Development 121, 2491-2500 (1995)
Printed in Great Britain © The Company of Biologists Limited 1995

The heterochronic gene *lin-29* encodes a zinc finger protein that controls a terminal differentiation event in *Caenorhabditis elegans*

Ann E. Rougvie[1,2,*] and Victor Ambros[1,†]

[1]Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138, USA
[2]Department of Genetics and Cell Biology, and Department of Biochemistry, University of Minnesota, St. Paul, MN 55108, USA

*Author for correspondence
†Present address: Department of Biology, Dartmouth College, Hanover, NH, USA

SUMMARY

A hierarchy of heterochronic genes, *lin-4, lin-14, lin-28* and *lin-29*, temporally restricts terminal differentiation of *Caenorhabditis elegans* hypodermal seam cells to the final molt. This terminal differentiation event involves cell cycle exit, cell fusion and the differential regulation of genes expressed in the larval versus adult hypodermis. *lin-29* is the most downstream gene in the developmental timing pathway and thus it is the most direct known regulator of these diverse processes. We show that *lin-29* encodes a protein with five zinc fingers of the $(Cys)_2$-$(His)_2$ class and thus likely controls these processes by regulating transcription in a stage-specific manner. Consistent with this role, a *lin-29* fusion protein binds in vitro to the 5′ regulatory sequences necessary in vivo for expression of *col-19*, a collagen gene expressed in the adult hypodermis. *lin-29* mRNA is detected in the first larval stage and increases in abundance through subsequent larval stages until the final molt, when *lin-29* activity is required for terminal differentiation.

Key words: *Caenorhabditis elegans*, heterochronic gene, *lin-29*, zinc finger, transcription, terminal differentiation.

INTRODUCTION

Metazoan development requires the precise scheduling and temporal control of cell division, differentiation and morphogenesis in diverse cell types. The timing of particular developmental events is ultimately specified by the stage-specific control of gene expression; therefore temporal coordination of complex developmental sequences must involve an organized schedule of gene activity. For example, genes whose products promote progression through the cell cycle must be downregulated after sufficient numbers of a given cell type have been produced. Furthermore, gene regulatory circuits must also coordinate the timing of cell division and differentiation throughout the animal.

In the nematode *C. elegans*, a hierarchy of heterochronic genes, *lin-4*, *lin-14*, *lin-28* and *lin-29*, controls the relative timing and sequence of many events during larval development, including certain stage-specific cell lineage patterns (Ambros and Horvitz, 1984), dauer larva development (Liu and Ambros, 1989) and terminal differentiation (Ambros, 1989). *lin-4*, *lin-14* and *lin-28* are general temporal regulators; mutations in these genes cause temporal alterations of events in diverse cell lineages, and at several stages during larval development. In contrast, phenotypic analysis suggests *lin-29* is a more specific regulator, controlling terminal differentiation of the lateral hypodermal seam cells during the final molt (Ambros, 1989). We refer to this *lin-29*-directed switch to the terminally differentiated adult state as the larval to adult switch (L/A switch; Fig. 1A).

Genetic analysis has ordered the heterochronic genes into a pathway that controls the timing of the L/A switch (Ambros, 1989; Fig. 1A). *lin-29* is the most direct regulator of the L/A switch within this hierarchy. Worms triply mutant for *lin-4*, *lin-14* and *lin-28*, but wild-type for *lin-29*, execute the L/A switch, but do so two molts early. The abnormally early occurrence of the L/A switch in the absence of *lin-4*, *lin-14* and *lin-28* suggests that these three genes normally control the timing of the switch by temporally restricting *lin-29* activity to the fourth molt. Thus, *lin-29* is the best candidate for the direct trigger of the switch, and the upstream heterochronic genes likely control the timing of *lin-29* action.

The timing and execution of the L/A switch can be analyzed at the level of individual cells. At each of the first three molts, the lateral hypodermal 'seam' cells synthesize cuticle and divide in a stem cell-like pattern (Sulston and Horvitz, 1977; see Fig. 1B), but during the fourth and final molt, they terminally differentiate: cell division ceases and the cells fuse to form bilateral syncytia that then contribute to the synthesis of the morphologically distinct adult cuticle. The adult cuticle is also biochemically distinct from larval cuticle, since certain collagen genes are expressed specifically in either larvae or adults (Cox and Hirsh, 1985).

lin-29 gene activity is critical for the coordinate execution of all aspects of the L/A switch, including cell cycle exit, cell fusion and the switch in cuticle type. *lin-29* mutant worms fail to execute the adult-specific terminal differentiation program at the L4 molt and instead reiterate the larval program. Their

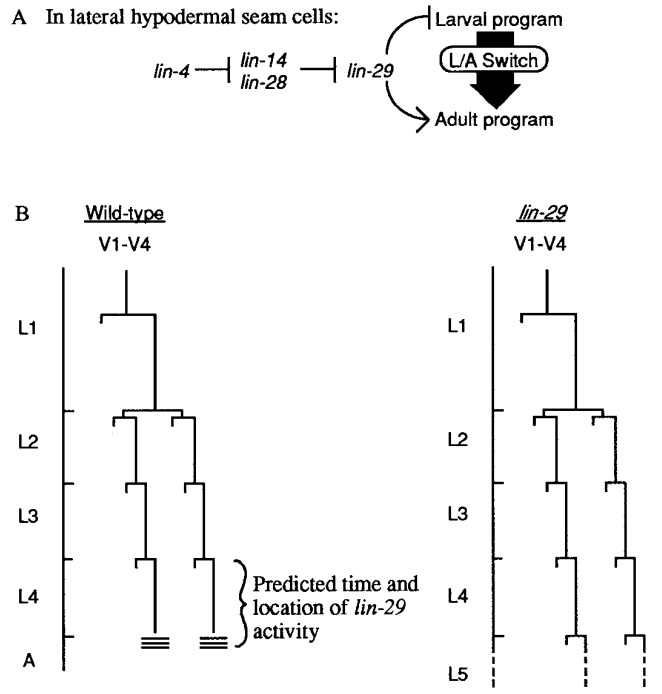


Fig. 1. Temporal control of the lateral hypodermal cell terminal differentiation. (A) Epistasis relationships among four heterochronic genes (Ambros, 1989). Briefly, *lin-29* activity is required at the final molt of the worm for the repression of larval-specific developmental programs and activation of adult-specific programs in lateral hypodermal cells. Genetically, *lin-14* and *lin-28* act as negative regulators of *lin-29* activity. Down-regulation of these two genes by the action of *lin-4* on *lin-14* is required for proper temporal activation of *lin-29*. (B) The cell lineage of lateral hypodermal seam cells V1-V4 in wild-type and *lin-29* mutant worms (Sulston and Horvitz, 1977; Ambros and Horvitz, 1984). The vertical axis represents time, with a hatch-mark to indicate each molt. L1-L4 indicate the 4 wild-type larval stages; A indicates the adult stage. L5 represents supernumerary larval stages that occur in the absence of *lin-29* activity. In the lineage diagram, the horizontal lines indicate cell divisions. In wild-type animals, seam cells divide during the first three molts and then terminally differentiate during the fourth molt, as indicated by the triple horizontal lines. In *lin-29* mutants, seam cell divisions continue during the fourth and subsequent (L5 etc.) molts, and the cells remain larval in character. Wild-type *lin-29* activity is predicted to be required during the L4 stage to trigger the L/A switch.

seam cells continue to divide in a larval-specific pattern (Fig. 1B) and to synthesize a larval-specific cuticle lacking adult-specific collagens and the characteristic adult alae (Ambros and Horvitz, 1984). These mutants undergo supernumerary molts (Fig. 1B), indicating that an additional function of *lin-29* is to limit the molts to four.

Our molecular analysis of *lin-29* has shown that *lin-29* encodes a zinc finger protein of the $(Cys)_2$-$(His)_2$ class and thus likely coordinates hypodermal cell terminal differentiation events by regulating the transcription of other genes. Furthermore, *lin-29* protein binds in vitro to the 5′ regulatory region necessary for in vivo expression of a temporally regulated hypodermal collagen gene. These data suggest that the genes directly controlled by *lin-29* include those encoding structural components of the terminal differentiation process.

MATERIALS AND METHODS

Nematode strains

The *C. elegans* strains used were: wild-type *C. elegans* var Bristol (strain N2) (Brenner, 1974), MT1835: *lin-29(n333)*, MT1176: *lin-29(n546)*, MT1834: *lin-29(n836)* and RG27: *lin-29(n1440)* (Ambros and Horvitz, 1984).

cDNA isolation

The cloned polymorphic 9.2 kb *Eco*RI fragment from *lin-29(n836)* (Papp et al., 1991) was gel-purified, radiolabeled and used to screen an L4-stage cDNA library (J. Ahringer and J. Kimble, personal communication). Several cDNA clones were obtained, each containing a 1.1 kb *Eco*RI fragment that hybridized only to the 2.4 kb *Eco*RI fragment in genomic Southern analysis (Fig. 2). The 1.1 kb fragment was used to identify clones from additional libraries (Barstead and Waterston, 1989; A. Fire; C. Martin; personal communications). An apparently full-length clone corresponding to lin-29A, pcBB71, was isolated from the Barstead library. cDNAs identified from the Fire library indicate two possible 5′ ends for lin-29B: *trans*-spliced to SL1 at the start of Exon V of lin-29A or *cis*-spliced at the start of Exon V to a 37 bp exon (5′-CTGAGATGACACACAGCTATTATTCT-GAAAAAATGAG-3′) that originates within the fourth intron, approximately 2.8 kb upstream of the start of Exon V (Fig. 2). The largest open reading frame encoded by these two lin-29B transcripts would be identical and would likely initate at the ATG codon beginning at nucleotide position 671 (Fig. 4). There are two ATGs present in the unique exon described above, but they are not in frame with the largest open reading frame. Use of the first of these two ATGs could result in the production of a 51 amino acid protein.

DNA sequencing and homology analysis

cDNA sequence was determined by the dideoxy chain termination method (Sanger et al., 1977) using double-stranded DNA (Kraft et al., 1988) and Sequenase 2.0 (USB) as directed by the manufacturer. Nucleotide and amino acid sequence analysis was performed using the GCG Program Manual for the Wisconsin Package, version 8.0 (Devereux et al., 1984). Homology searches were performed with both the GCG programs and the BLAST Network service of the National Center for Biotechnology Information.

We identified the conceptual translation of human clone hbc087 as most similar to Lin-29 through database searches using each Lin-29 zinc finger alone. Fingers 1, 2, and 3 all exhibited the highest identity with the predicted translation of the 288 b hbc087 sequence (GenBank accession #T10913), but in two different reading frames. Sequencing of the hbc087 5′ end (a generous gift of Drs J. Takeda and G. Bell, University of Chicago) identified a missing C at position 112 of the original sequence. With this correction, the three fingers fell into a single open reading and two additional zinc fingers were identified in the adjacent sequence. We determined the amino acid sequence of an additional 33 residues both 5′ and 3′ of the zinc finger domain, but we did not observe significant identities to Lin-29.

RNA analysis

Worm populations were synchronized by hatching eggs for 24-36 hours in the absence of food (Wood, 1988). The developmentally arrested L1s were then fed *E. coli* OP50 and grown until they reached the appropriate developmental stage as judged by examining vulval and gonadal development (Sulston and Horvitz, 1977).

Total RNA was isolated from synchronous populations of worms as described (Z. Liu, S. Kirch, V. Ambros, submitted). Poly(A)$^+$ RNA was selected by one round of oligo(dT) chromatography and northern blot analysis was performed as described (Ausubel et al., 1989) using formaldehyde as denaturant and nylon Biosbrane membranes (Bios Corp.). Following transfer filters were UV-irradiated for 2 minutes at 1200 $\mu W/cm^2$, and were hybridized with α^{32}P-labeled probes

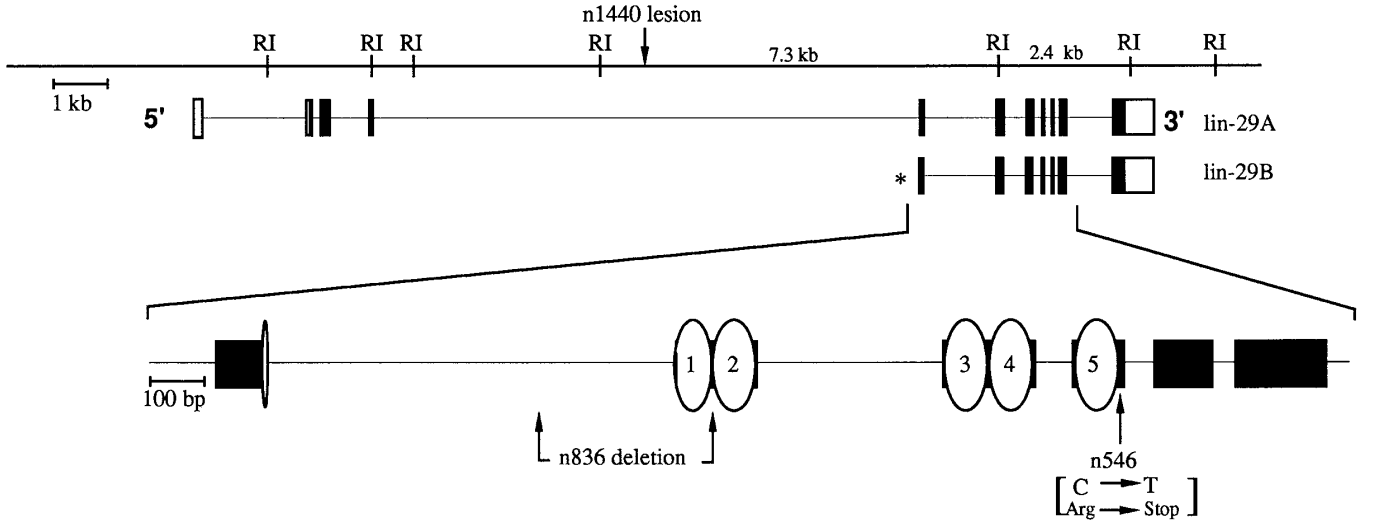


Fig. 2. *lin-29* genomic region. The top line shows the *lin-29* locus relative to *Eco*RI sites (RI). The two *lin-29* primary trancripts, lin-29A and lin-29B are indicated below the line. Transcription is shown left to right for ease of comparison with Figs 4 and 5. This is in opposite orientation relative to the genetic map (Papp et al., 1991). Exons are indicated by boxes. Open boxes indicate 5′ and 3′ untranslated regions and the shaded boxes indicate coding regions. The * indicates that two 5′-ends of the lin-29B transcription unit have been identified by cDNA analysis (see text). In one, SL1 is *trans*-spliced to the start of Exon V of lin-29A. In the second, a 37 bp exon is *cis*-spliced to Exon V. The numbered ovals in the region expanded at the bottom of the figure indicate the positions of the five zinc fingers, with the numbers corresponding to those in Figs 4 and 5. The first six amino acids of zinc finger 1 are separated from the rest by an intron. The positions of the molecular lesions of three *lin-29* mutations, *lin-29(n546), lin-29(n836)* and *lin-29(n1440)*, are indicated.

prepared as described by Feinberg and Vogelstein (1983). High stringency washes were performed in 36 mM NaCl, 2 mM NaH_2PO_4, 0.2 mM EDTA, 0.2% SDS at 65˚C.

Analysis of *lin-29* molecular lesions

The *lin-29(n546)* lesion was detected by Southern hybridization as a *Cla*I restriction site polymorphism. Oligonucleotide primers 5′ and 3′ to this site (AR4: 5′-CAATCTCACTCTCGATGCCA-3′; AR5: 5′-GAGTCATGTGCATTATCTCC-3′) were used to asymmetrically amplify and sequence a 385 bp region from genomic wild-type N2 or *lin-29(n546)* DNA (Gyllesten and Erlich, 1988). The sequence of *lin-29(n546)* was identical to wild type throughout the 385 bp amplified region (both strands) except for a single C to T transition in the *Cla*I recognition site corresponding to nucleotide position 1124 of the cDNA (Fig. 4).

The junction of the *lin-29(n836)* deletion was sequenced in pVT101RV, a 2.1 kb *Eco*RV fragment subcloned from pVT101 into pBluescriptSK⁻ (Stratagene). The sequence of the wild-type genomic region was determined from cloned PCR fragments. Primers AR7 (5′-ATGGCATCGAGAGTGAGATT-3′) corresponding to nucleotide positions 921-940 of the cDNA (non-coding strand) and AR22 (5′-GAACAAAAGCCGGACGT-3′) corresponding to positions 635 to 651 (coding) were used to amplify a 1.2 kb genomic region spanning the *lin-29(n836)* deletion site. The amplified fragment was cloned into pBluescript SK⁻ and sequenced.

DNA binding experiments

A *glutathione S-transferase-lin-29* fusion protein vector (Gst:Lin-29) was constructed by cloning a 1.7 kb *Eco*RV fragment of *lin-29* cDNA, beginning 26 amino acid residues upstream of the first zinc finger and extending through the poly(A) tail, into the *Sma*I site of pGex1 (Smith and Johnson, 1988). A Gst:Lin-29 fusion protein containing the entire lin-29A open reading frame has also been used in DNA-binding experiments (not shown), and confirms the results shown here.

Expression and purification of the fusion protein was performed as described (Smith and Johnson, 1988). The quantity of fusion protein absorbed from an initial culture volume of 5-10 ml was used in each binding reaction. DNA binding experiments were performed by a modification of procedures described by Desplan et al. (1985). To initiate a binding reaction, 20 μl of bead-protein complexes were added to 20 μl of 2× Binding Buffer [1× = 75 mM NaCl, 10 mM Tris-Cl (pH = 7.5), 1 mM EDTA, 1 mM DTT, 10 μM $ZnSO_4$] containing approximately 1 ng of end-labeled DNA and 2 μg of poly[dI:dC]. Reactions were incubated for 15 minutes on ice and the beads were washed 4 times in 1× Binding Buffer. The bound DNAs were purified from the beads by phenol extraction and ethanol precipitated. The DNA pellets were resuspended and electrophoresed through polyacrylamide gels in 1× TBE (90 mM Tris-borate, 2 mM EDTA), with 1/4 and 1/16 of the input DNA for comparison. The gels were dried and exposed to Kodak XAR-5 film. DNA fragments were restriction digested and end-labeled with α^{32}P-labeled deoxynucleotides using Klenow polymerase. When a *Hin*dIII-*Bam*HI digest was used to purify the parental *col-19* fragment for binding experiments (Fig. 6), the filled-in *Hin*dIII site ended at −845. This fragment is referred to as Δ-846 in the text for consistency with the in vivo assays in which the *Hin*dIII site is intact. The 162 bp *col-19* fragment was generated by PCR, using the *col-19*-specific primers AR23: 5′-TTGA-GATTAGTTATTGAACTTCAT-3′ and AR24: 5′-TAATGTGTTTC-CAAGAGAGATGGAC-3′. The fragment was purified and then end-labeled using T4 polynucleotide kinase and $\gamma[^{32}P]$ATP (Ausubel, 1987).

RESULTS

The *lin-29* transcribed region

Previous work mapped the *lin-29* locus to a defined region of chromosome II using a combination of chromosomal walking and restriction fragment length polymorphism (RFLP) mapping techniques (Papp et al., 1991). Distinct allele-specific RFLPs that each affected a 7.3 kb *Eco*RI restriction fragment were detected in two of the five *lin-29* alleles, *lin-29(n836)* and

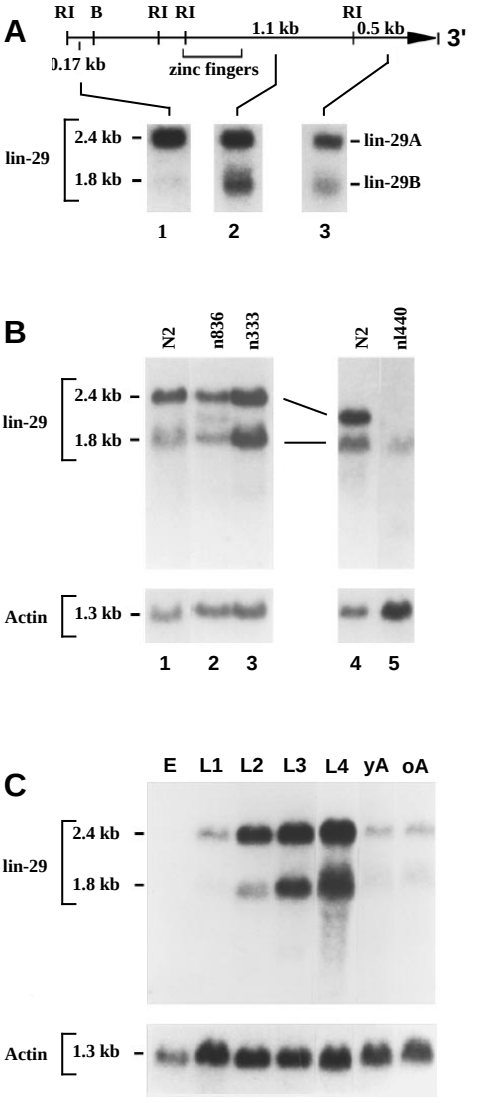


Fig. 3. Northern analysis of *lin-29* transcripts. Poly(A)$^+$ RNA was isolated from developmentally staged worm populations and analyzed by northern hybridization. Transcript sizes were estimated based on ethidium bromide staining of the 18S and 28S ribosomal RNAs in a sample of total RNA run on the same gel. (A) A restriction map of lin-29A cDNA is shown indicating the different fragments used to probe a northern blot. RI, *Eco*RI; B, *Bst*EII. Below the map are autoradiographs of a northern blot of L4-stage RNA hybridized sequentially with three different fragments: the 5′ 0.17 kb *Eco*RI-*Bst*EII fragment (Lane 1), the 1.1 kb *Eco*RI zinc finger containing fragment (Lane 2), and the 3′ 0.5 kb *Eco*RI fragment (Lane 3). (B) RNA isolated from synchronized L4-stage populations of wild-type N2 (lanes 1 and 4), and three *lin-29* mutants, *n836* (lane 2), *n333* (lane 3) and *n1440* (lane 5), was analyzed by northern blot. Lanes 1-3 and lanes 4 and 5 are from two separate experiments. In each case, the upper autoradiogram is an exposure of a filter hybridized with the 1.1 kb *Eco*RI fragment of *lin-29*, and the lower autoradiogram is the same filter rehybridized with an *act-1* probe (Files et al., 1983) as a control for amounts of RNA loaded. (C) RNA was isolated from staged populations as indicated and analyzed by northern hybridization using the 1.1 kb *Eco*RI fragment as probe. E, RNA isolated from mixed-stage eggs; L1-L4, the four larval stages; yA, young adults containing no or just a few eggs; oA, older gravid adults. The lower autoradiogram is the same filter hybridized with an *act-1* probe (Files et al., 1983).

lin-29(n1440). Since these mutant lesions were shown to define at least a portion of the *lin-29* gene by intragenic recombination experiments (Papp et al., 1991), genomic DNA from this interval was used to isolate cDNAs (see Materials and Methods). A single class of cDNAs was isolated, which mapped to genomic DNA spanning the allele-specific polymorphisms associated with *lin-29(n836)* and *lin-29(n1440)*, and a point mutation associated with *lin-29(n546)* (see below), indicating that they represent *lin-29* transcripts.

The *lin-29* cDNAs detected two poly(A$^+$) transcripts of 2.4 and 1.8 kb in RNA blot experiments (lin-29A and lin-29B, respectively; Fig. 3A). Additional blots hybridized with radiolabeled strand-specific probes indicate these two *lin-29* transcripts are transcribed from the same DNA strand, left to right as diagrammed in Fig. 2 (not shown). The longest isolated *lin-29* cDNA clone, pcBB71, is essentially full-length and corresponds to the 2.4 kb lin-29A transcript. The structure of the lin-29A transcript was deduced from a combination of restriction mapping and sequencing of many independent cDNA clones, and the sequencing of the corresponding genomic DNA and junction fragments (Fig. 2; see Materials and Methods). The lin-29A transcription unit contains 11 exons and spans approximately 17.5 kb of genomic DNA.

The DNA and deduced amino acid sequences corresponding to the lin-29A transcript are shown in Fig. 4. The nucleotide sequence contains an open reading frame that begins at the first ATG of the cDNA and is predicted to encode a 50×10^3 M_r protein. The open reading frame is flanked by a 228 bp 5′ untranslated region and a 700 bp 3′ untranslated region. The 5′-most 9 bp of this cDNA (boxed in Fig. 4) diverge from the genomic sequence and are a perfect match to the terminal 9 bp of the 22 bp *trans*-splice leader SL1, suggesting that *lin-29* is *trans*-spliced in vivo to SL1 (Krause and Hirsh, 1987). The position at which the cDNA and genomic sequences diverge coincides with a 3′ splice-acceptor consensus sequence (Emmons, 1988). With the additional 13 bp of SL1 taken into account, this 5′ end is in close agreement with primer extension results (not shown). The *lin-29* 3′ untranslated region lacks a canonical polyadenylation signal (AAUAAA), but it contains the sequence UAUAAA 16 nt before the polyadenylation site. UAUAAA is found at approximately the same position in numerous other *C. elegans* genes (T. Blumenthal, personal communication) and may serve as a polyadenylation signal.

Restriction mapping and sequence analysis of cDNA clones from independent libraries revealed no internal differences that would account for the size differences between the lin-29A and lin-29B transcripts, suggesting that they are not generated by alternative splicing. The site of poly [A] addition was the same in every polyadenylated *lin-29* cDNA examined, from a total of four independent cDNA libraries, suggesting that the lin-29A and lin-29B transcripts have the same 3′ end. In addition, Northern blots hybridized with various fragments of the largest cDNA, pcBB71, suggest that the two transcripts differ at their 5′ ends (Fig. 3A). A restriction fragment containing the 5′-most 170 bp of pcBB71 preferentially hybridized to the 2.4 kb species (Fig. 3A), as does a 5′ 380 bp fragment that contains Exons 1 and 2 (not shown). In contrast, fragments derived from the middle portion or the 3′ end hybridized to both messages with similar intensities (Fig. 3A). Taken together, these experiments suggested that the lin-29B

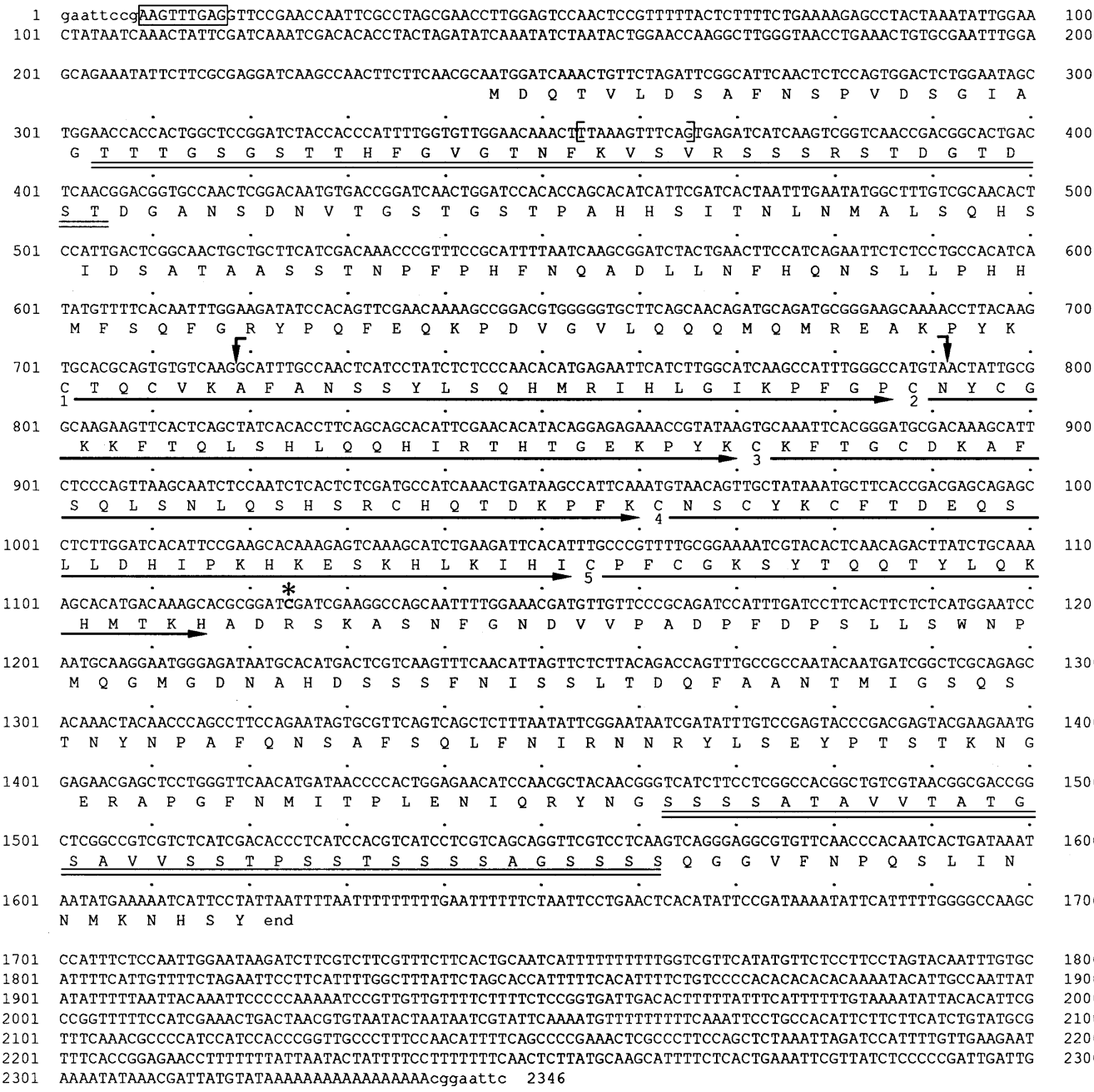


Fig. 4. Nucleotide sequence and conceptual translation of *lin-29* cDNA. The sequence of the coding strand of a full-length *lin-29* cDNA is shown, along with the predicted translation of the longest open reading frame. The first ATG encountered in the cDNA sequence is assumed to be the start codon. Stop codons are present in all three reading frames upstream of this ATG. The five zinc fingers are underlined with numbered arrows (as in Figs 2, 5), extending through the H/C link. Two serine/threonine-rich domains are indicated by double underlining. The 5′ terminal 9 bp that diverge from genomic sequence and match the 3′ end of the trans-splice leader SL1 are boxed. All cDNAs isolated were colinear with this sequence except that the 12 bp indicated by brackets were present only in one partial cDNA clone. In the full-length cDNA pcBB71, these 12 bp were spliced out as part of the second intron, resulting in the amino acids FKVSV being replaced by a leucine residue. The position of the C-to-T point mutation in *lin-29(n546)* is indicated by an asterisk. The borders of the *lin-29(n836)* deletion are indicated by vertical arrows. The 5′ border of the *lin-29(n836)* deletion is shown here at an exon junction, but resides within the intron sequence (see Fig. 2). Vector sequences are shown in lower case.

transcript begins at or near the start of Exon 5 of lin-29A. Either of two in-frame ATG codons within this 90 bp exon (beginning at nucleotide position 671 and 677; Fig. 4) could serve as an initiation codon for a smaller protein. We have identified two cDNAs that are consistent with this predicted 5′ end for lin-29B transcripts. One cDNA contains the terminal 11 bp of SL1 *trans*-spliced to the splice acceptor site of Exon V. The second cDNA contains a unique *cis*-spliced 37 bp exon at its 5′ end (see Materials and Methods). This exon originates approximately 2.8 kb upstream in the large

fourth intron of lin-29A (see Fig. 2). The relative abundance of these two transcripts in vivo is not known, but each transcript would encode a large open reading frame with the ATG at nucleotide position 671 of lin-29A likely serving as the translation initiation codon (Fig. 4).

Three *lin-29* mutations map to the transcribed region

The cDNA sequence and predicted open reading frame shown in Fig. 4 are likely to correspond to a functional *lin-29* gene product, because three of four examined *lin-29* alleles (judged to be nulls by genetic criteria) dramatically affect the abundance of *lin-29* transcripts (Fig. 3B) and/or alter the structure of a predicted *lin-29* protein (Figs 2, 4).

Southern blot analysis of genomic *lin-29(n546)* DNA identified a *Cla*I RFLP located within the 2.4 kb *Eco*RI fragment (Fig. 2). Sequence analysis of this region of *lin-29(n546)* DNA revealed a single C to T transition at nucleotide position 1124 (Fig. 4). This point mutation introduces a premature opal stop codon that eliminates the C-terminal 1/3 of the protein.

The *lin-29(n836)* lesion was identified as a deletion that removes the *Eco*RI site between the 2.4 and 7.3 kb fragments (Fig. 2; Papp et al., 1991). Sequencing of *lin-29(n836)* genomic DNA identified a lesion consisting primarily of a 320 bp deletion, that removes bases 719 through 791 relative to the cDNA sequence (Fig. 4). The 5′ border of this deletion resides within the fifth intron (see Fig. 2) and removes the splice acceptor site at position 719. The sequence of the deletion junction is 5′-aatgtcaCTATTTTactattg-3′, where the lower case letters represent nucleotides that match the wild-type sequence. The deletion resulted in the addition of 7 bp (capital letters), five of which contribute to a 6 bp direct repeat (underlined). In RNA of *lin-29(n836)* mutants, transcripts of 2.4, 2.0 and 1.8 kb are detected (Fig. 3B). Since the *lin-29(n836)* lesion is a deletion spanning a splice acceptor site, it is likely that the transcripts observed at 2.4 and 1.8 kb are the result of splicing to cryptic splice acceptor sites.

The *lin-29(n1440)* lesion maps to the 7.3 kb *Eco*RI fragment (Papp et al., 1991; see Fig. 2), within a large intron. This mutation increases the size of the *Eco*RI fragment to 9.2 kb, but its molecular nature is unknown. The *lin-29(n1440)* mutation has a dramatic effect on *lin-29* mRNA levels. In RNA isolated from *lin-29(n1440)* animals, the 2.4 kb mRNA is not detected and the smaller mRNA appears to be reduced in abundance relative to wild-type (Fig. 3B).

Finally, the two transcripts appear to be unaltered in *lin-29(n333)* mutant worms (Fig. 3B). We have not detected any alterations in genomic *lin-29(n333)* DNA by Southern analysis, suggesting this EMS-induced lesion is a point mutation or small rearrangement.

lin-29 encodes a zinc finger protein

The *lin-29* deduced amino acid sequence shown in Fig. 4 was compared with protein sequences in available databases (see Materials and Methods). These studies revealed the presence of the $(Cys)_2$-$(His)_2$ zinc finger nucleic acid binding motif of the Krüppel class (Schuh et al., 1986; reviewed in Klug and Rhodes, 1987). Zinc finger proteins of this class usually contain multiple copies of the consensus Cys-$X_{(2-4)}$-Cys-X_3-Phe-X_5-Leu-X_2-His-X_3-His separated by the so-called 'H/C link' sequence of consensus Thr-Gly-Glu-Lys/Arg-Pro-Tyr/Phe, where X = any amino acid (Schuh et al., 1986). The *lin-29* open reading frame is predicted to encode five consecutive $(Cys)_2$-$(His)_2$ zinc fingers. The amino acid sequence of the five *lin-29* zinc fingers is aligned with the consensus sequence in Fig. 5A. With the exception of a conservative change of Phe to Tyr in the fifth finger, all five Lin-29 fingers match the consensus exactly, and all five possess a Lys residue at position 8. Other positions show amino acid conservation in at least three of the five fingers. The H/C link region is less well conserved; only the link between fingers 2 and 3 shows a complete match to the consensus sequence, and two of the links contain additional amino acids (Fig. 5A). Limited homology to the H/C link is also present just upstream of the first zinc finger (---KPY-; Fig. 4). The lin-29B transcript is predicted to encode a smaller version of this protein containing all five zinc fingers, but lacking approximately 140 amino acids from the N terminus.

Of the $(Cys)_2$-$(His)_2$ zinc finger proteins identified to date, the *lin-29* zinc fingers are most similar to a predicted product of the human cDNA clone hbc087 (Takeda et al., 1993). Both Lin-29 and the predicted hbc087 protein contain five zinc fingers, and significantly, the individual fingers match best when aligned in order (Fig. 5B). The similarity does not extend outside of the finger domain in the limited regions of the hbc087 sequence that we have analyzed so far (see Materials and Methods). The function of the predicted hbc087 protein is currently unknown.

When amino acid sequences outside of the Lin-29 zinc finger region were used in database searches, no obvious similarities were found but two transcription factor hallmarks were noted (Ptashne, 1988). First, the predicted protein is very rich in serine (14.8%) and threonine (8.1%), with one stretch of 34 amino acids near the N terminus containing 50% Ser/Thr and a second 34 amino acid stretch near the C terminus reaching 65% (Fig. 4). Second, a short glutamine-rich region (7 of 21 amino acids) is located just upstream of the zinc fingers.

Developmental analysis of *lin-29* expression

Poly(A)$^+$ RNA was isolated from synchronous *C. elegans* populations at various developmental stages and analyzed by RNA blotting (Ausubel et al., 1989). A probe generated from the 1.1 kb *Eco*RI fragment of the *lin-29* cDNA (see map, Fig. 3C) detects both the 2.4 and 1.8 kb transcripts beginning in the L1 stage. These transcripts increase in abundance as development proceeds to the L4 stage, and then decrease at least 5-fold relative to the L4 levels in adults. The abundance of the 1.8 kb transcript increases slightly relative to the 2.4 kb transcript during larval development.

Lin-29 binds in vitro to collagen gene regulatory DNA

Since *lin-29* mutations affect stage-specific cuticle morphology, genes encoding cuticle components that are regulated in a stage-specific manner are candidate targets of *lin-29* activity. Indeed, adult-specific transcription of the collagen gene *col-19* (Cox and Hirsh, 1985) is temporally regulated by the heterochronic gene pathway (Liu and Ambros, 1991). A *col-19:lacZ* reporter gene has been shown to be regulated in vivo by the genes of the heterochronic pathway, including *lin-29*. In wild-type worms, a *col-19:lacZ* fusion is specifically expressed in hypodermal cells beginning at the L4 molt. The fusion is not

A

```
                       Finger                      H/C
Consensus C----C---F-----L--H---HTGE[K/R]P[F/Y]-

Finger 1  CTQ..CVKAFANSSYLSQHMRIHLGIKPFGP
Finger 2  CNY..CGKKFTQLSHLQQHIRTHTGEKPYK
Finger 3  CKFTGCDKAFSQLSNLQSHSRCHQTDKPFK
Finger 4  CNS..CYKCFTDEQSLLDHIPKHKESKHLKIHI
Finger 5  CPF..CGKSYTQQTYLQKHMTKH
```

B

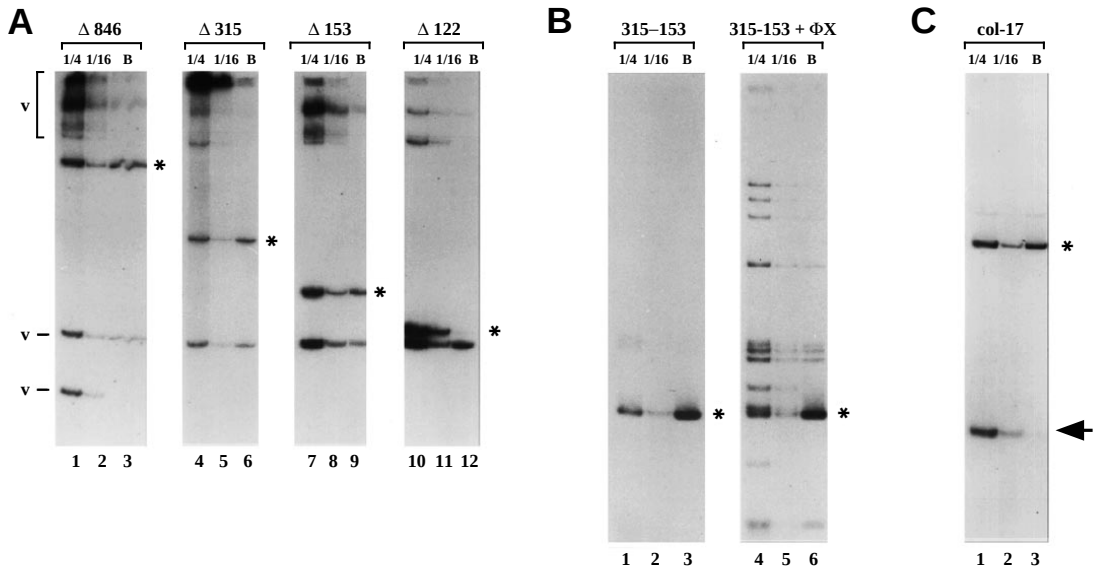
```
Finger 1: lin-29  CTQ..CVKAFANSSYLSQHMRIHLGIKPFGP
                  |    | | ||||||| || ||| | ||
          hbc087  CPH..CSKTFANSSYLAQHIRIHSGAKPYS.

Finger 2: lin-29  CNY..CGKKFTQLSHLQQHIRTHTGEKPYK
                  ||   | | | |||||||| | |||  |||
          hbc087  CNF..CEKSFRQLSHLQQHTRIHTGDRPYK

Finger 3: lin-29  CKFTGCDKAFSQLSNLQSHSRCHQTDKPFK
                  |   || ||| |||||||| | |  |||||
          hbc087  CAHPGCEKAFTQLSNLQSHRRQHNKDKPFK

Finger 4: lin-29  CNS..CYKCFTDEQSLLDHIPKHKESKHLKIHI
                  |    |    ||  ||  |    |   || |
          hbc087  CHN..CHRAYTDAASLEVHLSTH.TVKHAKVYT

Finger 5: lin-29  CPF..CGKSYTQQTYLQKHMTKH
                  |    |   ||  ||| ||| ||
          hbc087  CTI..CSRAYTSETYLMKHMRKH

Zinc finger consensus: C----C---F-----L--H---HTGE[K/R]P[F/Y]-
```

Fig. 5. Comparison of the five *lin-29* zinc fingers with the $(Cys)_2$-$(His)_2$ zinc finger consensus. (A) The consensus sequences for the $(Cys)_2$-$(His)_2$ zinc finger DNA binding motif and for the 'H/C link' sequence that joins adjacent zinc fingers (Berg, 1993; Schuh et al., 1986) are aligned with the five *lin-29* zinc finger domains (numbered as in Figs 2, 4). Identical residues are highlighted in bold. Residues that occur in at least three of the five *lin-29* fingers are outlined. (B) The amino acid sequence of the Lin-29 zinc finger motifs are aligned with those of the zinc fingers encoded by the human cDNA clone hcb087 (Takeda et al., 1993). Identical amino acids are indicated by vertical lines. In each figure, the dashes in the consensus represent any amino acid and the periods indicate gaps in the sequence.

expressed in hypodermal cells of *lin-29* mutant animals at any stage, whereas it is expressed precociously, beginning during an abnormally early molt, in *lin-14*(loss-of function) mutants (Liu and Ambros, 1991; Liu et al., 1995).

To determine if Lin-29 can regulate *col-19* directly, we assayed in vitro for binding of a Lin-29 protein to *col-19* promoter sequences. The portion of the *col-19* gene tested was a control region sufficient to program adult-specific expression of a *lacZ* reporter gene in vivo (Liu et al., 1995). A *glutathione-S-transferase:lin-29* fusion gene was constructed, the chimeric protein expressed in *E. coli* and purified by absorption to glutathione-coupled agarose beads (Smith and Johnson, 1988). The fusion protein/agarose complexes were incubated with end-labeled DNA restriction fragments derived from *col-19:lacZ* fusion constructs as described in Materials and Methods. DNA fragments bound to the fusion protein were purified and analyzed by polyacrylamide gel electrophoresis relative to dilutions of input DNA. The *lin-29* fusion protein binds to an 872 bp fragment extending from +26 to −846 relative to the *col-19* ATG (Fig. 6A, lane 3). This fragment is sufficient to program adult-specific *col-19:lacZ* expression in vivo (Liu et al., 1995). In these experiments, a high percentage of an input fragment that binds specifically to the fusion protein is recovered in the bound fraction (typically around 25%). A 300 bp fragment derived from *lacZ* also binds the fusion protein, whereas a smaller 181 bp vector fragment, and several larger vector and *lacZ* fragments do not bind. DNA binding requires the *lin-29* portion of the fusion protein, since no fragments were recovered from extracts containing the glutathione S-transferase alone or with a Gst:Lin-29 fusion protein lacking the zinc finger domain (not shown).

The *lin-29* fusion protein binding site(s) were further localized relative to the *col-19* sequences necessary for in vivo expression. 5′ deletions of the *col-19:lacZ* fusion that retain 315, 153 and 122 bp (Δ315, Δ153, and Δ122, respectively) upstream of the ATG have been examined for their ability to program *lacZ* expression in transgenic worms (Liu et al., 1995). Of these, only Δ315 retains the ability to program detectable *lacZ* expression in transgenic worms (see Fig. 6D). The binding of *lin-29* fusion protein to these deletions was examined. Both Δ315 and Δ153 bind to the fusion protein in vitro, whereas Δ122 does not (Fig. 6A, lanes 6, 9, and 12). These results suggest that the 31 bp interval between −122 and −153 contains a *lin-29* binding site(s), but this site alone is insufficient to program *col-19:lacZ* expression in vivo. The 162 bp region of the *col-19* gene from −122 to −153 is required for *col-19* expression in vivo. To test if the requirement for this interval could be due to additional *lin-29* binding sites, the 162 bp fragment was synthesized precisely using PCR amplification (see Materials and Methods) and its ability to bind to the fusion protein was assayed. This fragment binds to the *lin-29* fusion protein (Fig. 6B, lane 3), and furthermore, it is preferentially recovered from binding reactions when mixed with end-labeled *Hae*III-digested ϕX174 DNA (Fig. 6B, lane 6).



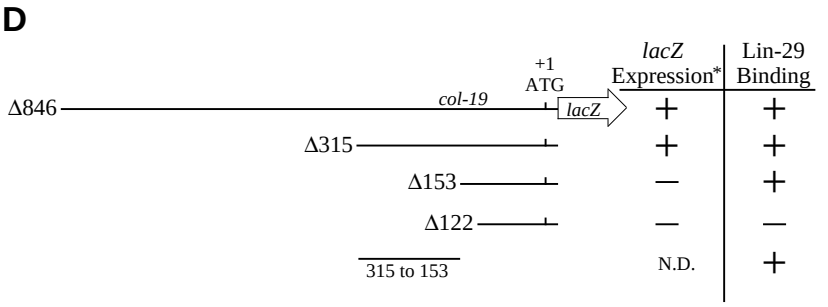


Fig. 6. In vitro binding of *lin-29* fusion protein to collagen gene sequences. Radiolabeled DNAs were incubated with Gst:Lin-29 absorbed to glutathione-coupled agarose beads. The bound DNAs were purified and analyzed by acrylamide gel electrophoresis and autoradiography (see Materials and Methods). (A) In vitro binding of the Gst:Lin-29 fusion protein to *col-19* sequences. In each panel, the first two lanes contain 1/4 or 1/16 of the input DNA as indicated. The third lane, marked B, shows the DNA that bound to the fusion protein. Asterisks indicate the *col-19* promoter fragment. 'v' indicates vector fragments. A series of *col-19* 5′ deletions in a *lacZ* vector were used for these binding experiments. In lanes 1-3, a *Hin*dIII-*Bam*HI fragment of *col-19* extending from +26 to −846 relative to the ATG was used (Δ846). The bound material in lane 3 is spread over the width of two lanes. The smallest (181 bp) vector fragment seen in lanes 1-3 and that does not bind Gst:Lin-29, is contained on the *col-19* fragment in the digests used in the remaining three panels. Lanes 4-6 are analysis of Δ315. Lanes 7-9 are analysis of Δ153. Lanes 10-12 are analysis of Δ122. (B) In vitro binding of the Gst:Lin-29 fusion protein to the 162 bp fragment extending from bp −153 to −315. Lanes 1-3 show results of binding to the PCR-amplified 162 bp fragment alone. Lanes 4-6 show results of a mixing experiment in which the 162 bp fragment was mixed at an equi-molar ratio with end-labeled ϕX174 DNA. In each panel, the first two lanes show 1/4 and 1/16 of input DNA repectively, and the third lane, (B), shows the DNA that bound to the fusion protein. The asterisk indicates the 162 bp fragment. (C) Binding of the Gst:Lin-29 fusion protein to *col-17* sequences. A 1096 bp *Hin*dIII-*Bam*HI fragment containing *col-17* promoter sequences from nucleotide −806 to +290 relative to the ATG was gel-purified, end-labeled and digested with *Xmn*I and used in a binding experiment. Two fragments result from this digestion, an 849 bp fragment extending from −806 to +43 (asterisk), and a 247 bp fragment from +44 to +290 (arrow). Lanes 1 and 2 are 1/4 and 1/16 of the input DNA. Lane 3 shows the DNA that bound to the fusion protein. (D) The results of in vitro binding of Gst:Lin-29 to DNA fragments from the *col-19* promoter are summarized along with data indicating the ability of each *col-19* fragment to program *lacZ* expression in vivo (Liu et al., 1995). N.D. = not determined.

Thus, at least one additional *lin-29* binding site is present within this 162 bp interval, indicating that there are at least two *lin-29* binding sites in the upstream regulatory sequences of *col-19*, one of which resides within a region required for *col-19* expression in vivo.

In addition to the adult-specific activation of collagen genes at the L4 molt, there is adult-specific repression of other collagen genes, including *col-17* (Cox and Hirsh, 1985). *col-17* transcripts accumulate only during larval stages in wild-type animals, but in *lin-29* mutants *col-17* transcripts continue to accumulate in the adult stage (Liu et al., submitted). To investigate a possible the role for *lin-29* in regulating *col-17* expression, binding of the *lin-29* fusion protein to *col-17* sequences was tested in vitro. The *lin-29* fusion protein binds to an 849 bp fragment extending from +43 to −806 relative to the ATG, but not to a 247 bp fragment extending from +44 to +290 bp downstream of the ATG (Fig. 6C).

DISCUSSION

lin-29 regulates the stage-specific transcription of a terminal differentiation gene

The two *lin-29* transcripts encode proteins with 5 tandem $(Cys)_2$-$(His)_2$ zinc finger motifs of the type that were first described in *Xenopus* transcription factor IIIA (Miller et al., 1985; see Fig. 4A). Several proteins containing the $(Cys)_2$-$(His)_2$ type of zinc finger motif have been shown to regulate gene expression through the direct binding to DNA (see Evans and Hollenberg, 1988). Thus, *lin-29* likely controls the switch to the terminally differentiated adult state of lateral hypodermal seam cells by directly regulating the transcription of other genes.

The terminal differentiation of seam cells at the L4-to-adult molt involves several apparently distinct cellular behaviors that are all coordinated by *lin-29*. These include cell cycle exit, cell fusion and the stage-specific regulation of genes expressed in

the lateral hypodermis, including collagens. *lin-29* could control these processes directly by activating or repressing the expression of genes encoding terminal differentiation gene products, such as cuticle structural proteins, components of the cell cycle machinery, and proteins that implement cell fusion and seam cell morphology. At the other extreme, *lin-29* could indirectly regulate the transcription of such terminal differentiation genes by controlling the expression of intermediate regulatory genes. Our finding that *lin-29* protein binds in vitro to a DNA regulatory region necessary for in vivo adult-specific activation of the collagen gene *col-19* suggests a direct role for *lin-29* in the transcriptional control of at least one terminal differentiation gene product, and possibly others. We expect that targets of Lin-29 will include genes encoding more cuticle components, and genes encoding proteins with critical roles in the control of cell cycle, cell fusion and molting, since all of these processes are misregulated in *lin-29* mutants.

We believe that the *lin-29* mutant phenotype (Ambros and Horvitz, 1984) reflects the developmental consequence of complete, or nearly complete, loss of *lin-29* function. Although we cannot rule out the possibility of residual *lin-29* function in the three *lin-29* mutants whose lesions we have mapped here, they behave as nulls by the criteria that we can measure. Their phenotypes are not enhanced when placed in *trans* to a deficiency (Ambros and Horvitz, 1984), and one mutant, *lin-29(n1440)*, exhibits a dramatic reduction in *lin-29* message levels (Fig. 3). Still, the possibility exists that Lin-29 plays a role in worm development other than suggested by analysis of these mutant alleles. Analysis of the Lin-29 spatial distribution should address possible additional roles for Lin-29 during wild-type worm development.

Positive and negative transcriptional control by Lin-29

In addition to activating genes encoding components of the adult cuticle at the L4 molt, *lin-29* is also required for stage-specific transcriptional repression of genes utilized in synthesis of the larval cuticle. In wild-type animals, *col-17* mRNA accumulates during the three larval molts but is undetectable during the L4-to-adult molt. The repression of *col-17* during the final molt is dependent on *lin-29* activity (Liu et al., 1995). The finding that the *lin-29* fusion protein binds in vitro to *col-17* DNA sequences as well as *col-19* sequences suggests that *lin-29* may have a direct role in both the activation of *col-19* transcription and the repression of *col-17* transcription during the L4 molt. There is precedent for transcriptional regulatory proteins having both positive and negative regulatory capabilities, mediated, for example, by ancillary proteins (Diamond et al., 1990) or concentration (Sauer and Jäckle, 1993). A similar mechanism could be responsible for *lin-29* protein activating or repressing different collagen promoters. Alternatively, *col-17* and *col-19* could be regulated by different forms of the *lin-29* protein, or by varied placement of *lin-29* binding sites within their promoters. Comparison of the DNA sequence of the fragments of *col-17* and *col-19* that are bound by Lin-29 protein reveals an A/T rich sequence, but no obvious conserved Lin-29 binding motif.

Transcription factors that promote terminal differentiation

One role of proteins that promote terminal differentiation is to halt cell cycle progression. For example, expression of MyoD causes cell cycle arrest of proliferating cells in culture (Crescenzi et al., 1990), and in worms lacking *lin-29* activity, seam cells that would otherwise exit the cell cycle instead continue to divide. The effect of these genes on cell cycle progression could be direct, by transcriptional control of cell cycle regulators, or it could be an indirect consequence of the transcription of certain terminal differentiation genes (Rastinejad et al., 1993).

The mechanism by which *lin-29* promotes cell cycle exit in the hypodermal seam cells is of particular interest as an example of strict cell lineage-specific control of the cell cycle. Since *lin-29* encodes a zinc finger protein, it seems likely that transcriptional regulation must be involved in the regulation of cell division cycles in the lateral hypodermis. We do not know at what stage of the cell cycle the seam cells terminally differentiate, so we cannot predict what key cell cycle regulator(s) may mediate *lin-29* cell cycle regulatory activity. However, there is ample precedent for transcriptional control of cell cycle arrest, such as G2 restriction by limiting *string* transcription in *Drosophila* (Edgar and O'Farrell, 1989), or control of G1 progression by transcriptional control of cyclin genes (Ogas et al., 1991). It will be interesting to test whether ectopic expression of *lin-29* in non-seam lineages is sufficient to impose hypodermal cell characteristics and/or to cause cell cycle exit.

Temporal regulation of *lin-29* to control the L/A switch

lin-29 is the most downstream gene identified in the heterochronic gene pathway and is specifically required for the execution of the L/A switch during the L4 molt. The appearance of active Lin-29 protein during the L4 stage could be the critical event responsible for triggering lateral hypodermal cell terminal differentiation. Alternatively, Lin-29 protein could be present and active at earlier stages, and the stage-specific activation of a hypothetical cofactor could be responsible for triggering the L/A switch. In the absence of direct evidence for the hypothetical temporally regulated cofactor, it is reasonable to propose that *lin-29* activity increases at, or shortly before, the final molt in wild-type worms to trigger the hypodermal cell terminal differentiation. A critical test of this proposition will require the use of modified *lin-29* constructs in transgenic worms to determine whether expression of Lin-29 protein in seam cells at abnormally early developmental times is sufficient to trigger precocious activation of the L/A switch.

Epistasis analysis suggests that a genetic pathway of heterochronic genes is responsible for controlling the timing of *lin-29* activity. By genetic tests, *lin-14* and *lin-28* are negative regulators of the L/A switch (Ambros, 1989), and so could act by inhibiting *lin-29* activity during early larval stages. Lin-14 is expressed and acts primarily during the L1 stage (Ruvkun and Giusto, 1989; Ambros and Horvitz, 1987), suggesting that *lin-14* protein regulates *lin-29* indirectly. *lin-28* and other heterochronic genes that apparently act between *lin-14* and *lin-29* in the genetic hierarchy (Ambros 1989; Moss and Ambros, unpublished) could include regulators of *lin-29*, but their gene products and stage-specificity of action have not yet been characterized.

lin-29 does not seem to be regulated by a sharp L4-specific increase of *lin-29* mRNA levels. The abundance of the two *lin-29* transcripts increases from the L1 through the L4 stage, and

then decreases dramatically in adults (Fig. 3C). If *lin-29* mRNA accumulation is a key component of *lin-29* control, then a slight increase of *lin-29* mRNA from the L3 through the L4 stage would have to be sufficient to cause activation of *lin-29*. There is evidence from analysis of one *lin-29* allele that a threshhold level of *lin-29* activity may signal lateral hypodermal cell terminal differentiation (Hodgkin et al., 1989), supporting the idea that a small temporal change in the amount of active Lin-29 may trigger the L/A switch.

Although the initial appearance of the *lin-29* transcripts in the L1 stage roughly correlates with the disappearance of *lin-14* protein toward the end of the L1 stage, preliminary results indicate that *lin-29* transcripts still accumulate in *lin-14* (gain-of-function) and *lin-4* (loss-of-function) animals that never execute the L/A switch. Thus, we favor a model in which *lin-14* acts indirectly, via other heterochronic genes, to regulate *lin-29* activity, and that post-transcriptional controls are at least partially, and perhaps primarily, responsible for the temporal control of *lin-29* activity. Examination of the spatial and temporal accumulation of *lin-29* protein during development will allow us to begin distinguishing among the various possible mechanisms for *lin-29* control.

We thank Jeff Simon, Sue Euling, Rhonda Feinbaum, Bob Kingston and Eric Moss for helpful discussions and critical reading of this manuscript. We thank Abby Telfer and Diane Levitan for advice and encouragement. This work was supported by a Damon Runyon-Walter Winchell Cancer Research Fund Fellowship, DRG-1028 and National Science Foundation IBN-9305208, to AER, and by Public Health Services Grant GM34028 and American Cancer Society Grant NP479A to VA. The accession number for the sequence data presented in Fig. 4 is L39212.

REFERENCES

Ambros, V. and Horvitz, H. R. (1984). Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226**, 409-416.

Ambros, V. and Horvitz, H. R. (1987). The *lin-14* locus of *Caenorhabditiselegans* controls the time of expression of specific postembryonic developmental events. *Genes Dev.* **5**, 1813-1824.

Ambros, V. (1989). A hierarchy of regulatory genes controls a larval-to-adult switch in *C. elegans*. *Cell* **57**, 49-57.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. G. and Struhl, K. (1989). *Current Protocols in Molecular Biology*. New York: John Wiley & Sons.

Barstead, R. J. and Waterston, R. H. (1989). The basal component of the nematode dense-body is vinculin. *J. Biol. Chem.* **264**, 10177-10185.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.

Cox, G. N. and Hirsh, D. (1985). Stage-specific patterns of collagen gene expression during development of *Caenorhabditis elegans*. *Mol. Cell. Biol.* **5**, 363-372.

Crescenzi, M., Fleming, T. P., Lassar, A. B., Weintraub, H. and Aaronson, S. A. (1990). MyoD induces growth arrest independent of differentiation in normal and transformed cells. *Proc. Natl. Acad. Sci. USA* **87**, 8442-8446.

Desplan, C., Theis, J. and O'Farrell, P. H. (1985). The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. *Nature* **318**, 630-635.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395.

Diamond, M. I., Miner, J. N., Yoshinaga, S. K. and Yamamoto, K. R. (1990). Transcription factor interactions: Selectors of positive or negative regulation from a single DNA element. *Science* **249**, 1266-1272.

Edgar, B. A. and O'Farrell, P. H. (1989). Genetic control of cell divisions patterns in the *Drosophila* embryo. *Cell* **57**, 177-187.

Emmons, S. (1988). 'The Genome'. In *The Nematode Caenorhabditis elegans*. W. B. Wood. ed. Cold Spring Harbor Monograph Series 17. Cold Spring Harbor: Cold Spring Harbor University Press.

Evans, R. M. and Hollenberg, S. M. (1988). Zinc fingers: gilt by association. *Cell* **52**, 1-3.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction fragments to high specificity. *Anal. Biochem.* **132**, 3-16.

Files, J. G., Carr, S. and Hirsh, D. (1983). Actin gene family of *Caenorhabditis elegans*. *J. Mol. Biol.* **164**, 355-375.

Gyllensten, U. B. and Erlich, H. A. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. *Proc. Natl. Acad. Sci. USA* **85**, 7652-7656.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V. and Anderson, P. (1989). A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* **123**, 301-313.

Klug, A. and Rhodes, D. (1987). 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**, 464-469.

Kraft, R., Tardiff, J., Krauter, K. S. and Leinwand, L. A. (1988). Using mini-prep plasmid DNA for sequenceing double stranded templates with sequenase. *Biotechniques* **6**, 544-546.

Krause, M. and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell* **49**, 753-761.

Liu, Z. and Ambros, V. (1989). Heterochronic genes control the stage-specific initiation and expression of the dauer larval developmental program in *Caenorhabditis elegans*. *Genes Dev.* **3**, 2039-2049.

Liu, Z. and Ambros, V. (1991). Alternative temporal control systems for hypodermal cell differentiation in *Caenorhabditis elegans*. *Nature* **350**, 162-165.

Liu, Z., Kirch, S. and Ambros, V. (1995). The *Caenorhabditis elegans* heterochronic gene pathway controls stage-specific transcription of collagen genes. *Development* **121**, 2471-2478.

Miller, J., McLachlan, A.D. and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609-1614.

Ogas, J., Andrews, B.J. and Herskowitz, I. (1991). Transcriptional activation of CLN1, CLN2, and a putative new G1 cyclin (HCS26) by SWI4, a positive regulator of G1-specific transcription. *Cell* **66**, 1015-1026.

Papp, A., Rougvie, A. E. and Ambros, V. (1991). Molecular cloning of *lin-29*, a heterochronic gene required for the differentiation of hypodermal cells and the cessation of molting in *C. elegans*. *Nucl. Acids Res.* **19**, 623-630.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. *Nature* **335**, 683-689.

Rastinejad, F., Conboy, M. J., Rando, T. A. and Blau, H. M. (1993). Tumor suppression by RNA from the 3′ untranslated region of α-tropomyosin. *Cell* **75**, 1107-1117.

Ruvkun, G. and Giusto, J. (1989). The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch. *Nature* **338**, 313-319.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Sauer, F. and Jäckle, H. (1993). Dimerization and the control of transcription by *Krüppel*. *Nature* **364**, 454-547.

Schuh, R., Aicher, W., Gaul, U., Côté, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R. and Jäckle, H. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Krüppel*, a *Drosophila* segmentation gene. *Cell* **47**, 1025-1032.

Smith, D. B. and Johnson, K. S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31-40.

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

Takeda, J., Yano, H., Eng, S., Zeng, Y. and Bell, G. I. (1993). A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones. *Hum. Mol. Genet.* **2**, 1793-1798.

Wood, W. B. (1988). (ed.)The nematode *Caenorhabditis elegans*. *Cold Spring Harbor Monograph Series 17*. Cold Spring Harbor: Cold Spring Harbor University Press.

(Accepted 20 April 1995)