an in-frame stop codon. We tested the ability of genomic fragments to rescue the sterility caused by *lin-9*(*n943*) using the strain MT1969 *lin-8*(*n111*) II; *lin-9*(*n943*)/*unc-36*(*e251*) III. Animals were judged to be fertile if an approximately normal number of eggs were visible in their uteri and many viable progeny were produced.

2.3. Northern blot hybridizations

We ^{32}P-labeled and hybridized the 7.0 kb *Xho*I–*Xho*I rescuing fragment to Northern blots of mixed-stage N2 animals.

2.4. cDNA library screens and DNA sequence analysis

We screened amplified cDNA libraries obtained from S. Kim (Kim and Horvitz, 1990) and R. Barstead (personal communication) using ^{32}P-labeled random-primed probes prepared from the 7.0 kb *Xho*I–*Xho*I genomic fragment (Fig. 4B). From a total of 540 000 plaques, we isolated 31 phage and determined the 5′ and 3′ end-sequences of the cDNA inserts using the forward and reverse primers for the pBluescript plasmid. Twenty-one of the 31 phage contained distinct inserts. We generated subclones for sequence determination of the longest Barstead and longest Kim library cDNAs by shotgun cloning and Exonuclease III deletion, respectively. We determined DNA sequences using single- and double-stranded templates processed using the Applied Biosystems Prism™ Cycle Sequencing kits and a Model 373A DNA sequencer (Applied Biosystems, Foster City, CA) or using Sequenase DNA sequencing kits (United States Biochemical, OH). In order to analyze the genomic sequences of *lin-9* alleles, we determined the complete genomic sequences of the exons and at least 15 bp on either side of exon splice sites using cycle sequencing with ^{33}P γ-ATP as a label. Sequencing reactions were analyzed on polyacrylamide gels. All differences from the wild-type (WT) sequence are shown in Fig. 5.

2.5. Microscopy and analysis of the germline phenotype

All animals for germline phenotypic analysis were raised at 19°C, except where noted in the text. *unc-32*(*e189*) *lin-9*(*n942*) animals were obtained from *unc-32*(*e189*) *lin-9*(*n942*)/*qC1* parents. We were unable to directly observe the hatching rate of *unc-32*(*e189*) *lin-9*(*n942*) animals from homozygous parents since, many *unc-32 (e189) lin-9 (n942)* animals were egg-laying

defective and larvae were initially trapped within the mother's cuticle. We therefore determined the percentage hatching of eggs of genotype *unc-32*(*e189*) *lin-9*(*n942*) among eggs obtained by cutting open *lin-9(n942) unc-32(e189)* hermaphrodites. As a control for this surgical procedure, we determined that eggs obtained by cutting open *fog-1*(*q253*); *unc-32*(*e189*) hermaphrodites had a 95% hatching rate ($n=351$).

lin-9(n942) and *lin-9(n943)* homozygotes were distinguished from their wild-type siblings and selected for high-magnification microscopy and/or cytochemical staining based on their relatively small sizes, and for hermaphrodites, protruding vulva and absence of organized rows of eggs in the uterus. Microscopy was performed using a Nikon microphot-SA equipped with standard Nomarski differential interference contrast optics and epifluorescence filters. Images were captured using a Dage CCD-300-RC camera connected to a Scion LGIII frame grabber installed in a Macintosh Quadra 700 running NIH Image (as modified by Scion). In most cases, a single round of sharpening was performed in NIH Image, and then images were transferred to Adobe Photoshop for assembly of composites, adjustment of brightness and contrast and labeling of key features. For Nomarski imaging, animals were mounted on agarose pads and immobilized by treatment with 2 mM levamisole or 30 mM sodium azide. We used the method of Church et al. (1995) for DAPI staining of whole animals and the method of Francis et al. (1995) for DAPI and anti-CEH-18 staining of gonads dissected from adult hermaphrodites.

3. Results and discussion

3.1. lin-9 *is required for hermaphrodite gonad development*

To determine the basis of the previously reported sterility caused by strong *lin-9* alleles (Ferguson and Horvitz, 1989), we examined the fertility and reproductive structures of *lin-9(n942)* and *lin-9(n943)* mutant animals. The defects caused by the *lin-9(n942)* and *lin-9(n943)* mutations are very similar and possibly identical, although we have characterized *lin-9(n942)* in more detail than *lin-9(n943)* (Table 1). Hermaphrodites homozygous for *lin-9(n942)* are not completely sterile, but rather have an average brood size of 1.1 [$n=59$, genotype *unc-32*(*e189*) *lin-9*(*n942*)], which is insufficient to allow this allele to be maintained in a homozygous strain (Table 1). This small brood size does not result from a failure to make functional sperm, because mating wild-type males with *lin-9(n942)* hermaphrodites did not dramatically increase the brood size (data not shown). At least three factors contributed to the reduction in brood size. First, the integrity of the gonad is not complete: sperm, and sometimes oocytes, were present within the body cavity in most animals (Table 1 and footnote). Second, 71% of the eggs that were fertilized in *lin-9(n942)* hermaphrodites arrested development as embryos and failed to hatch [$n=59$, genotype *unc-32*(*e189*) *lin-9*(*n942*)]. Third, more than 85% of *lin-9(n942)* and *lin-9(n943)* hermaphrodites had endomitotic oocytes (the 'Emo' phenotype described by Iwasaki et al., 1996) within the proximal oviduct, indicative of a defect in ovulation (Table 1, Fig. 1B and D).

The reproductive system abnormalities observed in *lin-9(n942)* hermaphrodites are also seen in animals that lack normal gonadal sheath-cell function (Greenstein et al., 1994; Myers et al., 1996; McCarter et al., 1997). We therefore examined the sheath cells in *lin-9(n942)* animals using an anti-CEH-18 antibody to stain sheath-cell nuclei in gonads dissected from adult *lin-9(n942)* hermaphrodites (Greenstein et al., 1994). We found that *lin-9(n942)* hermaphrodite gonads contained four instead of eight sheath cells (Table 1). Thus, *lin-9* is required for correct development of the hermaphrodite sheath cells. It is not known if cell signaling controls

Table 1
Lin-9 hermaphrodite phenotypes[a]

Genotype	Emo[b] (n)	Ectopic sperm[c] (n)	Fog arms[c] (n)	Sheath cell number[d] (n)
lin-9(n942)/+[e]	0% (76)	0% (76)	0% (152)	8 (13)
lin-9(n942)[f]	85% (41)	32% (41)	0.7% (136)	4 (24)
lin-9(n943)[g]	88% (52)	31% (52)	n.d.	n.d.

[a] n.d., not determined.

[b] Animals were stained with DAPI and scored by epifluorescence for the presence of endomitotic (Emo) oocytes in one or both gonad arms.

[c] Animals were stained with DAPI and scored by epifluorescence for the presence and location of sperm. Sperm were scored as ectopic if they were present in the body cavity and/or distal gonad. Fog, feminization of the germline in the anterior or posterior gonadal arms was indicated by the absence of sperm in the spermatheca/proximal oviduct.

[d] cGonads were dissected and stained with an anti-CEH-18 antibody (Greenstein et al., 1994). Sheath cell nuclei were identified by immunofluorescence and DAPI staining. The pair of sheath cell nuclei most proximal to the spermatheca could not be consistently identified, even in wild-type animals, so only the remaining eight sheath cells were scored for each gonadal arm.

[e] Genotype *lin-9(n942)/qCl*.

[f] Obtained from *lin-9(n942)/qCl* animals; 2/41 animals had endomitotic oocytes in the body cavity 60–84 h after hatching.

[g] Obtained from *lin-9(n943)/qCl* animals; 3/52 animals had endomitotic oocytes in the body cavity 60–84 h after hatching.

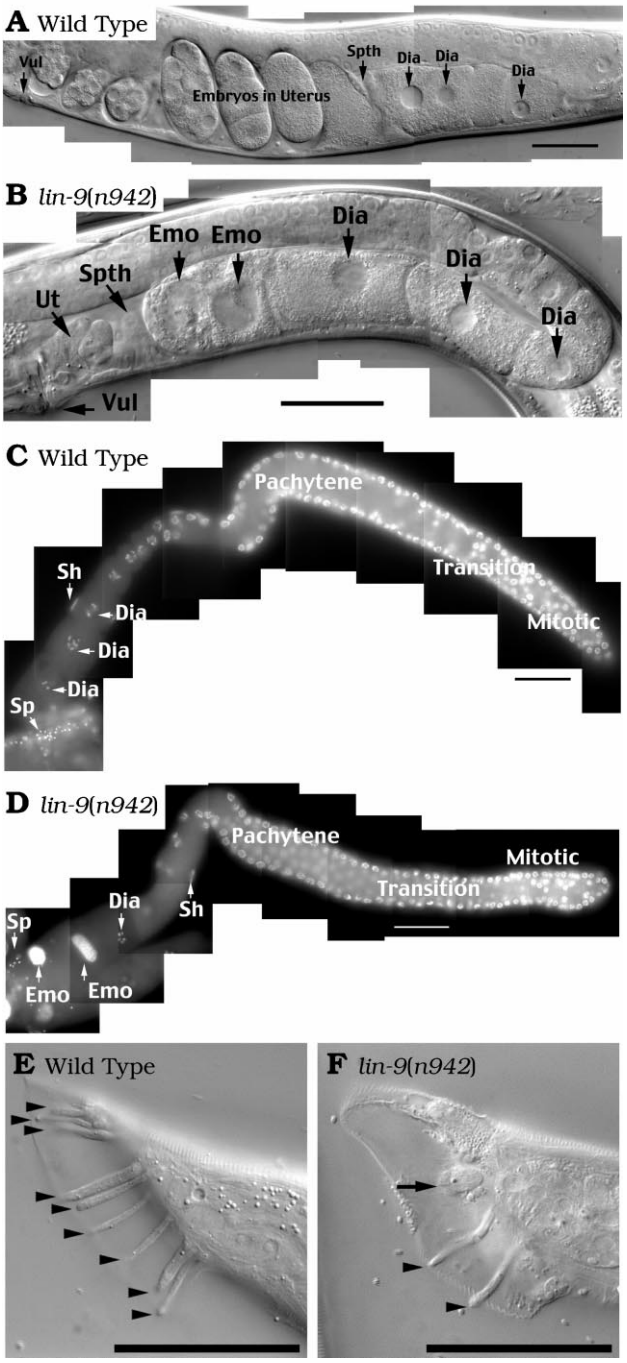


Fig. 1. Reproductive organ defects in adult *lin-9* mutant animals. Scale bars, 40 μm. (**A**, **B**) Nomarski optics photomicrographs of hermaphrodite gonads. The large nuclei in oocytes in the proximal oviduct of *lin-9(n942)* mutants are characteristic of the endomitotic (Emo) phenotype (see Section 3.1). Dia, diakinesis-stage oocyte; Spth, spermatheca; Vul, vulva; Ut, uterus; Emo, endomitotic oocyte. (**C**, **D**) Epifluorescent images of DAPI-stained hermaphrodite gonads. Note the brightly staining Emo oocytes in *lin-9(n942)* but not wild-type animals. Sh, somatic sheath-cell nucleus; Sp, sperm; mitotic, transition and pachytene zones of the germline are indicated. (**E**, **F**) Nomarski images of male tails. *lin-9(n942)* males have a reduced number of sensory rays. Arrowheads, sensory rays. Arrow, malformed sensory ray.

sheath cell development or if *lin-9* antagonizes the *let-23* RTK/*let-60* Ras pathway function in this process. Since *lin-9* and other class B synMuv mutations cause hermaphrodite sterility even in the presence of wild-type class A synMuv activity and since no class A synMuv mutations have yet been reported to cause a sterile phenotype (Ferguson and Horvitz, 1989; J. Thomas and H. R. Horvitz, in preparation), it appears that the class B synMuv genes act in a non-redundant pathway in hermaphrodite sheath-cell development.

3.2. lin-9 is required for male reproductive system development

lin-9 is also required for the development of the male reproductive system. The tails of *lin-9(n942)* mutant males had fewer than half the normal number of sensory rays (Fig. 1F, Table 2), and the copulatory spicules were crumpled and often reduced in size (Table 2). In many *lin-9(n942)* males, spermatids entered the space between the gonadal basement membrane and the germline epithelium and slipped back distally, away from the seminal vesicle (Fig. 2B, Table 2). Thus, *lin-9* is required for the development of the male reproductive system even in the presence of normal class A synMuv gene activity. There have been no previous reports of the synMuv or the *let-23* RTK/*let-60* Ras pathways being required for male gonad or ray development, so it is unclear what cells or signal transduction pathways are involved in the development of these structures. The synMuv genes were previously reported to negatively regulate the *let-23* RTK/*let-60* Ras pathway during male spicule development (Chamberlin and Sternberg, 1994). However, those experiments used *lin-15* mutants defective in both class A and class B synMuv activities (Ferguson and Horvitz, 1985; Clark et al., 1994; Huang et al., 1994). Since the strong *lin-9(n942)* mutation is defective only in class B synMuv activity and by itself disrupts male spicule development, presumably by affecting the same pro-

Table 2
Lin-9 male phenotypes

Genotype	Number of rays[a]	Spicule defects[b]	Ectopic sperm[c]	*n*
unc-32 lin-9(n942)/unc-32[d]	8.7	0%	0%	25
unc-32 lin-9(n942)[e]	2.0	81%	69%	16
lin-9(n942)[f]	3.8	89%	53%	19

[a] aNumber of wild-type rays visible on the upper side of the tail.
[b] bReduction in size and/or crumpling of spicules.
[c] cPresence of spermatids distal to the seminal vesicle.
[d] dGenotype *fog-1(q253)/+*; *lin-9(n942) unc-32(e189)/unc-32(e189)*.
[e] eObtained from *lin-9(n942) unc-32(e189)/qC1* parents.
[f] fObtained from *lin-9(n942)/qC1* parents.

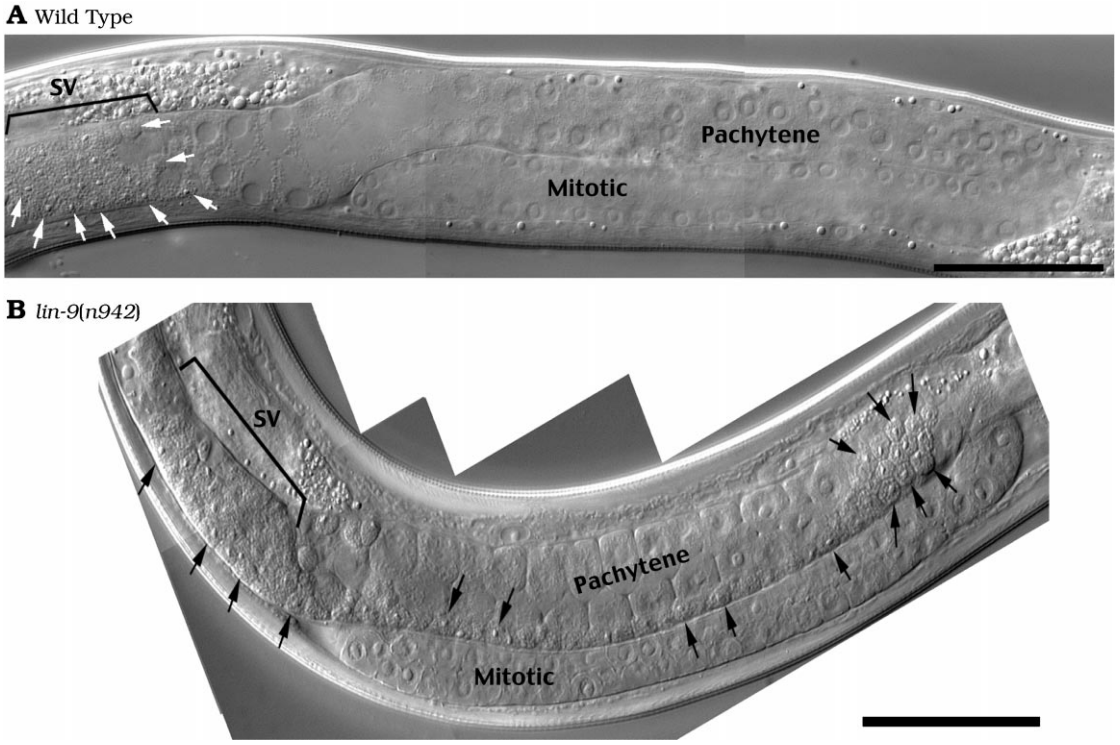


Fig. 2. Nomarski optics photomicrographs of adult male gonads. (**A**) Wild type (**B**) *lin-9(n942)*. In the *lin-9* animal, spermatids have slipped distally, away from the seminal vesicle. SV, seminal vesicle; arrows, representative spermatids; mitotic and pachytene germline zones are indicated. Scale bars, 40 μm.

cesses as the class B *lin-15* mutations, it is not clear whether the class A synMuv pathway functions at all in spicule development. If so, the class A and B synMuvs functions are not redundant in spicule development, in contrast to the case of hermaphrodite vulval induction.

3.3. Positional cloning of lin-9

To clone *lin-9*, we utilized the *C. elegans* genetic and physical maps. *lin-9* was previously mapped between two markers, *sup-5* and *unc-32*, positioned on both the physical and genetic maps (Fig. 3). We used cosmids from this region in germline transformation experiments to rescue the synMuv phenotype of *lin-9*(*n112*) animals. We chose the temperature-sensitive (ts) strain *lin-9*(*n112*); *lin-15*(*n433*) (Ferguson and Horvitz, 1989) for the transformation experiments because injecting *lin-9*(*n112*); *lin-15*(*n433*) animals reared at the permissive temperature avoided the difficulties of injecting Muv animals, which frequently rupture and die during transformation experiments (Fig. 3).

Injection of cosmids ZK637, C13C1 or C07D12 rescued the *lin-9*(*n112*) synMuv phenotype (Fig. 3). We constructed restriction maps of ZK637, C13C1 and C07D12, and tested subclones from the region of overlap among these cosmids for rescuing activity (Fig. 4). As this work was in progress, the *C. elegans* Genome Project determined the sequence of the cosmid ZK637

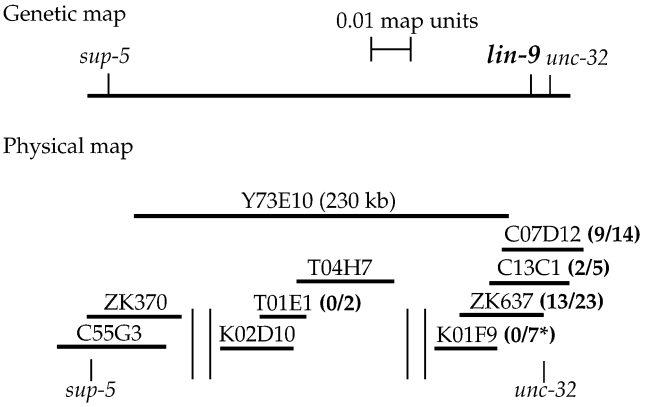


Fig. 3. Positional cloning of *lin-9*. The genetic and physical maps of the *lin-9* region are shown. The number of rescued transgenic lines and total number of transgenic lines tested are indicated beside each cosmid tested. The genetic map position of *lin-9* is from Ferguson and Horvitz (1989), and the physical map positions of *sup-5* and *unc-32* are from J. Sulston (personal communication) and Sulston et al. (1992). We used the temperature-sensitive strain MT990 *lin-9*(*n112*); *lin-15*(*n433*) to test for rescue of the *lin-9*(*n112*) synMuv phenotype. The *lin-15*(*n433*) allele is a weak synMuv class A mutation that in combination with *lin-9*(*n112*) causes a synMuv phenotype. ‖, regions of the physical map not covered by cosmids; *, transgenic lines after injection of K01F9 were difficult to obtain, suggesting that a 'poison' sequence is located in this cosmid. The failure of K01F9 to rescue the *lin-9* phenotype is therefore not a reliable indicator of whether *lin-9* is located on this cosmid. Y37E10 is a YAC clone, and the remaining clones shown on the physical map are cosmid clones. Physical map data are from the *C. elegans* Genome Project (Sulston et al., 1992).

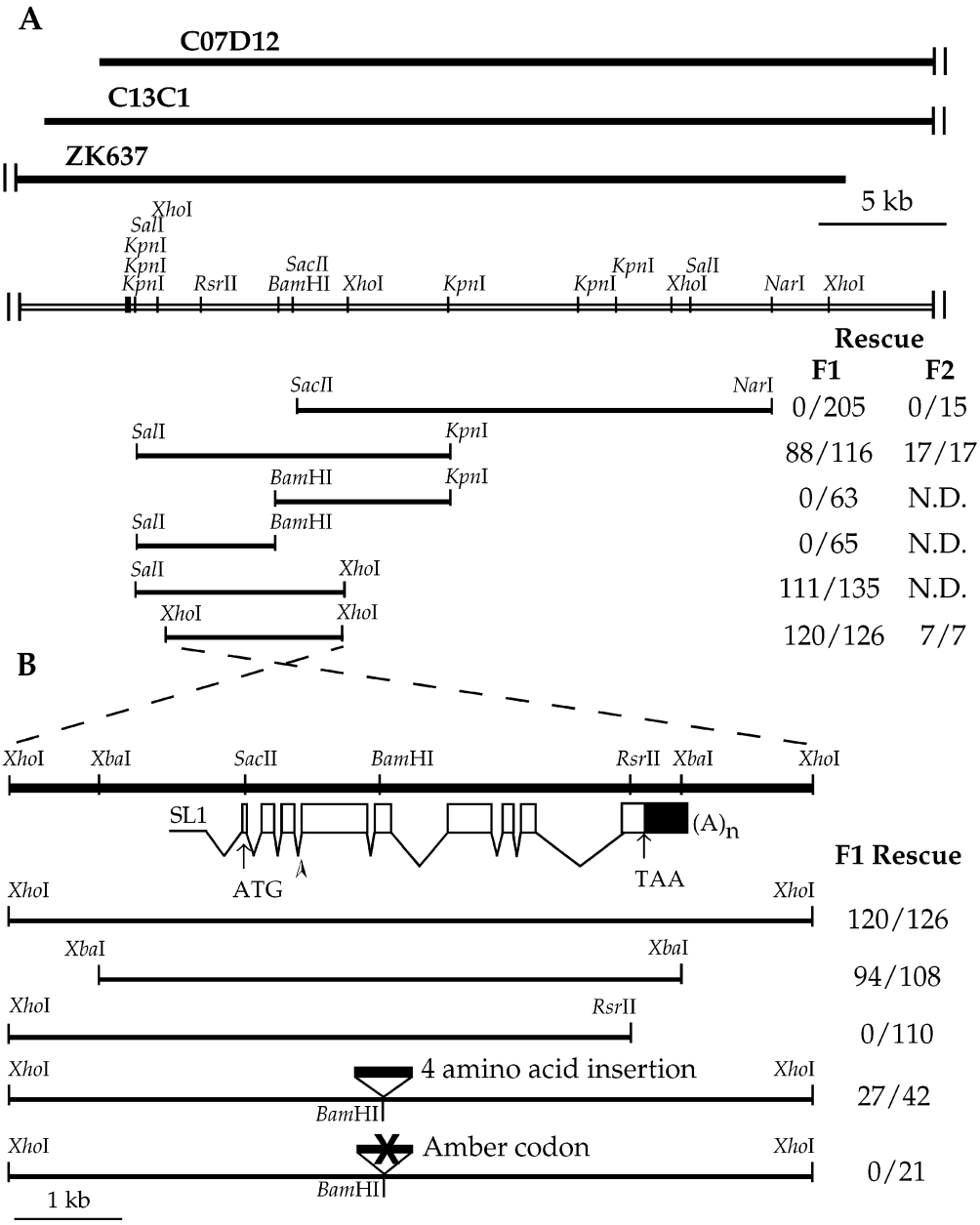


Fig. 4. Identification of genomic fragments with *lin-9* rescue activity. (**A**) Region of overlap among the cosmids ZK637, C13C1 and C07D12 as determined by restriction mapping and hybridization experiments (data not shown). We tested subclones of C07D12 for ability to rescue the synMuv phenotype of *lin-9* (*n112*); *lin-15*(*n433*) animals. The orientation of the cosmids in this figure corresponds to the standard orientation of the *C. elegans* physical and genetic maps. F1 rescue, fraction of transgenic F_1 animals that were rescued for the synMuv phenotype; F2 rescue, fraction of transgenic lines that were rescued for the synMuv phenotype; N.D., not determined. (**B**) Enlargement of the 7.0 kb *Xho*I–*Xho*I rescuing genomic fragment. Also shown is the exon/intron structure of the *lin-9* mRNA as deduced by comparing *lin-9* cDNA sequences (Fig. 6) with the genomic sequence obtained by the *C. elegans* Genome Project (Sulston et al., 1992). Exons are shown as boxes. The white areas of the boxes indicate the continuous open reading frame, while the black area indicates the 3′ untranslated region (UTR). SL1, the SL1 leader *trans*-spliced to the first exon of *lin-9*; ➢, location of alternative splice sites. ATG and TAA indicate the start and stop codons of the *lin-9* ORF, respectively. Other abbreviations and genotype of injected animals as in (**A**).

(Sulston et al., 1992), and we used this genomic sequence throughout our remaining analysis. Our germline transformation results defined a 5.1 kb *Xba*I–*Xba*I fragment that rescued the synMuv phenotype of *lin-9*(*n112*); *lin-15*(*n433*) animals (Fig. 4b).

3.4. *lin-9 probes hybridize to a 2.4 kb RNA on Northern blots*

To identify transcripts from the *lin-9* region, we used a 7.0 kb *Xho*I–*Xho*I fragment (Fig. 4A and B), which

contained the 5.1 kb *Xba*I–*Xba*I minimal rescuing fragment, to probe several Northern blots from mixed-stage wild-type animals. The 7.0 kb fragment strongly hybridized to RNA of approximately 2.4 kb on all blots (Fig. 5).

3.5. lin-9 *transcripts are alternatively spliced and polyadenylated*

We used the 7.0 kb *Xho*I–*Xho*I fragment (Fig. 4A and B) to probe cDNA libraries obtained from S. Kim and R. Barstead (Kim and Horvitz, 1990; R. Barstead, personal communication). We isolated and determined the 5′ and 3′ end sequences of a total of 31 hybridizing phage, of which 21 had independent inserts (see Section 2.4). Eight of the 21 independent inserts contained sequences that were not located within 20 kb of *lin-9* and thus appeared to consist of multiple cDNA fragments.

Two alternative polyadenylation (polyA) sites appear

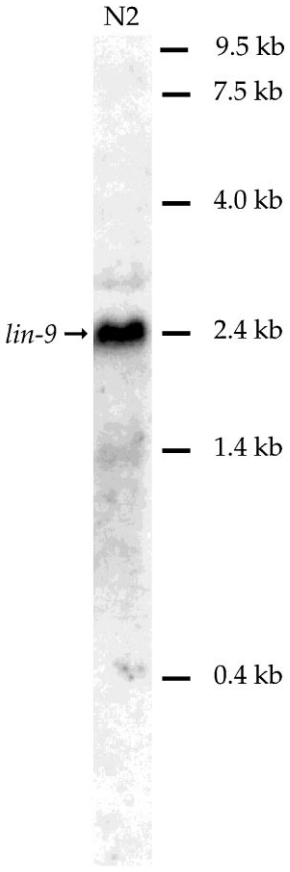


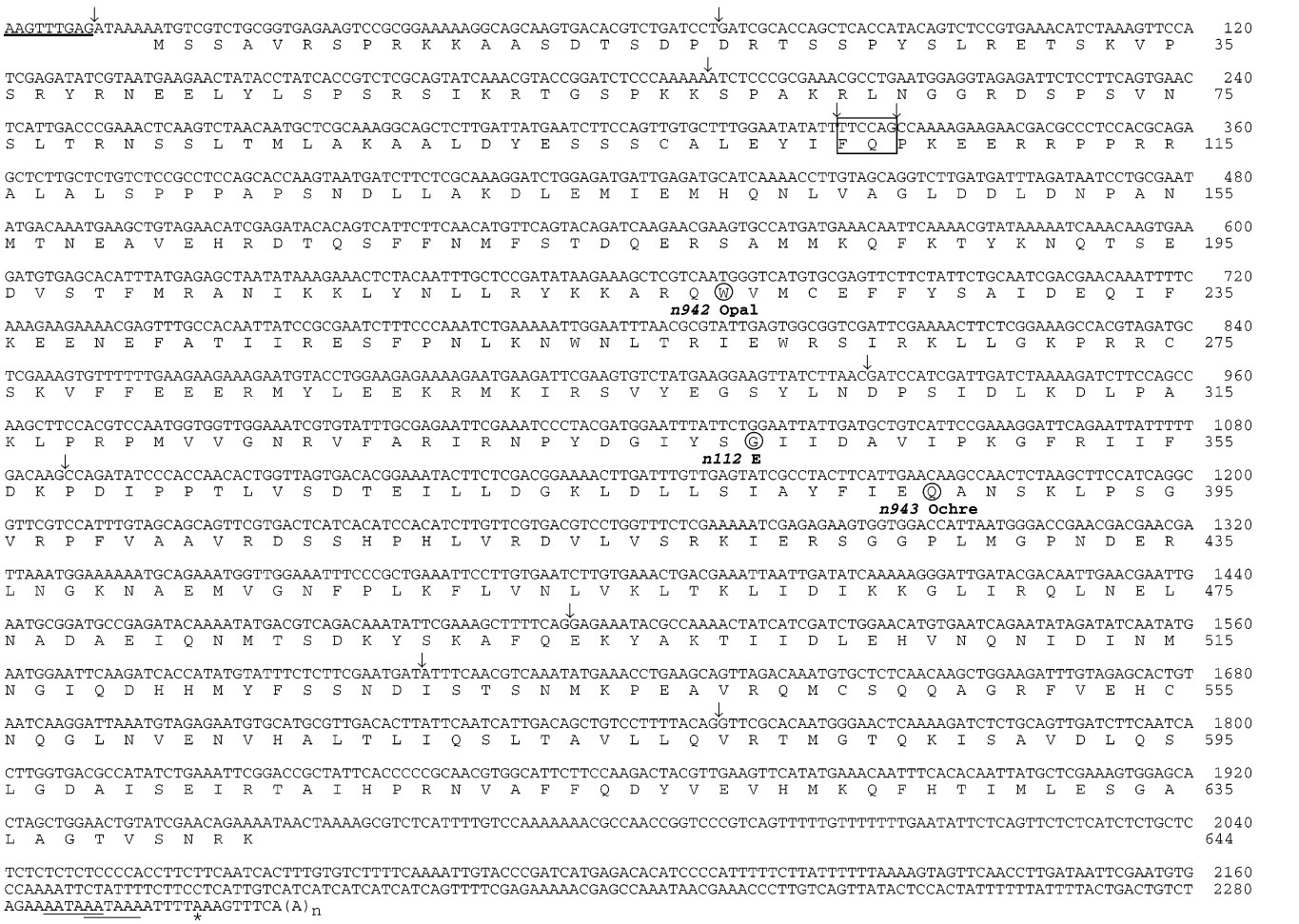
Fig. 5. Northern blot probed with the 7.0 kb *Xho*I–*Xho*I *lin-9* genomic rescue fragment showing one strongly hybridizing band of approximately 2.4 kb. The weakly hybridizing bands of other sizes were not consistently seen on other blots and were probably artifacts (data not shown). The positions and sizes of the RNA size markers that were run in an additional lane are indicated at the right-hand side of the blot. The 2.4 kb size of the strongly hybridizing bands was calculated from the positions of the size markers.

to be used to generate *lin-9* transcripts: of 14 independent inserts having *lin-9* region sequences at their 3′ ends, 10 inserts ended in TAAAATTTTAAAGTTTC$(A)_n$, while four were slightly shorter and ended in TAAAA TTTT$(A)_n$ (Fig. 6A). These polyA addition sites are separated by 5–8 bp (the exact cleavage/addition site is ambiguous) and are preceded by the consensus cleavage and polyA signal sequence AATAAA (Fig. 6A). We do not know whether there are any functional differences between these two distinct transcripts.

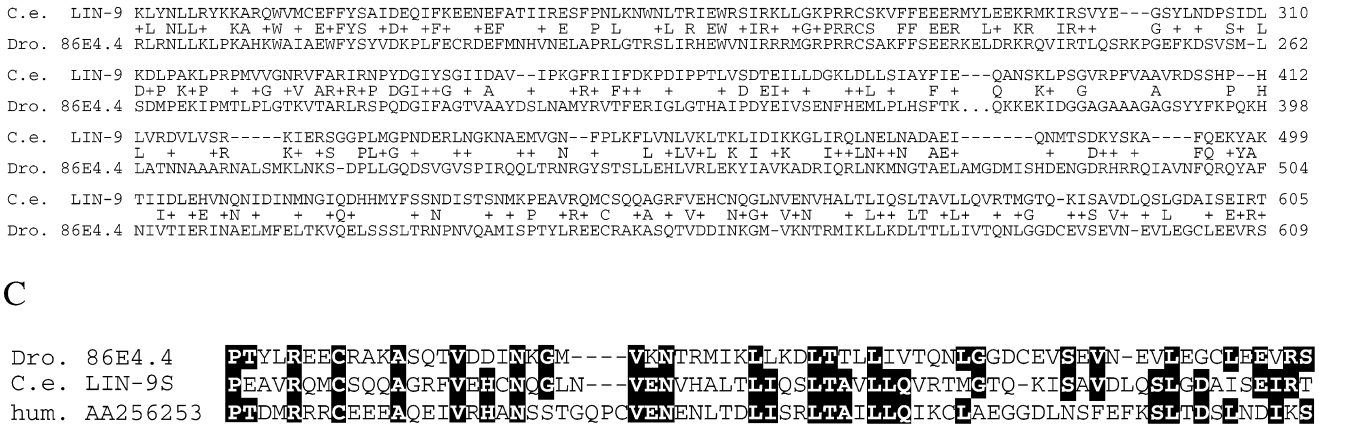
We determined the complete sequence of the longest cDNA from the Barstead library and of one strand of the longest cDNA from the Kim library (Fig. 6A). Comparison of these sequences to each other and to the genomic sequence determined by the *C. elegans* sequencing project revealed the features described below:

1. Two alternative splice-acceptor sites separated by 6 bp are used at the 5′ end of the fourth exon of *lin-9*. Three of the cDNAs that we isolated, including the longest Barstead cDNA and one cDNA isolated by Y. Kohara (GenBankID: D34812) had a 6 bp insert relative to the longest Kim cDNA and two cDNAs isolated by Y. Kohara (GenBank IDs: D35704 and D36526) (Fig. 6A and data not shown). The 6 bp insert did not interrupt the long open reading frames (ORFs) found in both cDNAs. Both splice-acceptor sites matched the minimal *C. elegans* splice acceptor sequence (Fields, 1990). The functional differences, if any, between these splice variants are not known.
2. There appeared to be no correlation between which alternative fourth exon splice-acceptor site was used and which polyA site was used. We found three of the four possible splice site/polyA site combinations among the four cDNAs that contained both the fourth exon and polyA tails.
3. The 5′ ends of the two longest cDNAs isolated from the Barstead library contained 9 and 7 bp of sequence (Fig. 6A and data not shown) that matched the *C. elegans* SL1 spliced leader sequence (reviewed by Blumenthal, 1995). The divergence between the cDNA sequences and the genomic sequence occurred at a site that matched the *C. elegans* splice-acceptor consensus sequence (Fields, 1990). Thus, the *lin-9* transcripts, like many other C. *elegans* transcripts, are SL1 *trans*-spliced (reviewed by Blumenthal, 1995). This result further indicated that the two longest cDNAs that we isolated from the Barstead library were full-length transcripts.
4. The calculated 2306 to 2320 bp sizes (including alternative splice and polyA sites, but not polyA tails) of the full-length *lin-9* cDNAs were consistent with the 2.4 kb size of the strongly hybridizing RNA seen on Northern blots.
5. The sequence of the longest Kim cDNA was in perfect agreement with the genomic sequence deter-

A



B



C

Fig. 6. *lin-9* cDNA sequences and protein alignments. (**A**) Sequence of the longest cDNA isolated from the Barstead cDNA library. The box indicates an alternatively spliced sequence absent in three of seven cDNAs. The amino acids altered by the three *lin-9* mutations are circled. The exact base changes are as follows: *n112*, GGA-to-GAA; *n942*, TGG-to-TGA; *n943*, CAA-to-TAA. ↓, intron/exon boundary; dark underline, SL1 *trans*-spliced sequence; light underline, consensus polyA signal sequences *, alternative polyadenylation site used by four of 14 cDNAs. The sequence of the cDNA that encodes the LIN-9L isoform and the deduced sequence of the cDNA that encodes the LIN-9S isoform have been submitted to GenBank and given the Accession Nos. AF269693 and AF269694, respectively. (**B**) Alignment between the predicted *Drosophila* protein 86E4.4 (GenBank Accession No. 2749752) amino acid 156–609 and LIN-9S amino acid 207–605. Dro., *Drosophila*; C. e., *C. elegans*; '-', gap introduced by alignment program; '…', 27 amino acids of 86E4.4 not shown; '+', conserved residue. (**C**) Alignment between the predicted *Drosophila* protein 86E4.4 amino acid 542–649, LIN-9S amino acid 535–603 and the translation of a human EST (GenBank Accession No. AA256253) amino acid 33–103. Black boxes indicate residues identical between at least two of the three sequences. hum., human.

mined by the *C. elegans* Genome Project. However, the longest Barstead cDNA had a T at position 2010, while the Kim cDNA and the Genome Project's genomic sequences had a C. As this discrepancy is in the 3′ untranslated region (UTR), we have not investigated this difference any further.

3.6. lin-9 *encodes the first members of a new gene family*

Using the first ATGs of the cDNA sequences as start codons, conceptual translation of the longest cDNAs from the Barstead and Kim libraries produced proteins of 644 and 642 amino acids, respectively (Fig. 6A). We designate the putative proteins from the longest Barstead and longest Kim cDNAs to be the LIN-9L and LIN-9S proteins, respectively.

The LIN-9L and LIN-9S amino acid sequences do not match any diagnostic motifs in the PROSITE database (Appel et al., 1994) provided by The Geneva University Hospital and University of Geneva (Geneva, Switzerland) and are not predicted to contain either transmembrane or coiled-coil regions based upon the criteria of Kyte and Doolittle (1982) and the PAIRCOIL program (Berger et al., 1995). However, gapped BLAST searches (Altschul et al., 1997) of National Center for Biotechnology Information non-redundant protein databases identified the predicted *Drosophila* protein 86E4.4 as having a significant similarity to both LIN-9 proteins over 60% their lengths ($P = 10^{-39}$) (Fig. 6B). The LIN-9 proteins also have extensive similarities to the *Drosophila* Aly protein (White-Cooper et al., 2000). In addition, searches of the nucleotide databases identified *Drosophila*, *Arabidopsis*, zebra fish, human and mouse sequences having a highly significant similarity to several regions of the LIN-9 proteins (P values $= 10^{-39}$, 10^{-6}, 10^{-19}, 3×10^{-7}, and 10^{-5}, respectively). An alignment of the region of shared similarity among the predicted amino acid sequences of LIN-9, 86E4.4 and the highest scoring human EST is shown in Fig. 6C. Note that in this region, 18% of the residues are identical among all three sequences, and a total of 44% of the residues are conserved between two of the three sequences. Since there are *lin-9* related genes in *C. elegans*, *Drosophila*, plants and vertebrates, *lin-9* defines a new gene family that may act in Rb-related pathways in different organisms.

3.7. *Nature of* lin-9 *point mutations*

To demonstrate that we had correctly identified the *lin-9* ORFs and to investigate the molecular nature of *lin-9* mutations, we determined the genomic sequences of the three *lin-9* mutant alleles *n112*, *n942*, and *n943*. For all three, we found single-base pair changes that resulted in alterations of the *lin-9* coding sequence (Fig. 6A). The weak allele, *n112*, causes a Gly-to-Glu substitution at positions 341 of LIN-9L and 339 of LIN-9S . The two strong *lin-9* mutations, *lin-9(n942)* and *n943*, which cause a non-synthetic Sterile defect in addition to a synMuv phenotype (Ferguson and Horvitz, 1989), are both stop codons that would result in truncated proteins less than half the lengths of the normal proteins (Fig. 6A). That we had correctly identified the *lin-9* ORFs is further evidenced by the observation that the rescuing activity of genomic fragments was blocked by insertion of an amber stop codon, but not by four in-frame amino acids, into the *lin-9* ORFs (Fig. 4B).

3.8. *Sequences required for rescue of* lin-9 *Sterile and synMuv phenotypes*

We tested the ability of the 5.1 kb *Xba*I–*Xba*I *lin-9*(*n112*) rescuing genomic fragment to rescue the Sterile and synMuv defects caused by the *lin-9(n943)* mutation (Section 2.2). This 5.1 kb fragment included 1.5 kb of DNA 5′ of the *trans*-spliced leader site and 333 bp of DNA 3′ of the *lin-9* stop codon (all but 23 bp of the 3′ UTR). The 5.1 kb fragment, which strongly rescued the *lin-9*(*n112*) synMuv phenotype (Fig. 4B), strongly rescued the *lin-9(n943)* sterile defect in transgenic lines (data not shown). However, only 20–40% of *lin-9(n943)* animals were rescued for the synMuv phenotype (data not shown). It is possible that the 5.1 kb fragment provides incomplete rescue because it lacks important regulatory sequences or because of mitotic loss of extrachromosomal arrays.

Since a genomic fragment missing the last 23 bp of the *lin-9* 3′ UTR could at least partially rescue the *lin-9* Sterile and synMuv defects, we tested the rescuing ability of a genomic fragment missing the entire 3′ UTR and 122 bp of C-terminal coding sequence (Fig. 4B). A 5.8 kb *Rsr*II–*Xho*I fragment, which contained an additional 780 bp 5′ of the *lin-9* SL1 *trans*-splice leader site compared with the 5.1 kb *Xba*I–*Xba*I fragment, did not rescue either the synMuv phenotype of *lin-9*(*n112*) animals or the synMuv and Sterile defects of *lin-9(n943)* animals (Fig. 4B and data not shown). Thus, some region of the 451 bp *Xba*I–*Rsr*II segment at the 3′ end of *lin-9* is required for the correct expression or function of the LIN-9 proteins in fertility and vulval development.

3.9. A lin-9 *rescuing fragment did not rescue* lin-8 *or* lin-36 *mutant defects*

To determine whether expression of *lin-9* from extrachromosomal arrays could rescue defects caused by loss-of-function mutations in other synMuv genes, we performed germline transformation experiments using the 7.0 kb *Xho*I–*Xho*I fragment and the temperature-sensitive synMuv strain MT3032 *lin-8*(*n111*); *lin-36*(*n747*) (Ferguson and Horvitz, 1989). The *Xho*I–*Xho*I fragment, which rescued the Muv phenotype of 95% of *lin-*

9(*n112*); *lin-15*(*n433*) F_1 animals (Fig. 4B), failed to rescue *lin-8*; *lin-36* animals (data not shown). Our results are similar to the observations of Clark et al. (1994) and Huang et al. (1994), who found that transgenic lines containing extrachromosomal arrays having *lin-15* class A activity did not rescue defects caused by mutations in the synMuv class A gene *lin-8* and the synMuv class B genes *lin-9*, *lin-36* or *lin-15*(B). Arrays having *lin-15* class B activity did not rescue defects caused by mutations in the class A genes *lin-8* or *lin-15*(A) and the class B genes *lin-9* or *lin-36*. It remains to be determined whether these results were a result of insufficient overexpression of the genes.

In a wild-type background, *lin-9* extrachromosomal arrays did not appear to cause any gain-of-function abnormalities, such as inappropriate vulval cell-fate determination (data not shown). By contrast, extrachromosomal arrays of some of the RTK/Ras signaling pathway genes, such as *let-60* Ras and *lin-3*, do cause gain-of-function phenotypes (Han and Sternberg, 1990; Hill and Sternberg, 1992). Further work will be required to determine whether these results reflect a limited ability of *lin-9* or the synMuv pathways to antagonize RTK/Ras pathway signaling or whether the expression levels in these experiments were not high enough to cause gain-of-function phenotypes.

4. Conclusions

The *C. elegans* synMuv genes provide an opportunity for analyzing intercellular signaling, genetic redundancy, and the antagonism of RTK/Ras signal transduction pathways and for identifying new genes that interact with or regulate Rb pathways. We characterized the synMuv gene *lin-9*. We found that *lin-9* is non-redundantly required for hermaphrodite sheath-cell development and the development of the male spicule, rays and gonad. *lin-9* encodes 642 and 644 amino acid proteins that differ by two amino acids and that have significant similarity to predicted proteins in other species. Thus, *lin-9* defines a new gene family. In *C. elegans*, *lin-9* functions in an Rb-related pathway that regulates gene expression via a NURD chromatin remodeling and histone deacetylase complex (Lu and Horvitz, 1998; Hsieh et al., 1999; Solari and Ahringer, 2000), which suggests that in other organisms, *lin-9* homologues will function in similar pathways. This prediction is supported by the recent observations of White-Cooper et al. (2000), who showed that the *Drosophila lin-9*-homologue *aly* encodes a chromatin-associated protein required for normal primary spermatocyte chromatin structure and terminal differentiation during male gametogenesis (Lin et al., 1996; White-Cooper et al., 1998). White-Cooper et al. (2000) also found that the wild-type Aly protein can have both nuclear and cytoplasmic subcellular localizations and that an amino acid substitution that caused the Aly protein to be cytoplasmically localized also strongly reduced *aly* activity. These observations suggest that *lin-9* activity could be regulated through subcellular localization and that *lin-9* could transduce the synMuv class B signal from the cytoplasm to the nucleus. If so, *lin-9* subcellular localization in synMuv mutants could be used as an assay for determining the order of action of the class B synMuv genes. In any case, our molecular characterization of *lin-9* provides important tools for a further analysis of the role of *lin-9* and its homologues in regulating signal transduction, gene expression and cellular differentiation.

Acknowledgement

We thank Jeffrey Thomas, Shai Shaham, Michael Koelle and Sander van den Heuvel for critical comments concerning this manuscript. We also thank current and past members of the Horvitz and Meyer laboratories for their helpful discussions and assistance including Jeffrey Thomas, Craig Ceol, Mark Metzstein, Michel Labouesse, Laird Bloom, Gian Garriga, and Leslie Lobel. We thank David Greenstein for the anti-CEH-18 antisera, and Helen White-Cooper and Minx Fuller for communicating unpublished data. Many appropriate references have not been cited because of limitations on the number of references allowed. This work was supported by United States Public Health Service grants GM24663 to H.R.H. and GM49785 to E.J.L. G.J.B. was a Predoctoral Fellow of the Howard Hughes Medical Institute. H.R.H. is an Investigator of the Howard Hughes Medical Institute.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zheng, Z.Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Appel, R.D., Bairoch, A., Hochstrasser, D.F., 1994. A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. Trends Biochem. Sci. 19, 258–260.

Berger, B., Wilson, D.B., Wolf, E., Touchev, T., Milla, M., Kim, P.S., 1995. Predicting coiled coils by use of pairwise residue correlations. Proc. Natl. Acad. Sci. USA 92, 8259–8263.

Blumenthal, T., 1995. *Trans*-splicing in polycistronic transcription in *Caenorhabditis elegans*. Trends Genet. 11, 132–136.

Chamberlin, H.M., Sternberg, P.W., 1994. The *lin-3*/*let-23* pathway mediates inductive signalling during male spicule development in *Caenorhabditis elegans*. Development 120, 2713–2721.

Church, D.L., Guan, K.L., Lambie, E.J., 1995. Three genes of the MAP kinase cascade, *mek-2*, *mpk-1*/*sur-1* and *let-60 ras*, are required for meiotic cell cycle progression in *Caenorhabditis elegans*. Development 121, 2525–2535.

Clark, S.G., Lu, X., Horvitz, H.R., 1994. The *C. elegans* locus *lin-15*,

a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. Genetics 137, 987–997.

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

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

Fields, C., 1990. Information content of *Caenorhabditis elegans* splice site sequences varies with intron length. Nucleic Acids Res. 18, 1509–1512.

Francis, R., Barton, M.K., Kimble, J., Schedl, T., 1995. *gld-1*, a tumor suppressor gene required for oocyte development in *Caenorhabditis elegans*. Genetics 139, 579–606.

Greenstein, D., Hird, S., Plasterk, R.H.A., Andachi, Y., Kohara, Y., Wang, B., Finney, M., Ruvkun, G., 1994. Targeted mutations in the *Caenorhabditis elegans* POU homeobox gene *ceh-18* cause defects in oocyte cell cycle arrest, gonad migration and epidermal differentiation. Genes Dev. 8, 1935–1948.

Han, M., Sternberg, P.W., 1990. *let-60*, a gene that specifies cell fates during *C. elegans* vulval induction, encodes a ras protein. Cell 63, 921–931.

Herman, R.K., Hedgecock, E.M., 1990. Limitation of the size of the vulval primordium of *Caenorhabditis elegans* by *lin-15* expression in surrounding hypodermis. Nature 348, 169–171.

Hill, R.J., Sternberg, P.W., 1992. The gene *lin-3* encodes an inductive signal for vulval development in *C. elegans*. Nature 358, 470–476.

Hodgkin, J., Edgley, M., Riddle, D.L., Albertson, D.G., 1988. . In: Wood, W.B. and the community of *C. elegans* researchers (Ed.), The Nematode *Caernorhabditis elegans*. Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 491–584.

Hsieh, J., Liu, J., Kostas, S.A., Chang, C., Sternberg, P.W., Fire, A., 1999. The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. Genes Dev. 13, 2958–2970.

Huang, L.S., Tzou, P., Sternberg, P.W., 1994. The *lin-15* locus encodes two negative regulators of *C. elegans* vulval development. Mol. Biol. Cell 5, 395–411.

Iwasaki, K., McCarter, J., Francis, R., Schedl, T., 1996. *emo-1*, a *Caenorhabditis elegans* Sec61p homologue, is required for oocyte development and ovulation. J. Cell Biol. 134, 699–714.

Kim, S.K., Horvitz, H.R., 1990. The *Caenorhabditis elegans* gene *lin-10* is broadly expressed while required specifically for the determination of vulval cell fates. Genes Dev. 4, 357–371.

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

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Lin, T.Y., Viswanathan, S., Wood, C., Wilson, P.G., Wolf, N., Fuller, M.T., 1996. Coordinate developmental control of the meiotic cell cycle and spermatid differentiation in *Drosophila* males. Development 122, 1331–1341.

Lu, X., Horvitz, H.R., 1998. *lin-35* and *lin-53*, two genes that antagonize a *C. elegans* Ras pathway, encode proteins similar to Rb and its binding protein RbAp48. Cell 95, 981–991.

McCarter, J., Bartlett, B., Dang, T., Schedl, T., 1997. Soma–germ cell interactions in *Caenorhabditis elegans*: Multiple events of hermaphrodite germline development require the somatic sheath and spermathecal lineages. Dev. Biol. 181, 121–143.

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

Myers, C.D., Goh, P.Y., Allen, T.S., Bucher, E.A., Bogaert, T., 1996. Developmental genetic analysis of troponin T mutations in striated and nonstriated muscle cells of *Caenorhabditis elegans*. J. Cell Biol. 132, 1061–1077.

Solari, F., Ahringer, J., 2000. NURD-complex genes antagonise Ras-induced vulval development in *Caenorhabditis elegans*. Curr. Biol. 10, 223–226.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., et al., 1992. The *C. elegans* Genome sequencing project: a beginning. Nature 356, 37–41.

Thomas, J.H., Horvitz, H.R., 1999. The *C. elegans* gene *lin-36* acts cell autonomously in the *lin-35* Rb pathway. Development 126, 3449–3459.

White-Cooper, H., Schafer, M.A., Alphey, L.S., Fuller, M.T., 1998. Transcriptional and post-transcriptional control mechanisms coordinate the onset of spermatid differentiation with meiosis I in *Drosophila*. Development 125, 125–134.

White-Cooper, H., Leroy, D., MacQueen, A., Fuller, M.T., 2000. Nuclear localization of a conserved chromatin associated protein regulates transcription of meiotic cell cycle and terminal differentiation genes. Submitted for publication.