Distinct Requirements for Somatic and Germline Expression of a Generally Expressed *Caenorhabditis elegans* Gene

William G. Kelly, SiQun Xu, Mary K. Montgomery and Andrew Fire

Carnegie Institution of Washington, Department of Embryology, Baltimore, Maryland 21210

Manuscript received December 18, 1996
Accepted for publication February 14, 1997

ABSTRACT

In screening for embryonic-lethal mutations in *Caenorhabditis elegans*, we defined an essential gene (*let-858*) that encodes a nuclear protein rich in acidic and basic residues. We have named this product nucampholin. Closely homologous sequences in yeast, plants, and mammals demonstrate strong evolutionary conservation in eukaryotes. Nucampholin resides in all nuclei of *C. elegans* and is essential in early development and in differentiating tissue. Antisense-mediated depletion of LET-858 activity in early embryos causes a lethal phenotype similar to characterized treatments blocking embryonic gene expression. Using transgene-rescue, we demonstrated the additional requirement for *let-858* in the larval germline. The broad requirements allowed investigation of soma-germline differences in gene expression. When introduced into standard transgene arrays, *let-858* (like many other *C. elegans* genes) functions well in soma but poorly in germline. We observed incremental silencing of simple *let-858* arrays in the first few generations following transformation and hypothesized that silencing might reflect recognition of arrays as repetitive or heterochromatin-like. To give the transgene a more physiological context, we included an excess of random genomic fragments with the injected DNA. The resulting transgenes show robust expression in both germline and soma. Our results suggest the possibility of concerted mechanisms for silencing unwanted germline expression of repetitive sequences.

GERMLINE totipotency is critical for reproductive function. To maintain totipotency, there must be mechanisms to ensure that the germline can be insulated from the somatic differentiation pathways that generate a diversity of tissues. A striking example of the early insulation of germline from somatic processes is seen in the nematode *C. elegans*. In the earliest divisions, a series of germline blastomeres (P lineage) each divide to give two daughter cells; one of which will give rise to somatic tissue lineages, while the other will maintain its germline characteristics (SULSTON *et al.* 1983). The generation of both germline and somatic fates from the division of a germline blastomere parent results in part from the asymmetric distribution of cytoplasmic components to the daughter cells (for review, see GUO and KEMPHUES 1996b). Although the germline cell and its newly formed somatic sisters appear only marginally different under the microscope, at the level of gene expression there is a striking difference. At these early stages of development, it appears that the germ cell nucleus fails to express many (and perhaps all) mRNA encoding genes (SEYDOUX and FIRE 1994). At the same time transcription of many different components is being activated in the somatic sisters and cousins of the germ lineage.

The germline cannot remain transcriptionally quiescent. After the animal hatches, there is a dramatic increase in both the number and the volume of germ cells in the animal, from two small cells to a population of several thousand nuclei, which comprises the adult germline. This increase in both cell number and mass presumably involves new expression of a wide variety of genes, including those responsible for housekeeping functions, cell proliferation, responses to environmental cues, and a multitude of factors destined for incorporation into developing oocytes (AUSTIN and KIMBLE 1987; SCHEDL 1996). With the onset of extensive new gene expression in the germline, there remains a need to prevent somatic (and perhaps irreversible) differentiation from occurring in the germ lineage. Thus it might be expected that the germline of *C. elegans* uses one or a few efficient mechanisms to distinguish its activity in gene expression from somatic neighbors.

One mechanism that could conceivably serve this function would be chromatin elimination. In certain nematode species, a massive rearrangement of chromosomes and elimination of certain interspersed germline sequences occurs at each successive germline-soma division, so that genes expressed in the soma do so in a strikingly different genomic "neighborhood" from the same genes in the germline (MULLER *et al.* 1996). This massive elimination of DNA in the soma is not seen in *C. elegans*, but it is quite possible that a change in the nature of chromatin between soma and germline might be as critical in *C. elegans* as in its parasitic relatives.

Transgenic studies in *C. elegans* suggest that such

Corresponding author: William G. Kelly, Carnegie Institution of Washington, Department of Embryology, 115 W. University Parkway, Baltimore, MD 21210. E-mail: fire@mail1.ciwemb.edu

soma/germline contextual differences may indeed exist. Transgenes are heritably transmitted as extrachromosomal arrays, which are the linear products of recombination and ligation of DNA injected into the syncytial gonad of the adult (STINCHCOMB *et al.* 1985; MELLO *et al.* 1991). A wide variety of both *lac-Z* and *Aequorea victoria* green fluorescent protein (GFP) reporter gene constructs have been introduced into *C. elegans* in this manner, with efficient expression of the reporter in most somatic tissues (FIRE *et al.* 1990; HOPE 1991; CHALFIE *et al.* 1994). An exception to this success, in our laboratory and others, has been the inability to observe any reporter transgene expression in the germline. This failure, coupled with notably poor transgenic rescue of maternal effect mutations, has led us to the conclusion that the transgenes are "viewed" differently by the germline than by the soma and suggests another mechanism whereby gene expression may be differentially repressed in the postembryonic germline. The most tractable feature of this mechanism is the selective silencing of transgenes in the germline. We have therefore used a generally expressed *C. elegans* gene, *let-858*, to begin to address the problem of germline transgenic expression and to try to understand the different rules that govern germline and somatic expression of genes.

MATERIALS AND METHODS

EMS lethal screen: To isolate lethal mutations in the mnDf90/cey-1 region of LG II, L4 stage *unc-4* (e120) animals were mutagenized with ethane methyl sulfonate (EMS) as described (BRENNER 1974) and mated with *unc-4(e120)/mnC1(dpy-10 unc-52)* males. Phenotypically wild-type F_1 progeny were cloned and screened for the failure to segregate fertile adult *unc-4* homozygotes. Sterile and lethal mutations linked to unc-4 were kept for further study if they were complemented by mnDf-77 but not by *mnDf-87* and *mnDf-90*. Thirty-four alleles belonging to seven separate complementation groups, designated *let-852* through *let-858*, were recovered from this screen.

Cloning and sequencing of *let-858*: Mutations in one of the genes recovered from the EMS screen, *let-858*, were rescued by injection with mixtures containing the cosmid F33A8 using described procedures (MELLO and FIRE 1995). The rescuing portion of the cosmid was further delineated by a combination of restriction digests and *Exo*III digestions of a parent 11-kb *Xba*I fragment, yielding a 6.8-kb rescue-competent subclone, which was sequenced. To determine the genomic arrangement of *let-858*, a *C. elegans* lambda-Zap (Stratagene) cDNA library was probed at high stringency with the *Xba*I fragment, and positive clones were isolated and sequenced. Additional sequences of cDNAs corresponding to *let-858* were obtained through searches of the expressed sequence-tagged databases. The 5′ end of the *let-858* transcript was identified by PCR amplification of a mixed-stage cDNA pool using an internal gene-specific antisense primer paired with primers corresponding to the *C. elegans* SL1 or SL2 trans-spliced leader sequences. A single product was generated in reactions containing the SL1 oligonucleotide. The structure of this product indicated that the *let-858* mRNA contains the SL1 splice leader, followed by a 5 bp 5′ UTR preceding the initiator codon.

Anti-LET-858 antibodies: A cDNA fragment encoding amino acids 137–333 of LET-858 was inserted in-frame into the *trp-E* bacterial expression vector, pATH1 (KOEHNER *et al.* 1991). Following IPTG induction, the resulting TRP-E::LET-858 fusion protein was purified for antibody preparation. Antigen injections into mice, and resulting antisera collections, were performed by Josman Laboratories (Napa, CA). The antisera was tested for specificity by Western blot analysis of bacterial and *C. elegans* lysates. *In situ* localization of the endogenous LET-858 protein was subsequently carried out on fixed *C. elegans* adults and embryos as described (MILLER and SHAKES 1995).

***In situ* hybridization to localize RNA transcripts:** Whole mount *in situ* hybridizations on wild-type (N2) *C. elegans* embryos were performed as described (SEYDOUX and FIRE 1994, 1995). Digoxigenin (DIG)-labeled single-stranded DNA probes were synthesized by multiple cycles of primer extension in the presence of DIG-dUTP using either a full-length *let-858* cDNA clone (pBK 4.1) or a partial *let-858* genomic clone (pMM215) as template essentially as described (PATEL and GOODMAN 1992). Antisense probes generated from either plasmid gave similar results; no signal was detected with sense probes (data not shown).

***let-858*::*gfp* fusions and deletion constructs:** The coding region for the *Aequorea victoria* green fluorescent protein (GFP) was inserted into a unique *Age*I site in a *let-858* genomic construct, pBK42, resulting in an internal, in-frame fusion of the fourth exon of *let-858* with GFP coding sequence. This construct became the parent vector for all deletions and insertions used to test for *cis*-acting elements driving germline expression of LET-858 (see text). In general, all 5′ deletions were generated by the following strategy: a *Bam*HI/*Age*I fragment from pBK48 (encompassing the genomic 5′ flanking DNA of *let-858* through the *Age*I *GFP* fusion site) was replaced by PCR-generated fragments encoding increasing deletions into the 5′ flanking region of *let-858*. These fragments were made using *Bam*HI site-containing sense primers corresponding to sequences increasingly 3′ within the *let-858* 5′ flanking region, coupled with antisense primers corresponding to *gfp* sequences through the *Age*I site.

All of the internal and 3′ deletions tested were constructed through the excision and ligation of internal restriction fragments, with the exception of an intron 5-intron 6 deletion; this construct was generated through the replacement of the genomic fragment with a corresponding fragment from a cDNA clone.

Transgenic analyses: *gfp* fusion reporter constructs were introduced using either of two methods. The first method involved the co-injection of circular reporter and transformation marker plasmids as previously described (MELLO and FIRE 1995). The resulting transformed lines generated from this "standard" transformation protocol carried the reporter and marker plasmids in heritable extrachromosomal arrays that proved incompetent for reporter expression in the germline (see text). To optimize for array complexity, another approach involving dilution with carrier DNA was used. Briefly, *gfp* reporter and transformation marker plasmid [pRF4; containing a dominant mutation in the collagen gene, *rol6(su1006)*] DNAs were linearized by restriction digestion at unique sites in the vector backbone. Where feasible, blunt end-generating enzymes were used; T4 polymerase was used to create blunt ends on digested plasmids if "sticky-end overhangs" were generated by the only unique backbone sites available. The linearized plasmids were then mixed with *Pvu*II-digested N2 (wild-type) strain genomic DNA that had been column-purified (Wizard Prep columns, Promega) and phenol-chloroform extracted. The mass ratio of N2 DNA to reporter or marker plasmid DNA in the injection mix was

routinely ~50–100:1, with the N2 DNA concentration at 50–100 μg/ml. The DNA injection mix was formulated in either dH2O or TE buffer (10 mM Tris, pH 7.5, 1 mM EDTA).

In all assays, transformed F_1 progeny of injected animals were cloned and checked for the transmission of heritable arrays, as indicated by the presence of marked transformants in their (F_2) progeny. Any clones segregating the arrays were then deemed to be independent lines for further study. For deletion and generational silencing studies, multiple animals (more than eight) from multiple independent lines (more than five) were analyzed. Relative fluorescence was determined by comparing the GFP fluorescence intensity of nuclei in test lines with that of nuclei in similar tissues in a "benchmark" line. The "benchmark" line used was a *let-858* (*cc534*) homozygous line rescued to fertility by co-injection of the pBK48.1 *let-858::gfp* reporter construct and N2 genomic DNA fragments.

Antisense RNA injections: Antisense RNA was generated *in vitro*, using a linearized *let-858* cDNA clone as template, according to instructions supplied by the RNA polymerase source (Promega). The transcription reaction was terminated on ice, and the products were purified by two successive phenol::chloroform extractions, followed by chloroform extraction and ethanol precipitation. Each antisense RNA was injected at a concentration of 0.5–1.0 mg/ml. The injected animals were allowed to recover overnight. Animals were transferred the following day to new plates and allowed to lay eggs for 4–6 hr before observation. The fraction of arrested offspring was determined from these plates by counting unhatched embryos the following day. The fraction of arrested embryos from antisense-injected animals typically exceeded 90%. Antisense RNA for other mRNAs, including *pie-1* and *mex-3*, were injected in parallel into N2 animals for comparison of arrest phenotype and to assess the efficiency of the injection protocol. Sense RNA controls were not attempted, given the known mimicry of antisense loss-of-function phenotypes by sense RNA injections in *C. elegans* (Fire *et al.* 1991; Guo and Kemphues 1995).

RESULTS

Essential genes in a defined region of chromosome *II*: We sought to identify genetic loci whose products might be required ubiquitously throughout development. Herman and colleagues have provided a number of genetic tools that make chromosome II an amenable region for isolation and genetic characterization of lethal mutations (Herman 1978; Sigurdson *et al.* 1984). From a chemical mutagenesis screen of a defined region of this chromosome, we obtained 34 alleles of seven essential genes (Table 1). For initial analysis, we chose one gene from this screen, *let-858*. All three *let-858* alleles show a uniform embryonic arrest phenotype in which 100% of the embryos fail to hatch.

Molecular structure of *let-858*: In initial rescue experiments, we found that cosmid F33A8 (Figure 1, A and B) was sufficient to rescue the embryonic lethal defect in *let-858* homozygotes derived from a heterozygous mother. Subsequent experiments using restriction fragments from F33A8 narrowed the rescuing activity to an 8.4-kb *Xba*I-*Apa*I fragment. This fragment was used to probe a *C. elegans* embryonic cDNA library, resulting in the isolation of four cDNAs, the longest of which defines a 2.8-kb open reading frame (Figure 1C). RT-PCR analysis showed the mRNA to be trans-spliced to the SL1 splice leader. Analysis of genomic sequence indicated that this open reading frame is interrupted by nine intron segments and spans the majority of the original 8.4-kb fragment (Figure 1C). Deletions at either end of this open reading frame prevent transgene rescue, demonstrating that this coding region indeed corresponds to *let-858*.

The LET-858 protein is predicted to be a 897 amino acid polypeptide that is highly hydrophilic, being richly scattered with both acidic and basic residues (42%

TABLE 1

Mutations in essential genes in a defined region of LGII

Complementation group	Alleles	Phenotype (arrest point)
let-852	*cc502*	Larval
	cc503	Larval
	cc504	Larval
	cc520	Larval
let-853	*cc505*	Embryonic
	cc506	Embryonic
	cc521	Larval
let-854	*cc507*	Embryonic
	cc508	Larval
let-855	*cc509*	
	cc510	
	cc511	
	cc512	All Larval
	cc513	
	cc522	
	cc523	
	cc524	
	cc525	
let-856	*cc514*	
	cc515	
	cc526	All Larval
	cc527	
	cc528	
	cc529	
	cc530	
let-857	*cc516*	Embryonic
	cc517	Embryonic
	cc518	Larval
	cc531	Larval
	cc532	Larval
	cc533	Larval
let-858	*cc500*	All embryonic
	cc501	
	cc534	

Thirty-four mutations in seven complementation groups recovered from the EMS screen are grouped by gene designation (*let-852* to *let858* and allele designation (cc no.). Embryonic arrest, alleles for which homozygous embryos from heterozygous mothers failed to hatch; arrested, those arresting after hatching but before the L4 larval molt.

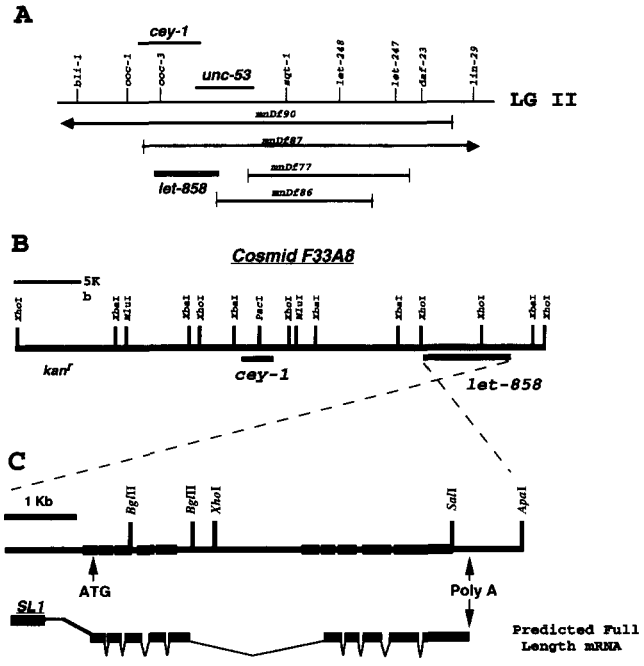

FIGURE 1.—Genetic and physical organization of the *let-858* locus. The position of *let-858* relative to nearby deficiencies and genetic markers on chromosome II, is shown in A. *cey-1* and *unc-53* are genes that have been physically mapped to the region. (B) The physical restriction map of cosmid F33A8; the locations of *let-858* and *cey-1* are indicated. (C) The genomic organization is shown of the *let-858* rescuing fragment. The minimal rescuing activity is confined within a *Exo*III deletion endpoint at the 5′ end and an *Apa*I site at the 3′ end; truncations at the *Xho*I (5′) and/or *Sal*I (3′) sites abolished rescuing activity. A larger construct from flanking *Xba*I sites was used to probe a mixed-stage *C. elegans* cDNA library, and four cDNAs were obtained and sequenced.

charged residues). The predicted protein contains a cluster of cysteine residues (five cysteines between positions 521 and 560) flanked by a number of histidine residues (505, 568, 584), which could be a metal binding domain (although not strictly resembling any known "zinc finger" motif). A striking serine and aspartate-rich stretch is found near the C terminus (residues 696–729; Figure 2). Searches for genes in Genbank with similarity to *let-858* (using BLAST; ALTSCHUL *et al.* 1990) have yielded several expressed sequences from other species [human, mouse, Arabidopsis, and Saccharomyces (Figure 2)]. The identified coding segments from these species have remarkable conservation at the amino acid level to corresponding regions of the LET-858 protein. A single putative homologue has been identified in the yeast genome sequencing project; likewise, there has thus far been only one *let-858* gene identified in the 50–60% of the worm genome sequenced to date. The coding regions from the other species were all derived from large-scale cDNA or genomic sequencing efforts; no information is available concerning the function of any of these species. The highly amphoteric nature and intracellular localization (below) of the *C. elegans* homologue led us to choose the name "nucampholins" for this group of homologous proteins.

Distribution of *let-858* products: A *let-858* genomic clone was used to probe a Northern blot containing mRNA from different developmental stages. A band of ~2.8 kb was detected; this was most abundant in embryos and larvae, with decreased abundance in adults (Figure 3). Much of the adult message appeared to be contributed by the germline, as evidenced by the decrease in adults homozygous for a germline deficient mutation, *glp-4* (*bn2*). *In situ* analysis of the distribution of the *let-858* mRNA in embryos showed that RNA to be present in the cytoplasm of all cells at all embryonic stages examined (Figure 4, A and B).

To assess the cellular and tissue distribution of the protein, antibodies were generated against a bacteria-expressed fusion carrying amino acids 139–336 of LET-858. These antibodies were used for immunofluorescence localization in whole mount preparations. As shown in Figure 4, antibody fluorescence is observed to be localized to the nuclei in most adult tissues and in all cells in embryos (Figure 4, C–E). The nuclear staining is nonuniform in adult tissue nuclei; the LET-858 protein is localized throughout the nucleus in a somewhat globular pattern that overlaps with but is not restricted to, the regions filled by chromosomes (as visualized by the DNA-binding dye DAPI; not shown). The protein appears to be excluded from the nucleolus. During mitosis, the protein is released from the nucleus (Figure 4D). In the adult germline, an association of LET-858 with chromatin is evident as the germ cells mature into oocytes; subsequently LET-858 disperses throughout the nucleoplasm as the oocytes become fully diakinetic (not shown). The antibody signal in germline nuclei is stronger than that in somatic tissues (Figure 4E, inset); this is consistent with the germline enrichment for *let-858* mRNA seen by Northern analysis (Figure 3). Pre-immune controls exhibited a diffuse background fluorescence with no nuclear concentration of signal (not shown).

Loss of function phenotypes for *let-858*: When homozygous *let-858* mutant animals are derived from a heterozygous mother, a uniform arrest phenotype is seen (Figure 5A). These embryos arrest after partial elongation, indicating the occurrence of early hypodermal differentiation (PRIESS and HIRSCH 1986). Functional muscle is made, as indicated by the fact that the animals move, and gut and pharyngeal differentiation is observed. The animals eventually stop moving, fail to hatch, and develop large vacuoles that arise most frequently in the head region.

Depletion of maternally donated *let-858* by antisense RNA results in a more severe arrest phenotype: Since maternal contributions of *let-858* RNA and protein have been observed, it is likely that the phenotypes of homozygous offspring from a heterozygous mother are not reflective of a complete loss of function but rather due

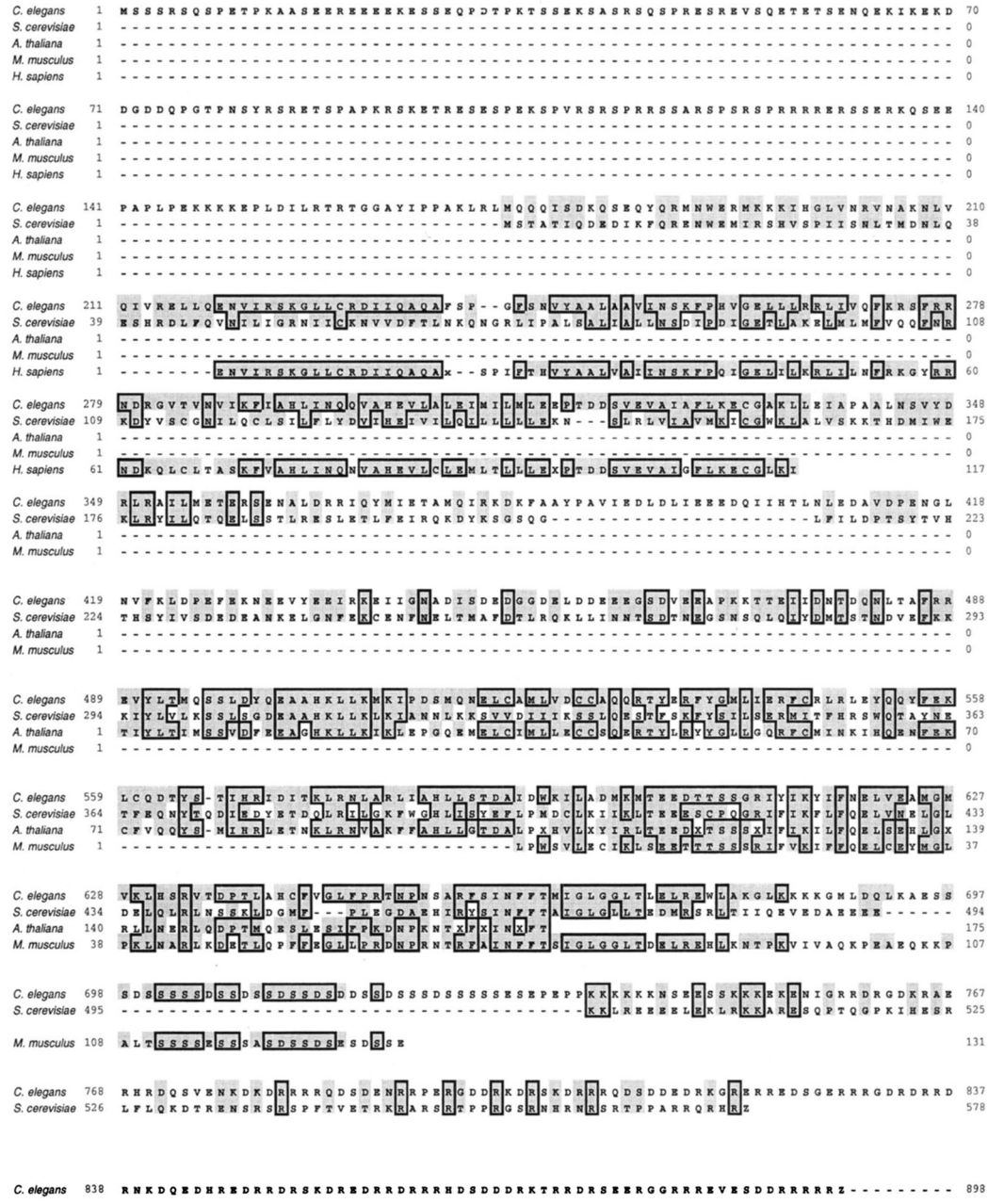

FIGURE 2.—Sequence and evolutionary conservation of nucamphilins. The predicted protein sequence for *C. elegans* LET-858 is depicted and aligned (using ClustalW software with the BLO-SUM weight matrix; THOMPSON *et al.* 1994) with predicted protein sequences of homologous sequences recovered from database searches. The *S. cerevisiae* sequence is a predicted open reading frame obtained from the *S. cerevisiae* Genome database; the other sequences are predicted from cDNAs in Expressed Sequence Tag databases. Sequence identities between at least two of the compared sequences are outlined; conservative amino acid changes are shaded. Residues identified as an "X" reflect ambiguously assigned amino acid residues from the cDNA sequence. Note that the possible metal binding residues in *C. elegans* are conserved in comparisons to the putative Arabidopsis homolog but are not seen in the Saccharomyces gene. The DNA sequence of *let-858* mRNA has been submitted to GenBank (accession no. U19615).

to the eventual depletion of maternally donated products. To more effectively examine development in the absence of *let-858* product, we have used an antisense RNA approach as pioneered by GUO and KEMPHUES (1995). Antisense RNA has been seen to specifically and effectively eliminate expression for a wide variety of maternally expressed genes (*e.g.*, GUO and KEMPHUES 1996a; MELLO *et al.* 1996). Antisense injections for *let-858* lead to a much earlier arrest phenotype than that observed in homozygous offspring (Figure 5B). Antisense-arrested embryos develop into a ball of ~120 cells, with no evident differentiation of gut, muscle, pharynx or hypodermis. Although some early cell movements were observed, we did not observe gastrulation, which would normally have occurred before the antisense-induced arrest of cell divisions (not shown). This gastrulation defect phenotype is similar to that observed in embryos arrested by antisense RNA from the large subunit of RNA polymerase II, *ama*-1 (POWELL-COFFMAN *et al.* 1996), and other mutations that presumably interfere with early zygotic gene expression (see BUCHER and SEYDOUX 1994).

A requirement for *let-858* in the proliferation of the germline is revealed by transgene rescue experiments: *C. elegans* transgenes are generally found to consist of large tandem arrays assembled from multiple copies of the injected DNA; these arrays will frequently give rise to structures that are transmitted as stable and heritable extrachromosomal arrays (STINCHCOMB *et al.* 1985). An observation made by several groups has been

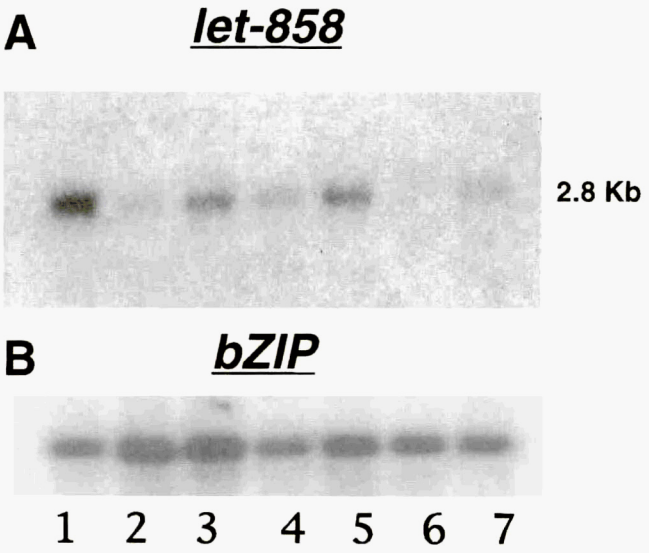

FIGURE 3.—Northern blot analysis of *let-858* mRNA. (A) Poly-A+ RNA from developmentally staged animals was purified, separated by gel electrophoresis, blotted and probed with a random priming-labeled *let-858* cDNA. Lane 1, embryos; lane2, L1 larvae; lane 3, L3/L4 larvae; lane 4 and 5, 1X and 2X loading of adult RNA, respectively; lane 6, *glp-4(bn2)* adults; lane 7, *fem-1* adults. (B) As a loading control, the same blot was reprobed with a labeled construct specific for *bZIP*, a factor that is present in all developmental stages (P. OKKEMA, personal communication).

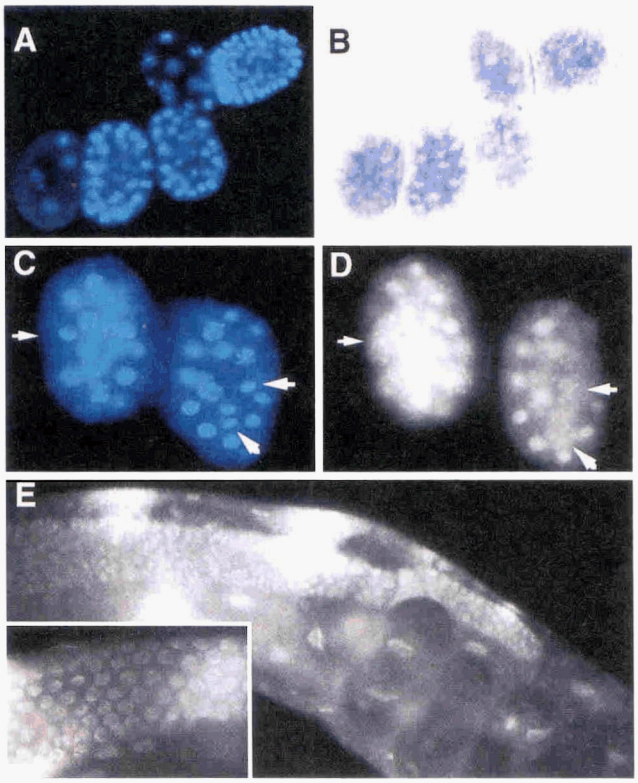

FIGURE 4.—Localization of *let-858* gene products. (A and B) *In situ* analysis of *let-858* mRNA in whole mount wild-type embryos. Embryos were hybridized to *let-858* antisense probe visualized using alkaline-phosphatase-mediated detection. *let-858* mRNA (purple color) appears uniformly localized in all stages of embryos; A, DAPI; B, alkaline phosphatase. (C–E) Antibody analysis of *let-858* protein in whole mount wild-type embryos (C and D) and adult (E). FITC-labeled secondary antibodies were used to locate specific binding of mouse anti-LET-858 antibodies. Arrow heads in C and D indicate mitotic cells; the inset in E is an enlargement of the pachytene-stage nuclei in the adult gonad.

that such arrays tend to express poorly in germline tissues (FIRE *et al.* 1990; HOPE 1991). We found that array transgenes made from *let-858* genomic constructs were capable of efficiently rescuing the embryonic lethality seen in homozygotes, such that these animals developed to adulthood. The resulting adult animals, however, were sterile (Figure 5D). Examination of these animals revealed a lack of germline proliferation: whereas the two primordial germ cells, Z2 and Z3, were observed in newly hatched larvae (not shown), no evidence for germ cells was seen in later larvae and in adults. This phenotype is similar to that seen in *glp* (germ line proliferation) mutants (SCHEDL 1996).

Using sequences from the *A. victoria* green fluorescent protein (GFP), a *gfp*-tagged version of *let-858* was used to further investigate the expression and function of the gene. This construct was produced by inserting the GFP coding region, in frame, into the fourth exon of *let-858*. All *let-858* coding sequences are still present in the resulting open reading frame. This construct, when introduced into standard transgene arrays in a wild-type genetic background, results in GFP fluorescence that is highly nuclear localized. In established transgenic lines, activity is observed in all somatic nuclei, yet a striking absence of GFP activity is seen in the germline (Figure 6, A and B). Like the parent construct, the *gfp*-tagged *let-858* rescues the embryonic defect but yields sterile adults.

In newly hatched animals that are homozygous for a chromosomal *let-858* mutation but carry the transgene array, GFP fluorescence is seen in all somatic cells, but is markedly absent in the primordial germ cells Z2 and Z3. Later in development, the gonad appears to consist entirely of GFP-positive somatic gonadal cells (not shown). The results are consistent with the hypothesis that lack of rescue in the germline is due to poor transgene expression in germline tissue. In addition, these results indicate that in the absence of *let-858* function, the germ cells are incapable of proliferation.

Generational silencing of *let-858* transgenes in the germline: We examined in detail the expression of *gfp*-tagged *let-858* constructs during the establishment of transgenic lines. In these experiments, we saw occasional low-level expression of GFP in the adult germline during the first generations following injection. After a number of generations, as the lines were established, no expression in the germline was observed. These results suggest a generational silencing of *let-858* transgenes in the germline. This silencing is restricted to the germline expression of the transgene, as the levels observed in

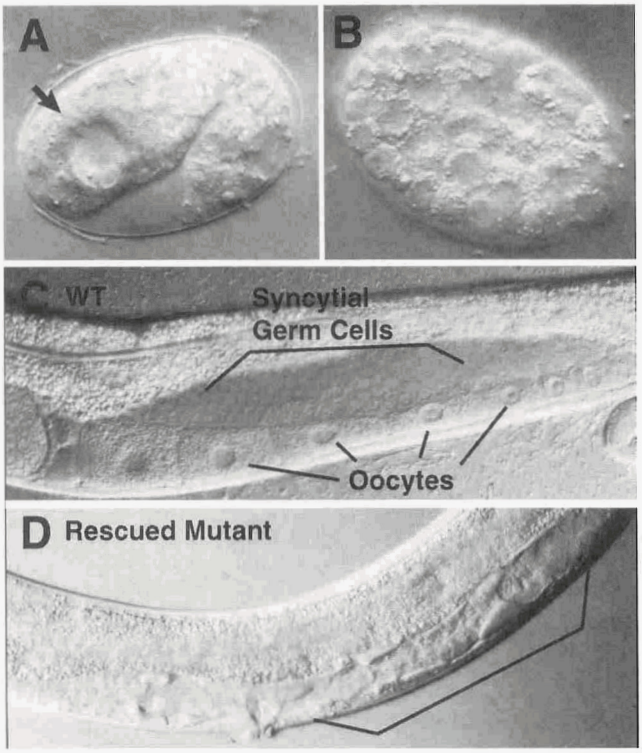

FIGURE 5.—*let-858* loss of function phenotypes. (A) Arrest phenotype of *let-858* homozygous offspring from heterozygous parent. Large vacuole in head region is indicated. (B) Arrest phenotype of F_1 embryo from wild-type parent injected with anti-*let-858* mRNA into its syncytial gonad. (C) Distal and proximal region of one arm of a wild-type adult hermaphrodite. (D) Similar region of the residual gonad seen in *let-858* homozygotes rescued from embryonic lethality in a line transgenic for the 6.8-kb *let-858* genomic construct shown in Figure 1C. While much of the somatic structures of the gonad appear to be present (bracket), descendants of the Z2 and Z3 germ cell lineage are not detected.

somatic tissue do not appreciably change (Figure 7). Pedigree analyses of lines in which low levels of germline expression are initially observed show that expression becomes increasingly mosaic, with expression often only detectable in one arm of the gonad. The frequency of silencing increases with each generation, until germline expression is effectively absent. The silencing does not appear to be easily reversible, since we have not observed the re-emergence of expression in silenced lines (data not shown).

Context-mediated desilencing of *let-858* transgenes in the germline: We undertook two approaches toward obtaining *let-858* transgenes with germline function. In initial experiments, we looked for animals in which the germline silencing of the array was relieved. This screen used a strain carrying a *let-858* transgene, with the initial genetic background being heterozygous for a chromosomal *let-858* mutation. A linked marker (*unc-4*) was used to obtain a population of animals homozygous for the chromosomal *let-858* mutation. As noted above, most of these animals are sterile. We screened this population for fertile animals, looking for rescue of the

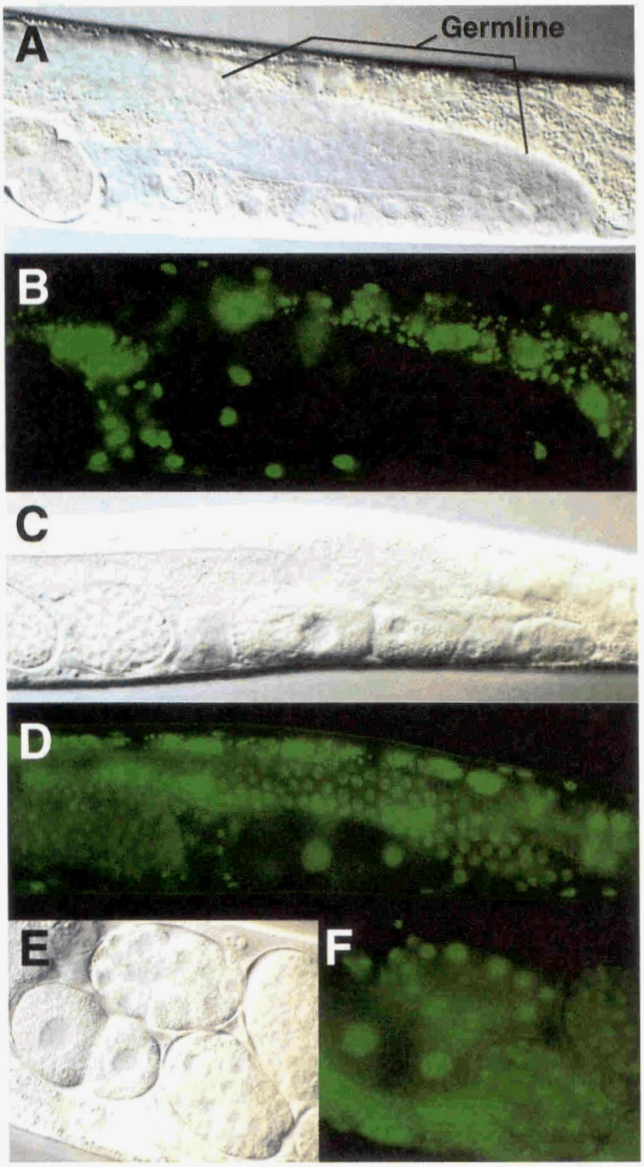

FIGURE 6.—Context-dependent germline expression of a *let-858::gfp* reporter construct. (A and B) Nomarski (A) and fluorescence (B) micrographs showing one gonadal arm of a heterozygous *let-858* adult transformed with an in-frame *let-858::gfp* fusion construct, pBK48.1, by the standard injection protocol [plasmids pBK48.1 and pRF4 (*rol-6d*) in a 1:5 ratio by mass]. Transgenic animals obtained by this protocol that are homozygous for the *let-858* mutation exhibit an identical phenotype to that illustrated in Figure 5D. The bracket indicates the syncytial germ cell region of the gonad. (C–F) Nomarski (C and E) and fluorescence (D and F) micrographs of one gonadal arm (C and D) and early embryos (E and F) in *let-858* homozygotes transformed by pBK48.1 coinjected with a 50- to 100-fold mass excess of random genomic DNA fragments as described in MATERIALS AND METHODS.

germline defect that was dependent on the transgene. One such line was obtained, and further analysis of this line indicated that the transgene array had been altered and was now capable of germline expression. Although the nature of the change in this transgene has not been further investigated, its properties indicated that cer-

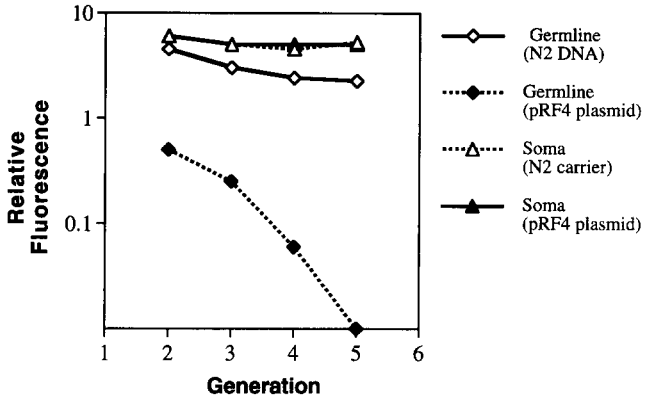

FIGURE 7.—Context-dependent generational silencing of extrachromosomal *let-858::gfp* reporter construct. Wild-type hermaphrodites were injected with the pBK48.1 *let-858::gfp* reporter construct either in the absence (♦ and ▲) or presence (◇ and △) of N2 DNA fragments as described (MATERIALS AND METHODS). Transgenic lines resulting from these injections were followed for four generations, starting with the F_2, and assessed for *gfp* fluorescence intensity in both soma (▲ and △) and germline (♦ and ◇). Relative intensity was determined by comparing nuclear *gfp* in the respective tissues with that observed in a "benchmark" line (see MATERIALS AND METHODS).

tain features of the transgene array might be critical for the observed silencing effects.

Therefore, as a second approach toward obtaining transgene activity in the germline, we hoped to modify the transformation protocol so that the repetitive nature of the transgene arrays was minimized. In standard transformations, the test construct is injected in mixtures composed of only the test plasmid and another plasmid (usually of similar vector backbone) containing a dominant transformation marker. This necessarily results in the *in vivo* construction of low complexity arrays of a highly repetitive nature. In various other genetic systems, tandem repeats have been shown to lead to heterochromatic effects resulting in gene silencing (DORER and HENIKOFF 1994; SABL and HENIKOFF 1996). In previous studies of somatic enhancer function, we have observed that certain transgenes exhibit decreased activity or are inactive in the context of these long tandem arrays (OKKEMA *et al.* 1993; KRAUSE *et al.* 1994). We have recently observed that some of these effects on somatic enhancer function can be mitigated if more complex carrier DNAs are used in the transformation protocol (A. FIRE, S. XU, J. HSICH and B. HARFE, unpublished observations). In an attempt to minimize the effect of repetitive arrays on gene expression in the germ line, we performed a set of injections in which short linear random segments of genomic DNA from *C. elegans* was used as carrier. It was hoped that the tested construct would be dispersed and become imbedded in regions of high sequence complexity within the array that might approximate euchromatin.

Numerous transgenic lines resulting from co-injection of *gfp*-tagged *let-858* with cleaved genomic DNA were obtained. Of these, the majority showed robust expression in both germline and soma (Figure 6, C–F). Examination of expression over several generations showed that the resulting transgenic lines could maintain their germline expression (Figure 7), although in some cases we have seen a decrease with continued passage (not shown). Several of the lines with germline expression were tested for mutant rescue and were found to rescue both the somatic and germline defects of *let-858* mutations. The fertile animals resulting from rescue could be maintained as homozygotes indefinitely (for ≤24 generations to date).

The improvement in germline expression in lines constructed using *C. elegans* DNA as carrier could conceivably be due either to the nature of the carrier DNA or to the effective dilution of *let-858::gfp* sequences with excess carrier. To distinguish these possibilities, we carried out a set of standard transformation experiments in which the *gfp*-tagged *let-858* plasmid was diluted with increasing ratios of the *rol-6* plasmid pRF4 molar ratios of 3:1, 48:1, and 192:1 (*rol-6/let-858gfp*) were used; no complex carrier DNA was included. We found that such simple dilutions of *let-858* sequences were not sufficient to produce transgenic lines with germline expression (five to six lines were examined for each concentration). Instead, dilution of *let-858* sequences resulted in lines that exhibited overall decreases in somatic expression while germline expression was not observed (data not shown). This indicated that the nature of the carrier DNA rather than the dispersal and dilution of *let-858::gfp* sequences was the key feature in allowing germline expression.

Proximal sequence requirements for germline and somatic expression of *let-858*: We looked for *cis*-acting sequences in the *let-858* rescuing fragment that might be specifically required for its germline expression. This analysis involved sequential deletions of the 5′ flanking DNA (Figure 8A), deletions of intron and exon sequences (Figure 8B), and replacement of the 3′ UTR with that of a non-germline gene (Figure 8B, pPD100.29). In all cases, we found similar requirements for germline and somatic expression; *i.e.*, removal of any element affecting germline expression had concomitant effects on somatic expression. At the 5′ end of the gene, deletion to −184 leads to a dramatic loss in activity for both soma and germline, while −201 deletions are fully active in both tissues. Although this data strongly suggests a similar set of transcription factors mediating germline and soma expression of *let-858*, we cannot rule out the possibility that separate germline and somatic activating sequences exist in the intervening region.

Context-dependent silencing in the germline may be a general feature of ubiquitously expressed genes: We used similar methods to examine the requirements for germline and somatic expression of a second broadly expressed gene, *unc-37*. MILLER and colleagues (PFLU-

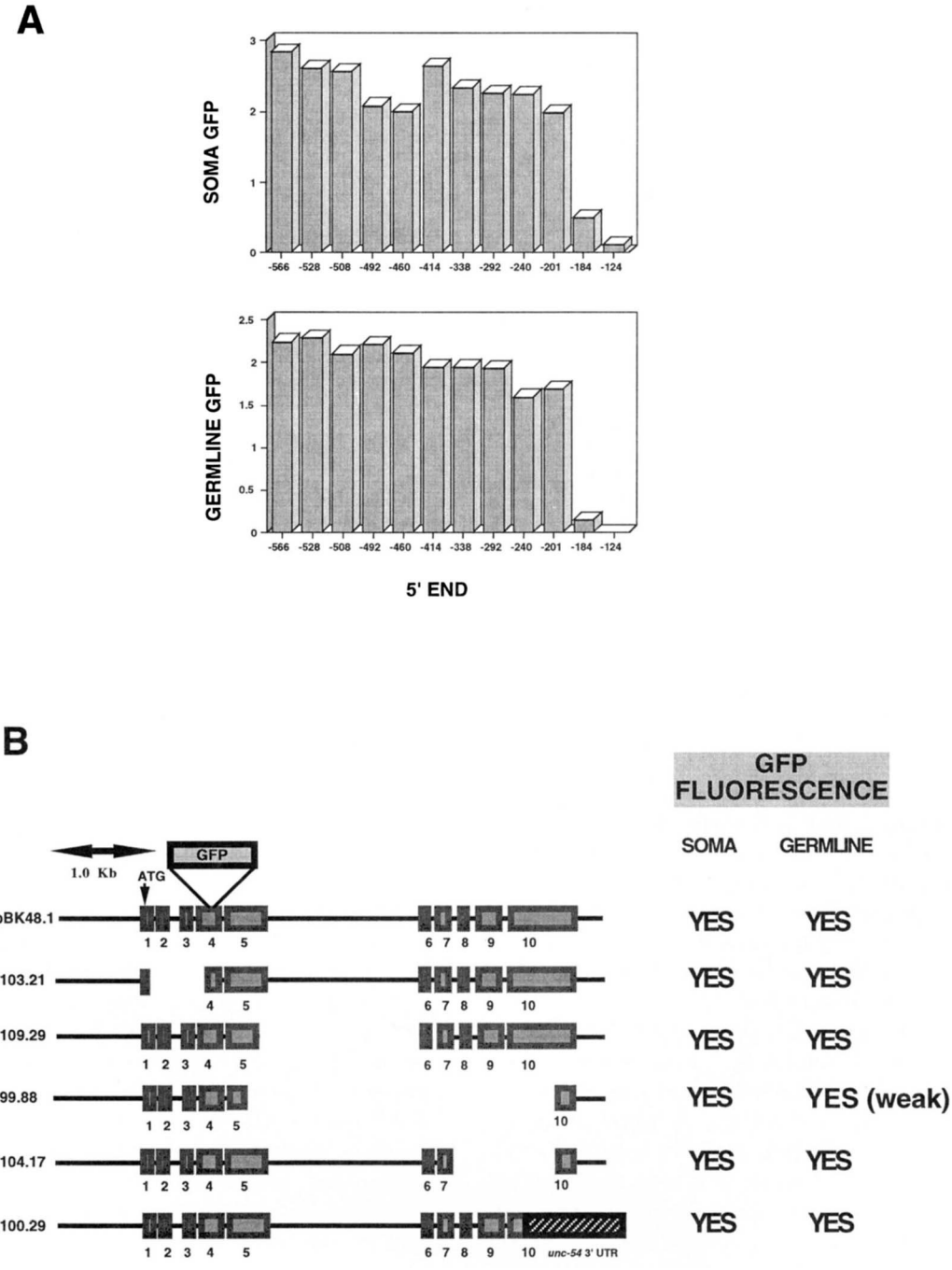

FIGURE 8.—Deletion analyses of *let-858* genomic sequences for the identification of germline and somatic specific control elements. (A) 5′ deletions: deletions in the 5′ flanking DNA of pBK48.1 were constructed, co-injected with N2 genomic fragments, and the resulting lines were assessed for *gfp* fluorescence in the soma and germline as described (MATERIALS AND METHODS), with the data presented as the average of all lines tested for each construct. Relative fluorescence refers to the nuclear *gfp* intensity in the respective tissue compared with the somatic nuclear intensity of the benchmark line, which was assigned a value of 3.0. *x* axes indicate the most 5′ bp from *let-858* genomic sequence remaining in each construct, with the first basepair of the initiator codon set to +1. (B) Exon/intron deletions, 3′UTR replacement: in-frame deletions of coding sequences and intervening sequences of constructs derived from the parent pBK48.1 plasmid are illustrated as gaps in the sequence diagram of each construct. pPD109.29 was generated by the replacement of the genomic sequence with most of the corresponding cDNA sequence, thereby creating an in-frame deletion of the 3′ end of exon 5, and all of introns 5 and 6. pPD100.29 is a truncation of exon 10 in which part of the COOH terminus, and all of the *let-858* 3′ UTR has been replaced with the 3′ UTR of the myosin heavy-chain gene, *unc-54*.

GRAD *et al.* 1997) have shown that *unc-37* encodes the *C. elegans* homologue of the Drosophila *groucho* transcription factor, and that this factor is present in the germline and in all somatic cell types. We tagged *unc-37* with *gfp*, using a similar approach to that used for *let-858*. When introduced into high copy tandem arrays

by standard transformation techniques, we observe initial low levels of expression in germline tissue, followed by complete silencing within a few generations as observed with *let-858*. Ubiquitous somatic expression is observed continuously. Injection of the *gfp*-tagged *unc-37* in the context of complex *C. elegans* carrier DNA results in strong expression in the germline. This expression persists upon passage. Thus *unc-37* transgenes exhibit many of the properties observed for *let-858*. Another broadly expressed gene, *cap-2*, has recently been assayed using our protocol for transformation with complex carrier (J. WADDLE, personal communication); this gene also appears to depend on the presence of complex carrier DNA for germline expression.

DISCUSSION

Distinct contextual requirements for somatic and germline gene expression: We have found that germline and somatic tissues of *C. elegans* display distinct criteria for efficient transgene expression. Expression in the soma has long been known to occur from large arrays of injected plasmid sequences; in at least some cases, this expression has been shown to reflect proper regulation. By contrast, definitive expression of reporter transgenes in the *C. elegans* germline has never been observed before this work.

Our data indicate that for at least two ubiquitously expressed genes, the context of the introduced transgene is critical in determining whether expression will occur in the germline. Embedding of these transgenes within complex sequences derived from the *C. elegans* genome is sufficient to allow stable expression in the germline. One possible explanation for this is that the germline might view more repetitive arrays as heterochromatic and act to silence their expression. Alternatively, specific features of *C. elegans* sequences might be required for the relief of silencing in the germline. If these features involve a specific sequence element, then these elements would have to be either abundant in chromosomal DNA and/or capable of acting over large distances, because the majority of N2 DNA-containing arrays (each with different DNA composition and structures) proved capable of germline expression.

The silencing of the repetitive arrays by the germline occurs over the course of several generations, by a process that is not understood and for which we know of no clear precedent. Silencing effects on repeated transgenes have been seen in a variety of species (MARTIN and WHITELAW 1996). The sensitivity to repetitive sequence, and the increasing mosaicism, are reminiscent of position effect variegation and repeat-induced silencing seen in Drosophila (DORER and HENIKOFF 1994; SABL and HENIKOFF 1996). These observations lead us to the hypothesis that the *C. elegans* germline is perceiving these arrays as heterochromatic, suggesting that different sets of heterochromatin sensing and/or establishment mechanisms may exist in soma and germline. Perhaps the germline maintains a generalized silencing of DNA, with germline specific and housekeeping genes protected from silencing by specific sequence elements.

This model would suggest that germline chromatin is structurally different from somatic chromatin. Indeed, chromatin morphology in midstage embryonic germline nuclei appears more compact than that of surrounding somatic cells. This difference is seen as early as the P4 blastomere and is maintained through into the early stages of larval germline proliferation (W. KELLY, unpublished observations; G. SEYDOUX, T. SCHEDL, J. PRIESS, personal communications). From this perspective, it is interesting to note that DPY-26, a structural protein involved in *X* chromosome-specific transcriptional repression in somatic tissue, is bound to both autosomal and *X* chromosomal DNA in germ cells (LIEB *et al.* 1996).

The *C. elegans* germline employs multiple selective mechanisms to prevent expression of somatic genes: We have demonstrated in this paper that the *C. elegans* germline is capable of silencing the expression of genes presented in a repetitive array context. We hypothesize that this is one of the means by which the germline prevents somatic gene expression and consequently avoids differentiation along somatic pathways. We have found, using *let-858* and *unc-37* reporter transgenes in repetitive arrays, that this germline silencing mechanism is active throughout larval and adult life. It is likely that this is one of several different mechanisms employed by the *C. elegans* germline to suppress the expression of somatic genes. In the early embryo, a different problem is presented to the germline. Rapid divisions of the germline precursors act in a stem-cell-like fashion to produce a series of somatic blastomeres that rapidly acquire aspects of their somatic identity. Previous studies (SEYDOUX and FIRE 1994) have suggested that at those early stages, a complete block to germline mRNA expression is in place. It is conceivable that at this sensitive stage in which germline and somatic cells diverge from a common cytoplasm, there may be a more stringent requirement for preventing early or inappropriate somatic expression. The maternal factor *pie-1* has been shown to play a key role in repression of new gene expression in germline blastomeres (SEYDOUX *et al.* 1996). *pie-1* products are not detectable following the middle stages of embryogenesis, so it seems unlikely that PIE-1 plays a direct role in later suppression of somatic gene expression (MELLO *et al.* 1996). Intriguingly, several genes have been identified with coding sequences similar to *pie-1* (MELLO *et al.* 1996; GUEDES *et al.* 1997). It is tempting to speculate that one or more of these might play a later role in context-dependent silencing.

Whatever the mechanism of context-dependent si-

lencing, the process must be applicable at a variety of stages of differentiation. This includes the mitotically quiescent germ cells in the late embryo and early L1 larva (Z2 and Z3), the mitotically active germ cells in later larvae, and the both mitotically and meiotically active stages of germ cells seen in adults. Some common feature of these nuclei must be responsible for their germ-line specific behavior.

In addition to the mechanisms described above, there may be additional mechanisms that protect the germline from expression of somatic or foreign genes. A suggestion of at least one such mechanism comes from our (so far unsuccessful) attempts to use *let-858* and *unc-37* to generate general vectors for expressing arbitrary coding sequences in the germline. We find that fusions containing just 5′ sequences from *let-858* or *unc-37* are sufficient to drive somatic, but not germline accumulation of GFP and β-galactosidase reporter proteins (W. KELLY and A. FIRE, unpublished observations). Thus sequences within these two genes, or features of the mRNA or protein sequences, are important for germline activity.

In addition to the potential roles for silencing in developmental distinctions, it is conceivable that silencing mechanisms help to prevent proliferation of viruses and transposons. From this perspective, it is intriguing that the dispersed multicopy transposon Tc1 is preferentially active in somatic tissue (EMMONS and YESNER 1984).

LET-858 defines an evolutionarily conserved and ubiquitous nuclear factor that is essential in somatic and germline development: We have given the name "nucampholin" to the *let-858* product to indicate the nuclear localization and amphipathic nature of this product. The apparent evolutionary conservation in single cell and multicellular eukaryotes underscores the importance of nucampholin in cellular events. At this point, we have defined the *C. elegans* nucampholin encoded by *let-858* at the genetic, molecular and cytological levels. The nuclear localization of the protein is suggestive of a key role in gene expression or in overall nuclear architecture. Consistent with this hypothesis, the *let-858* antisense phenotype is very similar in timing and severity to that reported for the elimination of RNA polymerase II function (POWELL-COFFMAN *et al.* 1996). The disruption of gastrulation is also observed in *zen-1* mutants. *zen-1* encodes the SL1 transplice leader precursor involved in the normal processing of the majority of *C. elegans* mRNAs (FERGUSON *et al.* 1996; see also BUCHER and SEYDOUX 1994). Injection of *let-858* antisense RNA can also interfere with the expression of some early zygotic genes (W. KELLY, unpublished observation), although it is not clear whether this interference is at a transcriptional or posttranscriptional level. From this perspective, it is interesting that changes in the subnuclear distribution of the LET-858 protein during the course of adult germ cell development parallel the transcriptional activity (STARCK 1977) of these nuclei: during transcriptionally active stages before diakinesis, LET-858 appears to colocalize with chromatin; in transcriptionally silent mature oocytes LET-858 is uniformly dispersed throughout the nucleus.

We thank G. SEYDOUX, L. COLES, A. COULSON, D. MILLER, J. MEIR, J. PRIESS, T. SCHEDL, J. WADDLE, P. OKKEMA, B. HARFE, J. GOODLIFFE, S. KOSTAS, J. GALL, A. PINDER, and members of this lab for their help, sharing of unpublished data, and suggestions through the course of this work. This work was supported by U.S. Public Health Service grants (A.F., R01-GM-37706; W.K., F32-HD-07794; M.M., F32-GM-17164) and the Carnegie Institution of Washington. Some strains used in this work were provided by the Caenorhabditis Genetics Center (funded by National Institutes of Health Center for Research Resources).

LITERATURE CITED

AUSTIN, J., and J. KIMBLE, 1987 *glp-1* is required in the germ line for regulation of the decision between mitosis and meiosis in *C. elegans*. Cell **51:** 589–599.

ALTSCHUL, S., W. GISH, W. MILLER, E.MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. J. Mol. Biol. **215:** 403–410.

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

BUCHER, E. A., and G. SEYDOUX, 1994 Gastrulation in the nematode *Caenorhabditis elegans*. Semin. Dev. Biol. **5:** 121–130.

CHALFIE, M., Y. TU, G. EUSKIRCHEN, W. W. WARD and D. C. PRASHER, 1994 Green fluorescent protein as a marker for gene expression. Science **263:** 802–805.

DORER, D. R., and S. HENIKOFF, 1994 Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell **77:** 993–1002.

EMMONS S. W., and L. YESNER, 1984 High-frequency excision of transposable element Tc1 in the nematode C. elegans is limited to somatic cells. Cell **36:** 599–605

FERGUSON, K. C., P. J. HEID and J. H. ROTHMAN, 1996 The SL1 *trans*-spliced leader RNA performs an essential embryonic function in *Caenorhabditis elegans* that can also be supplied by SL2 RNA. Genes Dev. **10:** 1543–1556.

FIRE, A. 1986 Integrative transformation of *Caenorhabditis elegans*. EMBO J. **5:** 2673–2680.

FIRE, A., S. W. HARRISON and D. DIXON, 1990 A modular set of lacZ fusion vectors for studying gene expression in Caenorhabditis elegans. Gene **93:** 189–198.

FIRE, A., D. ALBERTSON, S. W. HARRISON and D. G. MOERMAN, 1991 Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. Development **113:** 503–514.

GUO, S., and K. J. KEMPHUES, 1995 *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative ser/thr kinase that is asymmetrically distributed. Cell **81:** 611–620.

GUO, S., and K. J. KEMPHUES, 1996a A non-muscle myosin required for embryonic polarity in *Caenorhabditis elegans*. Nature **382:** 455–458.

GUO, S., and K. J. KEMPHUES, 1996b Molecular genetics of asymmetric cleavage in the early *C. elegans* embryo. Curr. Opin. Genet. Dev. **6:** 408–415.

GUEDES, S., and J. R. PRIESS, 1997 The C. elegans MEX1 protein is present in germline blastomeres and is a P granule component. Development **124:** 738–739.

HERMAN, R. K., 1978 Crossover suppressors and balanced recessive lethals in *Caenorhabditis elegans*. Genetics **88:** 49–65.

HOPE, I. A., 1991 "Promoter trapping" in Caenorhabditis elegans. Development. **113:** 399–408.

KIMBLE, J., and D. HIRSCH, 1979 The post-embryonic lineage of the hermaphrodite and male gonads in *Caenorhabditis elegans*. Dev. Biol. **70:** 396–417.

KOERNER, T. J., J. E. HILL, A. M. MYERS and A. TZAGOLOFF, 1991 High-expression vectors with multiple cloning sites for construction of *trpE* fusion genes: pATH vectors. Methods Enzymol. **194:** 477–490.

KRAUSE, M., S. WHITE-HARRISON, S. XU, L. CHEN and A. FIRE, 1994 Elements regulating cell and stage-specific expression of the *C. elegans* MyoD homologue *hlh-1*. Dev. Biol. **166:** 133–148.

LIEB, J. D., E. E. CAPOWSKI, P. MENEELY and B. J. MEYER, 1996 DPY-26, a link between dosage compensation and chromosome segregation in the nematode. Science **274:** 1732–1736.

MARTIN, D. I. K., and E. WHITELAW, 1996 The vagaries of variegating transgenes. Bioessays **18:** 919–923.

MELLO, C., and A. FIRE, 1995 DNA transformation. Methods Cell Biol. **48:** 451–482.

MELLO, C. C., J. M. KRAMER, D. STINCHCOMB and V. AMBROS, 1991 Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. **10:** 3959–3970.

MELLO, C. C., C. SCHUBERT, B. DRAPER, W. ZHANG, R. LOBEL *et al.*, 1996 The PIE-1 protein and germline specification in *C. elegans* embryos. Nature **382:** 710–712.

MILLER, D. M., and D. C. SHAKES, 1995 Immunofluorescence Microscopy. Methods Cell Biol. **48:** 365–394.

MULLER, F., V. BERNARD and H. TOBLER, 1996 Chromatin diminution in nematodes. Bioessays **18:** 133–138.

OKKEMA, P., S. WHITE-HARRISON, V. PLUNGER, A. ARYANA and A. FIRE, 1993 Sequence requirements for myosin gene expression and regulation in *C. elegans*. Genetics **135:** 385–404.

PATEL, N. H., and C. S. GOODMAN, 1992 Preparation of digoxigenin-labeled single-stranded DNA probes, pp. 377–381 in *Non-radioactive Labeling and Detection of Biomolecules*, edited by C. KESSLER. Springer-Verlag, Berlin.

PFLUGRAD, A., J. Y.-J. MEIR, T. BARNES and D. M. MILLER, 1997 The *groucho*-like transcription factor UNC-37 functions with the neural specificity gene *unc-4* to govern motor neuron identity in *C. elegans*. Development (in press).

POWELL-COFFMAN, J. A., J. KNIGHT and W. B. WOOD, 1996 Onset of *C. elegans* gastrulation is blocked by inhibition of embryonic transcription with an RNA polymerase antisense RNA. Dev. Biol. **178:** 472–483.

PRIESS, J. R., and D. I. HIRSCH, 1986 *C. elegans* morphogenesis: The role of the cytoskeleton in elongation of the embryo. Dev. Biol. **117:** 156–173.

SABL, J. F., and S. HENIKOFF, 1996 Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila. Genetics 142: 447–458.

SCHEDL, T., 1996 Developmental genetics of the germline, pp. 241–269 in *C. elegans II*, edited by D. L. RIDDLE, T. BLUMENTHAL, B. J. MYERS and J. R. PRIESS. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (in press).

SEYDOUX, G., and A. FIRE, 1994 Soma-germline asymmetry in the distribution of embryonic RNAs in *Caenorhabditis elegans*. Development **120:** 2823–2834.

SEYDOUX, G., and A. FIRE, 1995 Whole-mount *in situ* hybridization for the detection of RNA in *Caenorhabditis elegans* embryos. Methods Cell Biol. **48:** 323–337.

SEYDOUX, G., C. C. MELLO, J. PETTITT, W. B. WOOD, J. R. PRIESS *et al.*, 1996 Repression of gene expression in the embryonic germ lineage of *C. elegans*. Nature **382:** 713–716.

SIGURDSON, D. C., G. J. SPANIER and R. K. HERMAN, 1984 *Caenorhabditis elegans* deficiency mapping. Genetics **108:** 331–345.

STARCK, J., 1977 Radiographic study of RNA synthesis in *Caenorhabditis elegans* (Bergerac variety) oogenesis. Biol. Cell. **30:** 181–182.

STINCHCOMB, D. T., J. E. SHAW, S. H. CARR and D. HIRSH, 1985 Extrachromosomal DNA transformation of *Caenorhabditis elegans*. Mol. Cell. Biol. **5:** 3484–3496.

SULSTON, J. E., E. SCHIERENBERG, J. G. WHITE and J. N. THOMSON, 1983 The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev. Biol. **100:** 64–119.

THOMPSON, J. D., D. G. HIGGINS and T. J. GIBSON, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Communicating editor: I. GREENWALD