Development 126, 2483-2494 (1999)
Printed in Great Britain © The Company of Biologists Limited 1999
DEV7716

The *Caenorhabditis elegans* genes *egl-27* and *egr-1* are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning

Florence Solari[1,*], Alex Bateman[2] and Julie Ahringer[1,*,‡]

[1]University of Cambridge, Department of Genetics, Downing Street, Cambridge CB2 3EH, UK
[2]The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
*Present address: Wellcome CRC Institute, Tennis Court Road, Cambridge CB2 1QR, UK
‡Author for correspondence (e-mail: jaa@mole.bio.cam.ac.uk)

Accepted 19 March; published on WWW 4 May 1999

SUMMARY

We show here that two functionally redundant *Caenorhabditis elegans* genes, *egl-27* and *egr-1*, have a fundamental role in embryonic patterning. When both are inactivated, cells in essentially all regions of the embryo fail to be properly organised. Tissue determination and differentiation are unaffected and many zygotic patterning genes are expressed normally, including HOX genes. However, *hlh-8*, a target of the HOX gene *mab-5*, is not expressed. *egl-27* and *egr-1* are members of a gene family that includes MTA1, a human gene with elevated expression in metastatic carcinomas. MTA1 is a component of a protein complex with histone deacetylase and nucleosome remodelling activities. We propose that EGL-27 and EGR-1 function as part of a chromatin regulatory complex required for the function of regional patterning genes.

Key words: SANT, ELM1, ELM2, *vab-7*, NURD, Histone deacetylation, Nucleosome remodelling, *Caenorhabditis*

INTRODUCTION

During embryonic development, cells are determined, differentiate and become organised into tissues. These events are regulated by extrinsic and intrinsic signals that are often translated into changes in the expression and/or activity of transcription factors. Therefore, how transcription factor activities are regulated is important for understanding most developmental processes. For example, the HOX genes encode transcription factors that pattern tissues along the anteroposterior axis of most or all animals (reviewed in McGinnis and Krumlauf, 1992). Recently, the identification of co-factors that act with HOX proteins have given insights into the mechanism by which they regulate their targets (for recent reviews, see Fisher and Caudy, 1998; Mann and Affolter, 1998). However, how developmental transcription factors act in the context of chromatin is still largely unknown.

Recent studies have revealed that a gene's transcriptional activity can be regulated through chromatin modifications: hyperacetylated chromatin is generally associated with active genes, whereas hypoacetylated chromatin is associated with repressed genes. Protein complexes that carry out ATP-dependant nucleosome remodelling also affect the activity of particular promoters. Multiprotein complexes with the above activities can interact with sequence-specific transcription factors, suggesting a model whereby the transcriptional activity of a promoter is regulated by locally perturbing chromatin structure (for recent reviews, see Kadonaga, 1998; Struhl, 1998).

Here we show that the *C. elegans* gene *egl-27* encodes a protein similar to MTA1, a component of a protein complex with ATP-dependent nucleosome remodelling and histone deacetylation activities (Xue et al., 1998). No previous members of this gene family have been identified by mutation. The *C. elegans* genome contains a related gene, *egr-1* that has a redundant role with *egl-27*; in embryos where both are inactivated, the pattern of cells in all regions of the embryo is abnormal. Furthermore, although HOX genes have normal expression, a HOX gene target gene fails to be expressed. We suggest that protein complexes containing EGL-27 family members regulate the activity of transcription factors involved in embryonic patterning.

MATERIALS AND METHODS

Strains and alleles

Cultivation of *C. elegans* was as in Brenner (1974). The following mutant alleles and deficiencies were used:

LG II: *bli-2(e768)*, *dpy-10(e128)*, *egl-27(e2394)*, *egl-27(n170)*, *egl-27(we3)*, *unc-4(e120)*, *mnDf30*, *mnDf96*, *mnDf39*, *mnDf68*, *mnDf61*, *mnDf88* and *maDf4*.

LGIII: *vab-7(ed6)*, *unc-49(e362)* and *vab-7(ed6)*.

The following strains were used as markers in expression studies: *ayIs7* [*hlh-8::gfp*] IV (Harfe et al., 1998), *eIs24* [*vab-7::lacZ*] II

(Ahringer, 1996), *muIs6* [*lin-39::lacZ + rol-6(su1006)*] (Wang et al, 1993), *muIs13* [*egl-5::lacZ + rol-6(su1006)*] V (kindly provided by C. Kenyon), *muIs16* [*mab-5::gfp + dpy-20(+)*] II; *dpy-20(e1282)* IV (kindly provided by C. Kenyon), *weEx20* [*lin-44::gfp + hsp16-2-vab-7 + hsp16-41-vab-7 + rol-6(su1006)*] (F. S. and J. A., unpublished; *lin-44::gfp* plasmid (pMHE004) kindly provided by M. Herman), *wIs1* [*SCM::lacZ + rol-6(su1006)*] (Terns et al., 1997), *hlh-1::gfp* (PD7963; K. Dej, S. Xu and A. Fire, personal communication.).

Isolation and mapping of *we3*

unc-49(e362) vab-7(ed6) animals were mutagenized using EMS as in Wood (1988). 484 F_1s were cloned (representing 968 chromosomes) and their progeny scored for stronger posterior defects. 5 enhanced strains were isolated, the strongest of which contained *we3*. *we3* was mapped to position +0.09 on chromosome II: from the strain + *we3* + /*dpy-10(e128)* + *unc-4(e120)*, 23/24 Unc non-Dpy recombinants carried *we3* and 1/15 Dpy non-Unc recombinants carried *we3*. The deficiencies *mnDf96* and *mnDf30* remove *we3*; *mnDf39*, *mnDf68*, *mnDf61*, *mnDf88* and *maDf4* do not remove *we3*. *we3* is maternally rescued: whereas 92% of *we3* mutants descended from *we3* homozygotes have abnormal body morphology at 15°C, only 4% of homozygotes are abnormal if they inherit wild-type maternal product.

When grown as a homozygote, *we3* behaves as a strong loss-of-function mutation: *we3* mutants are 89.5% lethal at 15°C. *we3*/*mnDf30* animals are 78% lethal at 15˚C (*n*=118) when they inherit wild-type maternal product, but 88% (*n*=84) lethal at 15°C when the mother was *we3*/*mnDf30*. Unless otherwise stated, all analyses of *we3* were done at 15°C.

egl-27(we3) enhances the phenotype of *vab-7(ed6)* zygotically. *vab-7(ed6)* animals range from having nearly wild-type morphology (31%) to having blunt, round tail regions (16%). Zygotic enhancement was tested using *unc-4(e120)* as a marker for *egl-27(we3)*: Only 1% of *egl-27(we3) unc-4(e120); vab-7(ed6)* that come from *egl-27(we3) unc-4(e120)* /+ +*; vab-7(ed6)* appear wild type and 50% have blunt tail regions. The progeny of the next generation have slightly stronger defects than *egl-27(we3)* single mutants (not shown). Only 4% (1/27) of *egl-27(we3) unc-4(e120)* that come from *egl-27(we3) unc-4(e120)* /+ + had posterior defects.

Complementation tests

Phasmid dye filling was assayed in adult hermaphrodites as in Herman and Horvitz (1994). 0% (0/532) of wild-type hermaphrodites have defective phasmid dye filling, whereas 100% of phasmids from *egl-27*(*e2394*) (39/39) or *egl-27*(*n170*) (51/51) fail to fill with dye. At 15°C, phasmids fail to fill in 66% of surviving *we3* mutants. Male tail morphology was assayed using DIC optics: 0/20 wild-type, 8/12 *egl-27*(*n170*) and 3/7 surviving *we3* males have abnormal tail morphology.

we3 was found to be an allele of *egl-27* by complementation tests: *egl-27*(*e2394*) or *egl-27*(*n170*) hermaphrodites were crossed with *we3 unc-4(e120)*/++ males and assayed for dye-filling defects (Dyf), abnormal male tail morphology (Mtl) and abnormal body morphology (Mor). In the *egl-27*(*e2394*) cross, 9/15 hermaphrodites were Dyf, 16/41 males were Mtl and no Mor animals were found. Progeny testing confirmed that one-half were *we3 unc-4(e120)*/*egl-27*(*e2394*). In the *egl-27*(*n170*) cross, 4/8 hermaphrodites were Dyf, 18/39 males were Mtl and no animals were Mor. Therefore, *we3* fails to complement *egl-27*(*n170*) and *egl-27*(*e2394*) for Dyf and Mtl, but not for Mor in this test. However, *we3* did fail to complement *egl-27*(*n170*) for Mor when the mother was *egl-27(we3)*. In the reciprocal cross of *we3 unc-4(e120)* hermaphrodites with *egl-27*(*n170*)/+ males: 21/47 non-unc hermaphrodites were Mor, 5/37 were Dyf and 8/61 males were Mtl. For comparison, in a cross of *we3 unc-4(e120)* with wild-type males, 24/144 were Mor, 0/80 non-unc hermaphrodites were Dyf and 0/30 males were Mtl.

The *egl-27*(*n170*) phasmid-filling defect has a maternal effect. Whereas 51/51 of *egl-27*(*n170*) homozygotes from homozygous mother have phasmids that fail to fill with dye, none out of 30 progeny from an *egl-27(n170)* /+ mother had phasmid dye-filling defects.

egl-27 cloning

egl-27 (we3) animals were injected with pools of cosmids covering the region where *we3* was mapped: T24B4+F41G3, F41G3+T13C2, T13C2+F31E8, F31E8+C04A2, C04A2+C44B7, C44B7+B0252, B0252+F22D3, F22D3+C15F1 at a concentration of 5 μg/ml with the plasmid pRF4 (100 μg/ml) which confers a dominant Roller phenotype (Mello et al., 1991). Transgenic lines were scored for the rescue of *egl-27(we3)* embryonic lethality at 15˚C. Two pools containing C04A2 (C04A2+C44B7 and F31E8+C04A2) showed rescue. These cosmids were injected individually into *egl-27(we3)* and lines tested for phasmid dye filling. Two lines carrying C04A2 showed rescue (phasmids failed to fill in 0% and 20% of rollers compared to 66% of *egl-27(we3)* mutants).

Tissue assays

Antibody staining of embryos was essentially as in Albertson (1984). The primary antibodies used are: anti-LIN-26 rabbit polyclonal serum (marks all epidermal and other non-neuronal ectodermal cells; Labouesse et al., 1996) mAb MH27 which recognizes a component of epithelial adherens junctions (Francis and Waterston, 1991), mAb NE2/1B4 against an antigen expressed in seam cells (Schnabel, 1991) and mAb 3NB12, which detects a subset of pharyngeal muscle cells (Priess and Thomson, 1987).

For the time course, the development of wild-type and double mutants was followed by DIC microscopy for 8 hours at 1 hour time intervals at 22˚C, the embryos were placed at 15˚C overnight and the terminal phenotype viewed 12 hours later.

RNAi

Templates for RNA synthesis contained T3 and T7 phage polymerase promoter sites; single strands were synthesized separately and then annealed as described (Fire et al., 1998). Templates were prepared by PCR using the following primers: for *egl-27(exon 11)*, 5′-ATTA-ACCCTCACTAAAGGGAGACTTCGATGAGCGAGACTCC and 5′-AATACGACTCACTATAGGGAGATCGCATTTGATGTTGGAATGC using cosmid C04A2 as a template; for *egl-27(exons 1-8)*, T3 and T7 primers using pJA54 (a RACE clone containing exons 1-8) as a template; *egr-1(RNAi)*, T3 and T7 primers using yk394g5 (an *egr-1* cDNA of 436nt covering exons 1, 2 and 199 bp of exon 3) as a template. Double-stranded RNA was injected into the body cavity of adult hermaphrodites at a concentration of 0.5-1 mg/ml. Injected animals were singled on plates for 12 hours, then transferred to new plates every 24 hours. Phenotypes were assayed on the first new plate. Control injections of dsRNA to *egl-27(exon 11)* into *weEx33* (carrying the translational *egl-27::gfp* reporter gene) resulted in the absence of GFP expression indicating that RNAi removes *egl-27* proteins.

Isolation and sequencing of cDNA clones and northern blotting

For RACE experiments, the template was mixed stage *C. elegans* cDNA to which 5′ and 3′ anchors were ligated (Clontech, Marathon cDNA Amplification Kit), kindly provided by Howard Baylis. PCR was carried out using a primer in *egl-27* (5′ CATGTTGTCAACG-TCTGTCG) and a primer to the anchor (5′ CCATCCTAATACGA-CTCACTATAGGGC). Several products were obtained and cloned. The major product contained exons 1-8; 2/10 clones sequenced had part of SL1 at the 5′ end of exon 1 (one was named pJA54). Several smaller minor products were sequenced; two had part of SL1 (TTTGAG) at the 5′ end of exon 6. Further PCR reactions using wild-type mixed stage cDNA confirmed additional exons. Using JA45 (in exon 8; 5′ TCACATACTTCATTTTCAGCT) and JA42 (in exon 11; 5′ GGACCATGACCATTTGTGAG) a product containing exons 8, 9, 10 and 11 was obtained. PCR reactions, using a primer to SL1 and one in exon 11 (JA42), identified products with SL1 on the 5′ ends of

exon 11 and of exon 10. *egl-27* cDNA clones yk27d3 and yk1e2 (kindly supplied by Yuji Kohara) were completely sequenced and found to begin within exon 10 (yk27d3 at bp 9726 of C04A2 and yk1e2 at bp 9800) and end with poly(A) tails in exon 13 at position 15204 of C04A2. Because the small major RNAs are the most abundant messages (northern blot, Fig. 3), and SL1 is trans-spliced to exons 10 and 11, this suggests that these are the true 5′ ends of the small major messages.

The start and stop positions of the *egl-27* exons in cosmid C04A2 are as follows: exon 1 (2318-2462), exon 2 (2752-2870), exon 3 (3924-4064), exon 4 (5149-5274), exon 5 (5324-5388), exon 6 (7490-7652), exon 7 (7748-7927), exon 8 (7980 or 7989-8207), exon 9 (8782-9010), exon 10 (9707-9902), exon 11 (12659-13784), exon 12 (13828-14337) and exon 13 (14388-15203).

The northern blots were done as in Ahringer (1996), using approximately 1 μg of mixed-stage wild-type poly(A)+ RNA per lane. RNA markers (Promega) were used as size standards.

Construction of *egl-27*::*gfp*

A genomic fragment covering exons 1-11 (encoding amino acids 1-845) of the *egl-27* large major transcript were fused in frame to the coding region of GFP: a 13.4 kb *Sma*I-*Bam*HI fragment from C04A2 (see Fig. 2) was ligated to the promoterless GFP vector pPD95.75 (A. Fire, J. Ahnn, G. Seydoux and S. Xu, personal communication) that had been cut with *Hin*dIII, filled in with Klenow, then cut with *Bam*HI. The resulting plasmid is called pJA53. Three independent transgenic lines containing pJA53 (20 μg/ml) and pRF4 (100 μg/ml) were produced as described (Mello et al., 1991): *weEx33*, *weEx34* and *weEx35*. All three express GFP in essentially all somatic nuclei.

Computational analyses

Regions of sequence similarity were found using Blast and PSI-Blast (Altschul et al., 1997) using the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST/) and default parameters. The zinc finger was found using ProfileScan (Bairoch and al., 1997; http://www.isrec.isb-sib.ch/software/PFSCAN_form.html).
Alignments were done using ClustalW (Thompson et al., 1994) with manual editing.

RESULTS

Isolation of *egl-27(we3)*

Strong or null mutations in the *C. elegans even-skipped* homologue *vab-7* cause abnormal patterning of posterior muscle and epidermal cells, which results in a disorganised posterior end (Ahringer, 1996). To identify additional genes important for embryonic patterning, we screened for genetic enhancers of a weak *vab-7* allele. *vab-7(ed6)* mutants have very mild posterior defects (Fig. 1A). We obtained 5 enhanced strains with stronger posterior defects from 968 chromosomes screened and from the strongest isolated the mutant *we3*. When separated from *vab-7(ed6)*, *we3* mutants are 89.5% lethal and have variable body defects which are most severe in mid- and posterior regions (Fig. 1B; Table 1). These defects are cold sensitive (Table 1) and maternally rescued (see Materials and Methods). *we3* retains some zygotic activity, but when grown as a homozygote at 15˚C, it behaves as a strong loss-of-function mutation (see Materials and Methods).

We retested the enhancement of *vab-7(ed6)* by *we3* and found that it has a zygotic component. *we3* homozygotes derived from *we3*/+ mothers have essentially wild-type morphology due to maternal rescue. However, *we3; vab-7(ed6)* homozygotes derived from *we3*/+; *vab-7(ed6)* mothers have stronger posterior defects than *vab-7(ed6)* alone (see Materials and Methods).

From complementation tests, we found that *we3* is a new allele of the gene *egl-27* (see Materials and Methods). The canonical allele of *egl-27*, *n170*, has pleiotropic defects shared by *we3*: hermaphrodites are variably egg-laying defective (Egl), male tails (the copulatory structure) have abnormal morphology and the phasmid sensory neurons, which are normally exposed to the environment, fail to fill with dye (Desai et al., 1988; Garriga et al., 1993; Trent et al., 1983; Herman et al., 1999; Table 1, Materials and methods, and data not shown). However, unlike *we3*, *egl-27*(*n170*) mutants have normal body morphology.

Muscles and epidermal cells are disorganised in *egl-27(we3)* mutants

To investigate the body morphology defects in *egl-27(we3)* embryos, we assayed the organization of two tissues where *vab-7* mutants are known to have patterning defects: body wall muscles and the epidermis (Ahringer, 1996). In wild-type embryos, muscles are arranged in 4 rows running from anterior to posterior (Fig. 1C,D). In *egl-27(we3)* mutants, although muscle cell number is normal (Table 2), muscle cells are found

Table 1. *egl-27(we3)* and RNAi phenotypes

Strains	Embryo-L1 lethal	Viable with body defects	Wild-type morphology	Dye-filling defective*
wild-type (*n*=532)	0.4%	0%	99.6%	0% (*n*=450)
egl-27(we3) 15°C (*n*=456)	89.5%	8%	2.5%	66% (*n*=41)
egl-27(we3) (*n*=734)	7%	42%	51%	68% (*n*=28)
*egl-27(exon 11)*15°C (*n*=187)	96.8%	0%	3.2%	nd
egl-27(exon 11) (*n*=425)	20.5%	61%	18.5%	100% (*n*=104)
egl-27(exons 1-8) (*n*=510)	0%	0%	100%	47% (*n*=98)
egr-1(RNAi) (*n*=671)	0%	3%	97%	0% (*n*=109)
egr-1(RNAi) +egl-27(exon11) (*n*=679)	100%	0%	0%	na
egr-1(RNAi) +egl-27(we3) (*n*=342)	100%	0%	0%	na
egr-1(RNAi) +egl-27(exons 1-8) (*n*=152)	96.7%	0%	3.3%	nd
egr-1(RNAi) +egl-27(n170) (*n*=241)	100%	0%	0%	na

The first column specifies the strain and the number of individuals scored for the phenotypes in columns 2-4. When no temperature is given, the experiment was done at 22°C.

*Dye filling was scored in adult hermaphrodites; *n* specifies the number of individuals examined; na, not applicable; nd, not done.

in clusters and often do not form rows (Fig. 1E,F). This phenotype is similar to, but stronger than that seen in *vab-7* null mutants (Ahringer, 1996). These defects are not due to loss of *vab-7* protein since *vab-7* is expressed normally in *egl-27(we3)* mutants (not shown). In addition, the abnormal pattern of muscle cells in *egl-27(we3)* mutants is not confined to the posterior as in *vab-7* mutants.

To analyse epidermal cell patterning, we looked at seam cells, which are lateral epidermal cells. Late stage embryos have 20 seam cells that lie in two rows along the animal (Fig. 1G; Sulston et al., 1983). In *egl-27(we3)* mutants, the seam is often branched and interrupted (Fig. 1H), indicating that seam cell organization is abnormal. However, seam cells are differentiated as shown by their expression of a seam antigen (Fig. 1H), and their number is normal (Table 2), as counted by the nuclear seam cell marker SCM (Terns et al., 1997).

Table 2. Muscle and seam cell number in *egl-27(we3)* and double mutants

Strains	Muscle cell count	Seam cell count
wild-type	61.2 (*n*=6; range=55-68)	15.8 (*n*=39, range=10-20)
egl-27(we3)	62.5 (*n*=6; range=54-70)	17.8 (*n*=56, range=8-20)
egl-27(exon 11) + *egr-1 (RNAi)*	59.4 (*n*=7; range=46-69)	4.5* (*n*=23, range=1-12)

Muscle cell numbers were counted from the expression of *hlh-1*::GFP (Krause et al., 1994); muscle numbers counted are lower than the 81 expected in wild-type embryos, probably due to mosaic expression of the reporter. Seam cell number was counted using *SCM::lacZ* (Terns et al., 1997).

*Seam cell number appeared normal based on the expression of a seam antigen (i.e., many seam cells not expressing SCM did express the seam antigen in Fig. 7D; data not shown).

These data show that *egl-27* is not required for the production or differentiation of muscle or seam epidermal cells, but is needed for their proper positioning in the embryo. As shown previously for *vab-7* mutants (Ahringer, 1996), such disorganisation is an indication that patterning within a tissue is defective. Postembryonic defects of *egl-27(we3)* mutants (in egg-laying, male tail morphology and phasmid dye filling) suggest that *egl-27* also has additional developmental roles.

Cloning of *egl-27*

The genetic mapping of *egl-27* narrowed the gene to a region of 9 cosmids. We first injected pools of cosmids and then single cosmids into *egl-27(we3)* mutants and tested for rescue of the mutant phenotype. We found that cosmid C04A2 had rescuing activity (see Materials and Methods). To determine which gene on C04A2 encodes *egl-27*, we took advantage of RNA-mediated gene inactivation (RNAi; Fire et al., 1998). Injection of double-stranded RNA (dsRNA) corresponding to a gene into a wild-type hermaphrodite will inactivate that gene in her progeny, mimicking a null phenotype (Fire et al., 1998). We found that *C04A2.2(RNAi)* resulted in progeny

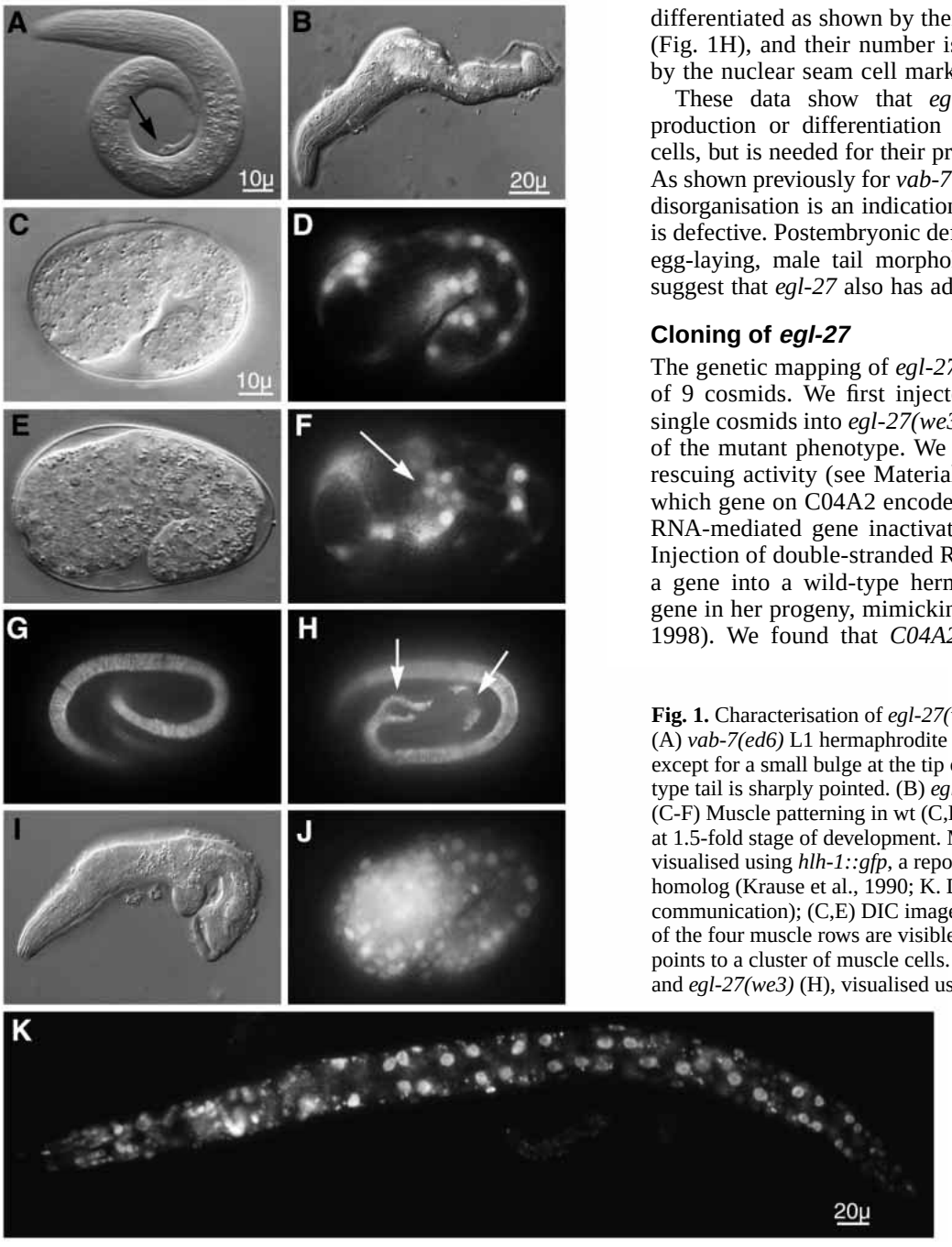

Fig. 1. Characterisation of *egl-27(we3)* and *egl-27*::*gfp* expression. (A) *vab-7(ed6)* L1 hermaphrodite is nearly identical to wild type except for a small bulge at the tip of the tail (arrow), where the wild-type tail is sharply pointed. (B) *egl-27(we3)* L1 grown at 15°C. (C-F) Muscle patterning in wt (C,D) and *egl-27(we3)* (E,F) embryos at 1.5-fold stage of development. Muscle cells in D and F are visualised using *hlh-1::gfp*, a reporter for the *C. elegans* MyoD homolog (Krause et al., 1990; K. Dej, S. Xu, and A. Fire personal communication); (C,E) DIC images of the same embryos. In D, two of the four muscle rows are visible in this focal plane. Arrow in F points to a cluster of muscle cells. (G,H) Seam patterning in wt (G) and *egl-27(we3)* (H), visualised using antibody NE2/1B4.14; arrows in H shows forked and disrupted posterior seam. (I) L1 larvae induced by *egl-27(exon 11)* RNAi. (J,K) *egl-27*::*gfp* expression in 1.5-fold embryo (J) and L3 hermaphrodite (K). All somatic cells appear to express the reporter gene. Scale bar in B is for B,I; scale bar in C is for C-H and J.

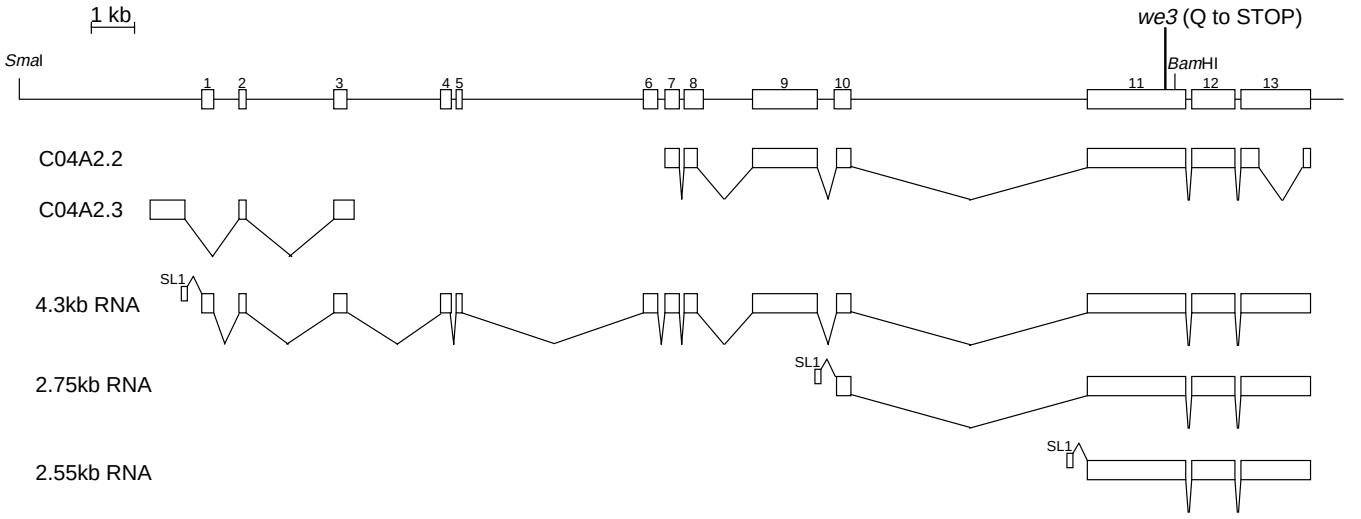

Fig. 2. *egl-27* gene structure. The top line shows 15 kb of C04A2, containing all of the *egl-27* exons (open boxes). Sizes of introns and exons are to scale. C04A2.2 and C04A2.3 are gene structures predicted (The *C. elegans* Sequencing Consortium, 1998). Below are our predicted major RNA gene structures, based on the sequences of cDNA and RACE clones and on northern blotting. Excluding a poly(A) tail, the lengths of the predicted major RNAs are 4051 nt, 2664 nt and 2468 nt (including 22 nt from SL1), in good agreement with the sizes of 4.3 kb, 2.75 kb and 2.55 kb from northern blotting (see Fig. 3; Materials and Methods). The *Sma*I and *Bam*HI sites marked were used to make the *egl-27*::*gfp* translational fusion.

with a phenotype very similar to that of *egl-27(we3)* (Fig. 1I). By sequencing this locus from *egl-27(we3)* DNA, we found that the *we3* mutation introduces a stop codon into the gene (see below), confirming that C04A2.2 corresponds to *egl-27*. Interestingly, like *egl-27(we3)*, the *C04A2.2(RNAi)* phenotype is cold-sensitive (Table 1, *egl-27(exon 11)* rows). The similarity in phenotype between *egl-27(we3)* and *egl-27(RNAi)* supports the view that *egl-27(we3)* is a strong loss-of-function mutation.

egl-27 encodes multiple transcripts

As C04A2.2 was only a gene prediction (The *C. elegans* Sequencing Consortium, 1998), we sequenced cDNA clones and performed RACE experiments to determine the gene structure. *egl-27* extends much farther 5′ than the C04A2.2 prediction (Fig. 2). We found evidence for multiple *egl-27* transcripts from our cDNA clones: in different cDNAs, the SL1 trans-spliced leader was found on exon 1, exon 6, exon 10 and exon 11 (see Materials and Methods). Spliced leaders are added to the 5′ ends of 70% of mRNAs in *C. elegans* (Zorio et al., 1994). We also found micro-heterogeneity in the 3′ splice site for intron 7 (see legend to Fig. 4).

To explore further the transcripts produced from the *egl-27* locus, we performed northern blots. Fig. 3 shows that at least 7 transcripts are produced: three major transcripts of 4.3 kb, 2.75 kb and 2.55 kb, and four minor ones of 3.5 kb, 3.2 kb, 3.0 kb and 1.5 kb; an exon 11 probe detects all transcripts, but exons 1-8 detect only the 4.3 kb major transcript and some minor transcripts (Fig. 3). The predicted structures of the three major transcripts is shown in Fig. 2, based on the results from northern blotting and cDNA sequencing.

We sequenced *egl-27* exons from *egl-27(we3)* DNA and found that *we3* introduces a stop codon into exon 11 (Figs 2, 4), so *we3* should affect all *egl-27* products. By northern blotting, this mutation does not result in the loss of any *egl-27* RNAs (Fig. 3). This suggests that *egl-27* is not needed for its own expression.

An *egl-27* reporter gene is expressed ubiquitously in nuclei

To learn where *egl-27* is expressed, we constructed a translational fusion of the coding region of GFP to exon 11 of *egl-27*. Since exon 11 is contained in all *egl-27* transcripts, we expect this reporter gene to reflect expression of all products made from the locus. *egl-27*::*gfp* is expressed in the somatic nuclei of most or all cells from the 50-cell stage of embryogenesis through to adulthood (Fig. 1J,K; additional data

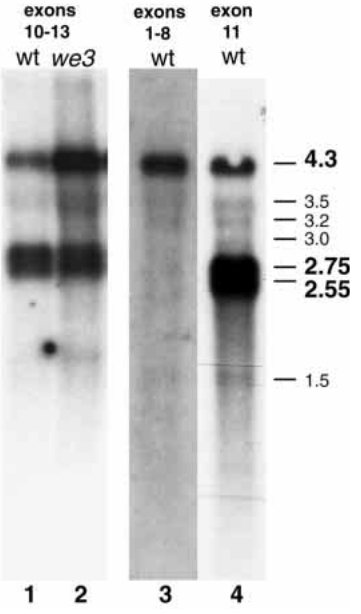

Fig. 3. *egl-27* produces multiple transcripts. Northern blots of wild-type (wt) or *egl-27(we3)* mixed stage poly(A)+ RNA, probed with exons 10-13 (lanes 1 and 2), exons 1-8 (lane 3), or exon 11 (lane 4). Lanes 1 and 2 were run on a different gel from lanes 3 and 4. Sizes are estimated based on standards (not shown).

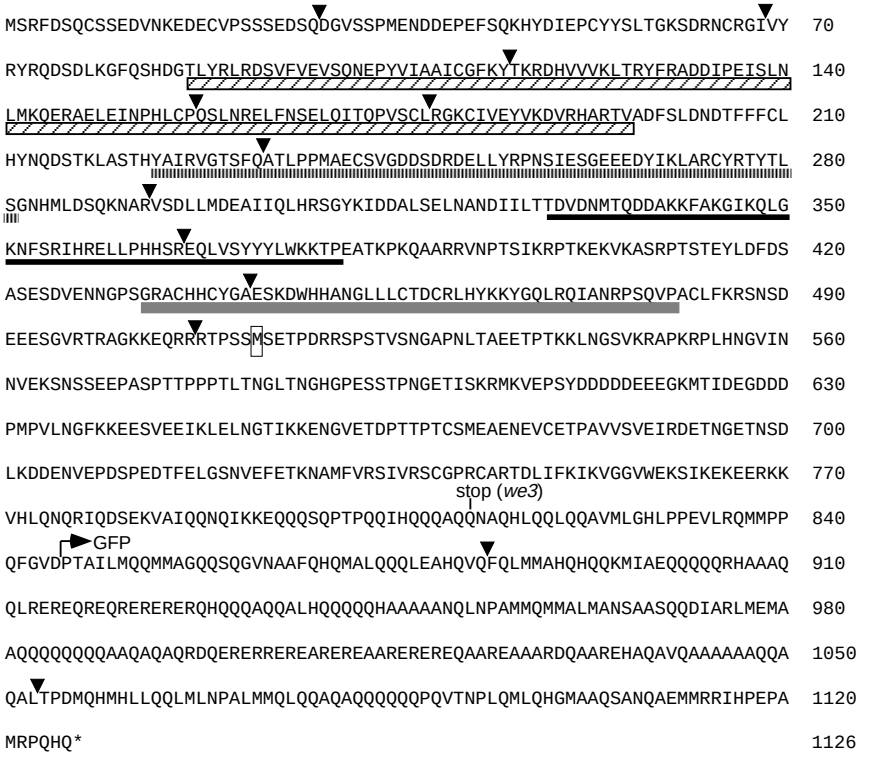

Fig. 4. EGL-27 sequence. Shown is the translation of the 4.3 kb major *egl-27* RNA. The small major RNAs are both predicted to encode a protein that begins at methionine 513 (boxed). The four domains described in the text are underlined: ELM1 with diagonals, ELM2 with verticals, SANT with black, and the GATA-like zinc finger with grey. Shown is the *egl-27(we3)* mutation, identified in three independent PCR clones, which changes a CAA codon to a TAA (stop) and the position where GFP is fused in the *egl-27*::*gfp* reporter gene. Positions of introns are marked with filled triangles. In our RACE clones, either of two 3′ splice sites for intron 7 were used; for the protein shown, the second splice was used as all of our long RACE clones (i.e., those that had ends at or near the beginning of exon 1) used this site. The splice sites differ by 9 nucleotides; splicing at the first one results in a protein with SLQ inserted at the position marked for intron 7.

not shown). *egl-27* is likely to be expressed maternally as well, since *egl-27(we3)* has a maternal effect. This would not have been seen with our reporter construct because transgenes in *C. elegans* usually do not report germline expression (Kelly et al., 1997). Since the GFP coding sequence introduced did not contain a nuclear localisation signal, nuclear expression of the reporter gene suggests that *egl-27* encodes nuclear protein(s).

EGL-27 is similar to a component of a chromatin regulatory complex with histone deacetylase and nucleosome remodelling activities

Conceptual translation of the major 4.3 kb RNA yields a protein of 1126 amino acids (Fig. 4). The 2.75 kb and 2.55 kb major transcripts would encode an identical protein that begins at methionine 513 of the large protein (boxed in Fig. 4).

EGL-27 has highest similarity (by Blast; Altschul et al., 1997) to two human genes and a *C. elegans* gene: KIAA0458, identified from a human brain cDNA library (Seki et al., 1997), MTA1, which has elevated expression in metastasizing mammary adenocarcinomas (Toh et al., 1995), and T27C4.4, a predicted *C. elegans* gene (Figs 5, 6). Comparisons between EGL-27, T27C4.4, KIAA0548 and MTA1 showed that EGL-27 is most related to KIAA0458 whereas T27C4.4 is most similar to MTA1 (Fig. 5A).

MTA1 has recently been shown to be a component of the NURD complex, which has histone deacetylase and ATP-dependent nucleosome remodelling activities (Xue et al., 1998). Histone deacetylation is usually associated with transcriptional repression, whereas nucleosome remodelling is usually associated with transcriptional activation. However, neither the in vivo function of the NURD complex nor transcription factors regulated by it are yet known. The finding that EGL-27 is similar to a component of this complex argues that EGL-27 has a role in the regulation of chromatin structure.

EGL-27 family members are multidomain proteins with a myb-like SANT DNA-binding domain and a zinc finger

Sequence analyses of EGL-27 identified four conserved domains (Fig. 5): (1) A SANT domain, which is similar to the DNA-binding domain of myb (Aasland, 1996; Fig. 5B). The SANT domain was first found in SWI3 (a yeast component of the SWI/SNF complex; Peterson and Tamkun 1995), ADA2 (a component of a histone acetylase complex; Horiuchi et al., 1995), N-CoR (a nuclear hormone co-repressor; Horlein et al., 1995) and the B″ subunit of TFIIIB (a basal pol III transcription factor in yeast; Kassavetis et al., 1995). We propose that this region of EGL-27 has a structure essentially the same as that of myb. (2) A zinc finger related to the GATA family (Fig. 5C). (3) A conserved N terminus that is shared with MTA1 and T27C4.4, which we call the ELM1 (EGL-27 and MTA1 homology) domain. This domain is also found in a number of nuclear proteins, including cytosine-5 methyl transferases, ORC1 (part of the origin recognition complex), the *Drosophila absent* gene (a *trithorax* group protein) and several bromodomain-containing proteins (Fig. 5D), but the region of similarity does not cover any previously defined domains of these proteins. (4) A further conserved region (ELM2), also found in additional proteins, many of which contain SANT (Figs 5E, 6). This region is not similar to any domains of known function. Finally, the C terminus of EGL-27 is acidic and glutamine rich (Fig. 4). The four named domains are encoded by the large major transcript but not the small ones.

Fig. 6 shows a cartoon of the domains that we defined in

A Pairwise sequence identities

	EGL-27	KIAA0458	T27C4.4
KIAA0458	40%		
T27C4.4	22%	27%	
MTA1	26%	36%	45%

B SANT domain

```
1mbe                LGKTRWTREEDEKLKKLVEQNGTDDWKVIANYLPNRTDVQCQHRWQKVLNPE
ss                  .......HHHHHHHHHHHHH.....HHHHHHH.....HHHHHHHHH......
EGL-27/329-380      TDVDNMTQDDAKKFAKGIKQLGKNFSRIHRELLPHHSREQLVSYYYLWKKTP Q09228
KIAA0458/123-174    LIEKCWTEDEVKRFVKGLRQYGKNFFRIRKELLPNKETGELITFYYYWKKTP g3413878
T27C4.4/479-530     DQLEEWSTPEMNLFEDALDKVGKDFNEIRAEYLPWKSIRDIVEYYYLMKASN g3165588
MTA1_HUMAN/283-334  DEMEEWSASEANLFEEALEKYGKDFTDIQQDFLPWKSLTSIIEYYYMWKTTD Q13330
N-CoR/435-485       QFMNVWTDHEKEIFKDKFIQHPKNFGLIASYLE.RKSVPDCVLYYYLTKKNE g2137603
N-CoR/622-672       VETSRWTEEEMEVAKKGLVEHGRNWAAIAKMVG.TKSEAQCKNFYFNYKRRH g2137603
Er1_xenla/272-323   EELSVWTEEECRNFEQGLKAYGKDFHLIQANKVRTRSVGECVAFYYMWKKSE g2529737
```

C GATA-like domain

```
1gat                GTVCSNCQTSTTTLWRRSPM...GDPVCNACGLYYKLHQVNRPLTMRKDGIQ
ss                  ..............EEEE......EEEEHHHHHHHHHH.....HHH......
EGL-27/433-480      GRACHHCYGAESKDWHHAN....GLLLCTDCRLHYKKYGQLRQIANRPSQVP Q09228
KIAA0458/236-283    GYACRHCFTTTSKDWHHGG..RENILLCTDCRIHFKKYGEL..PPIEKPVDP g3413878
T27C4.4/584-635     DNPCENCGTLDALNWYQWGGVGDKKVLCSTCWIQWKKFAGLNQKHELERFDK g3165588
MTA1_HUMAN/390-441  GRACESCYTTQSYQWYSWGPPNMQCRLCASCWTYWKKYGGLKMPTRLDGERP Q13330
```

D ELM1 domain

```
EGL-27/87-187       TLYRLRDSVFVEVSQN.......EPYVIAAICGFKYTKRDHV.VVKLTRYFRADDIP.
T27C4.4/13-245      VTYAVGDFVYFDDTSA.(24).CVVYLRRRDIPQHLLKIADQAQRRFDNYYEVDKKK.
MTA1_HUMAN/4-131    NMYRVGDYVYFENSSS......NPYLIRRIEELNKTANGNV..EAKVVCFYRRRDIS.
Q43479/26-112       GVIRPGDSVLMKAPDS.(1)..KPPYVAKIEEIEAAGPRGAN.VKVKVRWYYRPEES.
ORC1_HUMAN/45-138   IHIQIGQFVLIEGDDD......ENPYVAKLLELFEDDSDPPPKKRARVQWFVRFCEV.
P70049/45-138       ITVTPGDFVLIEGENE......ERPFVAKLQELYDDGNEKHTSKHAIVQWFLRYEEV.
YG23_YEAST/368-452  EKYQIGDWVLLHNPND.(1)..NKPIVGQIFRLWSTTDGN...KWLNACWYFRPEQT.
Q18210/734-819      TKYVAPCYAYVSRSDE.(2)..TPLHIFRIERTFKDENGE...KALQGHWVYRPEET.
Q90941/954-1038     SMYHVGDYVYVEPAEA.(1)..LQPHIVCIERLWEDSAGE...KWLYGCWFYRPNET.
Q24189/1914-2021    LQVRQGDAVYVLRDIP.(24).QECDIFRVEHLWKNELGK...RFIFGQHFLRPHET.
Q18210/942-1026     KFFWLGQCVLVFNNMK.(1)...LCDVMKINKIWREKDGS...EWFSGCWFARPSET.
Q90941/1155-1237    MWLKVGDCVFIKSHGL......VRPRVGRIEKMWVRDGA....AYFFGPIFIHPEET.
MTDM_ARATH/735-817  EMVAVGGAVTLEVDDP.(1)..EMPAIYFVEYMFESTDHC...KMLHGRFLQRGSMT.
Q27746/743-825      EKIEIGDCVLIHPDDP.(1)..KPLFMARVIYMWQESQGE...MMFHAQWFVYGSET.

EGL-27/87-187       .(18)..HLCPQSLNRELFNSE...LQITQPVSCLRGKCIVEYV Q09228
T27C4.4/13-245      .(124).DQRLKLRQHEIFMTR...QSEILPAAAIRGKCRVVLL g3165588
MTA1_HUMAN/4-131    .(35)..KLKHQLRHRELFLSR...QLESLPATHIRGKCSVTLL Q13330
Q43479/26-112       .(2)...GRRPFHCEKEVFLSD...HQDVQSADTIECKCNVYSF Q43479
ORC1_HUMAN/45-138   .(6)...LLGRKPGAQEIFWYDYPACDSNINAETIIGLVRVIPL Q13415
P70049/45-138       .(6)...LLGREPHQEEIFLYDVPSCENDIDAETIIGSVKVTQL P70049
YG23_YEAST/368-452  .(2)...RVDRLFYKNEVMKTG...QYRDHPIQDIKGKCYVIHF P53236
Q18210/734-819      .(2)...LASRKFMKQEVFLTP...FRDTVLAERLRGRCVVISL Q18210
Q90941/954-1038     .(2)...LATRKFLEKEVFKSD...YYNKVPVSKILGKCVVMFV Q90941
Q24189/1914-2021    .(2)...EPSRRFYPNEVVRVS...LYEVVPIELVIGPCWLLDR Q24189
Q18210/942-1026     .(2)...DEGRLFFKNEVIAVYR..NDETRKLCEIQRVCDVMPA Q18210
Q90941/1155-1237    .(2)...EPTKMFYKKEVFLSN...LEETCPMSCILGKCAVLSF Q90941
MTDM_ARATH/735-817  .......VLGNAANERELFLTN...ECMTTQLKDIKGVASFEIR P34881
MTDM_HUMAN/634-716  .......VLGATSDPLELFLVD...ECEDMQLSYIHSKVKVIYK P26358
Q27746/743-825      .......VLGETSDPLEVFPID...ECQDTYLGSVNAKCTVIYK Q27746
```

E ELM2 domain

```
EGL-27/224-281      YAIRVGT.SFQATLPPMA.(12).LLYRPNSIESG...EEEDYIKLARCYRTYTLS Q09228
KIAA0458/16-74      GEIRVGP.SHQAKLPDLQ.(14).ELVWMPGVNDC...DLLMYLRAARSMAAFAG. g3413878
T27C4.4/278-401     GAIRVGE.KYQAVVDEWM.(78).VWHPHHALTDR...DIDQYMIVARSVGLFARA g3165588
MTA1_HUMAN/165-226  GEIRVGN.RYQADITDLL.(16).VWEAHNPLTDK...QIDQFLVVARSVGTFARA Q13330
KIAA0071/14-75      GGMRVGP.QYQAVVPDFD.(17).VWSPNQNLSEA...KLDEYIAIAKEKHGYNM. g505104
F10E7.11/81-137     RSIRVDPVLFQADVPLFN.(14).LWTIDQTNQPSD.EVIDNYLKDVVGLRKAH.. g1086837
Er1_xenla/169-229   KEIMVGS.MFQAEIPVGI.(16).LWNPEYVMEER....VIDFLNEASRRTCEERG g2529737
D1014.9/107-167     DEINVGT.EFQAKIADLN.(16).IWNTPETIDDE...KLEAFIRESSDRYLIPI. g1256279
F53H10.2/501-565    PHINLGK.NYQARVKKWC.(17).IVFSSEILQDIDPEQITAFELLACSQACPRA. Q20733
```

Fig. 5. Alignments of EGL-27 domains. (A) Percent identity between EGL-27, KIAA0458, T27C4.4 and MTA1, summed over the ELM2, GATA and SANT domains.
(B-E) Alignments of EGL-27 domains. Conserved residues are highlighted using the belvu program (E. Sonnhammer, personal communication) in default mode. White letters on a black background are the most conserved, white letters on a grey background show intermediate conservation, and black letters on a grey background are the least conserved. The secondary structure is marked in the ss row, with E denoting β-strand and H denoting α-helix. For each sequence, the start and end points follow the identifier of the sequence, with the TrEMBL or SWISSPROT accession number or genbank PID following the alignment. (B) Alignment of myb-like SANT domains with the structurally solved myb domain 1mbe (Ogata et al., 1994). (C) Alignment of the GATA-like domain with the structurally solved gata domain 1gat (Omichinski et al., 1993). (D) Alignment of the ELM1 domain.
(E) Alignment of the ELM2 domain. Not all identified matching proteins are included in the alignments.

EGL-27 with other proteins that contain them. T27C4.4 is likely to be the only other *C. elegans* gene that contains all the identified domains of EGL-27, as >99% of the *C. elegans* genomic sequence is available (The *C. elegans* Sequencing Consortium, 1998); we have named this gene *egr-1*, for *egl-27*-related gene.

egr-1 and *egl-27* have a redundant embryonic function

Because *egr-1* shares the domain structure of *egl-27*, we wondered whether it had a similar function. To investigate this, we used RNAi to inactivate *egr-1*, alone or in combination with *egl-27*. *egr-1(RNAi)* animals are completely viable, with a

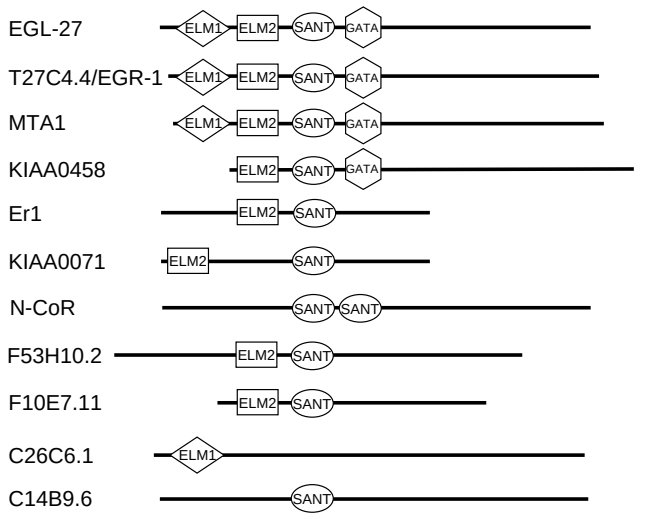

Fig. 6. Proteins with domains similar to those in EGL-27. Proteins shown were found to contain domains similar to EGL-27 using Blastp or PSI-blast (Altschul et al., 1997). Not every protein identified is shown.

small number having mild posterior defects (Table 1; Fig. 7A). However, *egr-1(RNAi); egl-27(exon 11)* or *egr-1(RNAi); egl-27(we3)* double mutants are 100% embryonic lethal (Table 1), indicating that *egl-27* and *egr-1* have a redundant function in the embryo.

Further RNAi experiments showed that this redundancy is at the level of the large *egl-27* RNA. Since exon 11 of *egl-27* is contained in all transcripts, *egl-27(exon 11)* should interfere with all *egl-27* transcripts. To interfere with only the large *egl-27* RNA, we injected dsRNA to exons 1-8 (*egl-27(exons 1-8)*), which are not contained in the small RNAs. In contrast to *egl-27(exon 11)*, *egl-27(exons 1-8)* animals do not have body morphology defects, but do have dye-filling defects (Table 1), a phenotype similar to that of *egl-27*(*n170*) mutants. This is consistent with the findings of Herman et al. (1999), who showed that *egl-27*(*n170*) deletes exons in the large transcript, but not the small ones. We next examined *egr-1(RNAi); egl-27(exons 1-8)* and found that these double mutants arrested as embryos with the same phenotype as *egr-1(RNAi); egl-27(exon 11)* (Table 1 and data not shown). This shows that the large *egl-27* RNA, which contains the similarity to *egr-1* and MTA1, has a shared function with *egr-1*.

Double mutant embryos are abnormally patterned

Double mutant embryos arrest with a striking uniform phenotype: tissues are not properly organised, but they are well differentiated and morphogenesis fails to occur (Fig. 7B-F). To find out when defects initially occur in double mutants, we compared their development to those of wild-type embryos developing at the same time. Wild-type embryogenesis can be conveniently divided into three phases: first, during the initial cell divisions, the axes of the embryo are established and blastomere fates are determined through the actions of maternal genes (Bowerman, 1998). Second, at the 28-cell stage, gastrulation begins. During this phase, the cells of the body are patterned and most of embryonic cell proliferation occurs. At the end of this second phase, cells lie in precise positions where they will form tissues, but they do not yet express terminal differentiation markers. Disruption of regional patterning during this time causes cells to be abnormally positioned (see e.g., Ahringer, 1996; Chisholm and Horvitz, 1995). During the final stage, morphogenesis causes the organised ball of cells to be transformed into a long thin worm as tissues and organs differentiate and become functional.

We found that double mutant embryos develop at the same rate as wild type (Fig. 8). However, cells of the embryo are not properly organised during gastrulation (Fig. 8). For example, intestinal cells in wild type are located in an ordered row but, in double mutants, they are found clustered together (compare Fig. 8C to H). Other cells in double mutants appear similarly disordered (Fig. 7B-F and data not shown).

As development proceeds, double mutants do not undergo morphogenesis (compare Fig. 8D,E with I,J), but tissues differentiate normally. For example, intestinal cells in double mutants produce gut granules (a marker of differentiation; Fig. 7O). Muscle cells are found in a wild-type number (Table 2) and are functional, as contractions are evident in double mutants, but they are scattered and are not organised into rows as they are in wild type (compare Fig. 7C to 1E). Pharyngeal tissue is located within a basement membrane and forms a recognizable grinder, but the pharynx is not elongated as in wild type (compare Fig. 7H to 7J). Epidermal cells express the apical membrane antigen MH27 (data not shown), but they do not enclose the embryo (Fig. 7F). Seam cells (lateral epidermal cells) in double mutants are also disorganised but differentiated, as visualised by a seam antigen expressed near the end of embryogenesis (compare Fig. 7D to Fig. 1H). However, most seam cells that express the seam antigen fail to express the nuclear seam cell marker SCM (Table 2 and data not shown). This suggests that SCM may be a target of *egl-27*/*egr-1*. Muscle, epidermal, pharyngeal, neuronal and intestinal cells appear to be well differentiated and present in the correct number based on antibody staining and visual inspection (Table 2; Fig. 7 and data not shown). The mispositioning of cells in all tissues coupled with normal

Table 3. Patterning gene expression in wild-type and double mutant embryos

Strains	*vab-7*	*lin-39*	*mab-5*	*egl-5*	*lin-44*
wild-type	21.6 (*n*=7) range=20-25	12.8 (*n*=12) range=10-16	17.7 (*n*=6) range=15-20	20.4 (*n*=7) range=16-25	4.7 (*n*=11) range=2-8
double mutant	21.9 (*n*=15) range=18-31	12.8 (*n*=10) range=11-16	18.1 (*n*=9) range=15-21	20.2 (*n*=16) range=17-25	5.0 (*n*=8) range=1-8

Double mutants were generated by injecting ds RNA to *egr-1(RNAi) + egl-27(exon11)* . The number of cells expressing the indicated reporter genes are shown. Counts in wild-type and double mutants were done in embryos of similar ages maintained at 22°C.

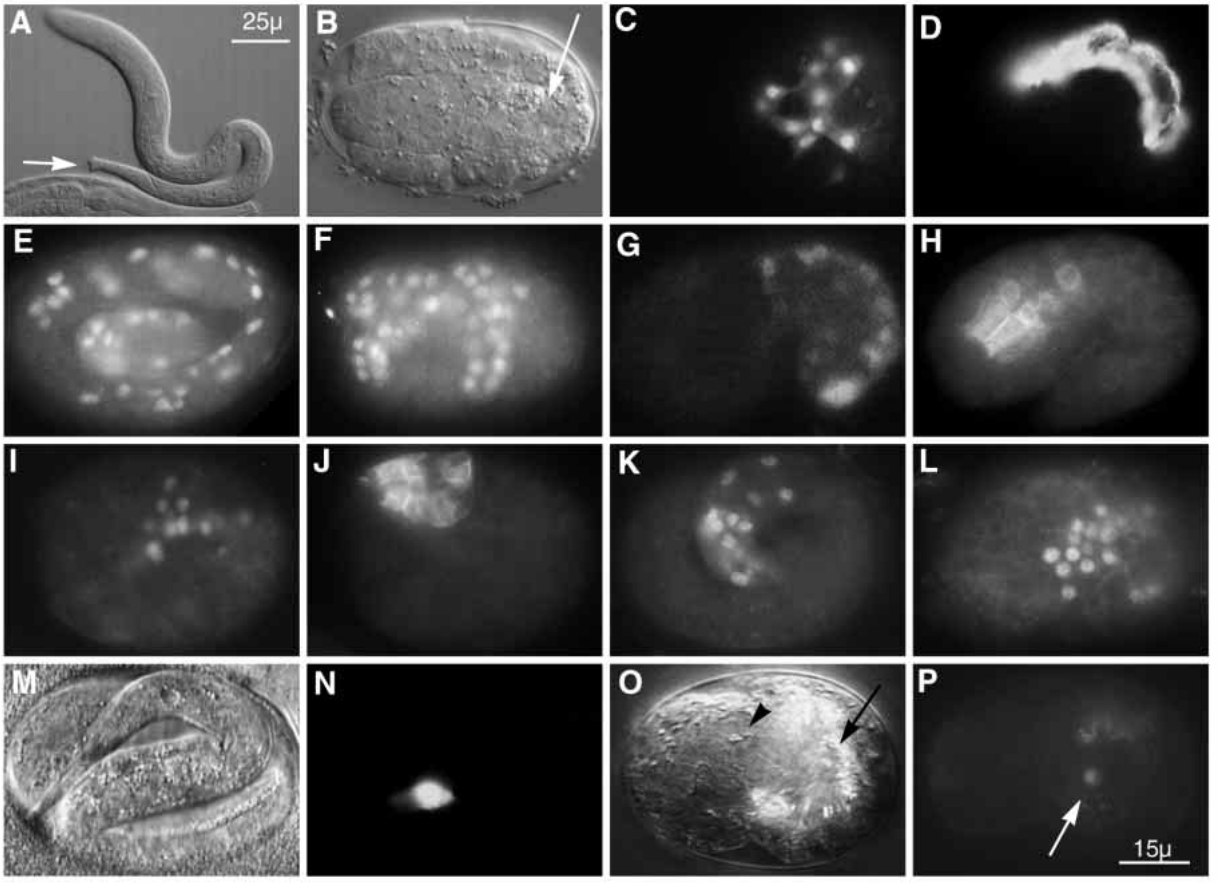

Fig. 7. *egl-27* and T27C4.4 have a redundant global patterning function. (A) L2 larvae from an hermaphrodite injected with *egr-1* dsRNA; arrow points to slight tail truncation. (B) Double mutant embryo from an hermaphrodite co-injected with *egr-1* dsRNA and *egl-27* dsRNA; arrow points to a cluster of intestinal cells. (C) *hlh-1::gfp* expression of (B), showing disorganised muscle cells. (D) Double mutant seam antigen expression; compare to Fig. 1H. (E) LIN-26 expression in a wt threefold embryo in rows of epidermal cells. (F) LIN-26 expression in a double mutant; epidermal cells are clustered on the surface. (G) wt *vab-7*::lacZ expression. (H) Pharyngeal muscle cells of wt embryo. (I) Double mutant *vab-7*::lacZ expression. (J) Pharyngeal muscle cells of embryo in I. (K) wt *egl-5*::*gfp* expression. (L) Double mutant *egl-5::gfp* expression. (M) wt *hlh-8::gfp* containing embryo. (N) *hlh-8::gfp* expression in the M cell of embryo in M. (O) Double mutant *hlh-8::gfp* containing embryo; arrow points to a round cluster of intestinal cells with visible birefringent gut granules, arrowhead points to the pharyngeal grinder. (P) Weak *hlh-8::gfp* in the M cell of double mutant embryo in O (arrow). This expression was not visible using the photographic settings for the wild-type expression in N. Scale bar in P is for B-P.

differentiation suggests that *egl-27* and *egr-1* specifically function in cell patterning in diverse tissue types.

A HOX gene target is not expressed in double mutants

The global disorganisation of cells of double mutants is similar to that seen regionally in the posterior of *vab-7* mutants. Since *egl-27*/*egr-1* are likely to affect transcription, we reasoned that the phenotype could be due to changes in the expression of many regionally acting developmental genes. To test this, we first assayed the expression of *vab-7*. Double mutants express *vab-7* in a normal number of cells, but in a disorganised pattern (Table 3; Fig. 7G,I). Because *egl-27* and *egr-1* are not required for *vab-7* expression, they either act with or downstream of *vab-7*. Likewise, *egl-27*::*gfp* expression is normal in *vab-7* mutants suggesting that *vab-7* does not regulate *egl-27* transcription (not shown).

We next asked whether the expression of other regionally acting genes might be affected in double mutants. We assayed the expression of four zygotic patterning genes that are first expressed in the embryo: three HOX genes (*lin-39*, *mab-5* and *egl-5*; Costa et al., 1988; Clark et al., 1993; Wang et al., 1993) and *lin-44*, a *wnt* homologue required for the correct pattern of cell divisions in the tail (Herman and Horvitz 1994; Herman et al., 1995). Like *vab-7*, all are expressed in a normal number of cells (Fig. 7K-L; Table 3). This shows that many developmental patterning genes are expressed normally in double mutants.

Since EGL-27 and EGR-1 are likely to regulate chromatin structure, thereby affecting the ability of transcription factors to function, one possibility is that they act in conjunction with transcription factors involved in patterning. To test this idea, we assayed the expression of *hlh-8*, a target of the HOX gene *mab-5* (Harfe et al., 1998) in double mutants. So far, this is the only known target of any of the genes assayed above. In wild-type four-fold embryos, where the pharyngeal grinder is well formed, *hlh-8*::*gfp* is highly expressed in the M cell (20/20 embryos; Fig. 7N). In contrast, in double mutants with a visible

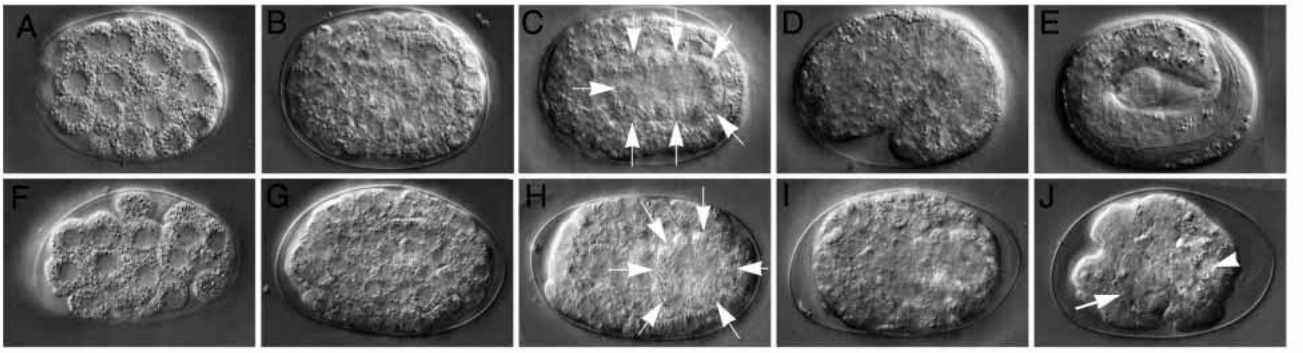

Fig. 8. Time course of wild-type and double mutant development. The development of wild-type (A-E) and *egr-1 (RNAi); egl-27 (exon 11)* double mutant (F-J) embryos was followed at 22°C until wild type was undergoing morphogenesis (D); embryos were then placed at 15°C overnight and photographed 12 hours later (E,J). In this figure, the zero time point is at the 50-cell stage (A,F). From this reference, other time points are (B,G) 2 hours; (C,H) 4 hours; (D,I) 8 hours. Cell number in wild-type and double mutant embryos is similar at all time points. (B,G) Wild-type and double mutant embryos are undergoing gastrulation. (C,H) End of gastrulation. Intestinal cells in wild type are arranged in a long anteroposterior tube (arrows in C); in double mutants, intestinal cells form a ball (arrows in H). (D,I) Wild-type embryo undergoing morphogenesis and elongating (D); no elongation occurs in the double mutant embryo (I). (E,J) The wild-type embryo (E) has fully elongated and was moving within the eggshell; the double mutant embryo (J) has not undergone morphogenesis but tissues are well-differentiated (arrow points to pharyngeal tissue, arrowhead to intestinal tissue, muscle contractions and neuronal processes were also evident). Anterior, left.

grinder, *hlh-8*::*gfp* is essentially absent (5/10 embryos had no expression and 5/10 had the barely detectable expression shown in Fig. 7P). We believe that double mutants survive long enough for *hlh-8* to be expressed, since a late embryonic marker first expressed at this time (the seam antigen in Fig. 7D) is present. This supports the hypothesis that *mab-5*, and perhaps other zygotic patterning genes such as *vab-7*, require *egl-27* or *egr-1* for activity.

DISCUSSION

egl-27 and *egr-1* are *C. elegans* genes with similarity to MTA1, a protein found in the NURD chromatin remodelling complex. These genes have an important redundant function in the organisation of embryonic cells during the phase of body patterning. Based on our phenotypic analyses and the known role of MTA1 in chromatin remodelling, we propose that EGL-27 and EGR-1 are components of a protein complex that specifically functions with sequence-specific transcription factors involved in embryonic patterning.

Multiple developmental roles of *egl-27*

In addition to its redundant function with *egr-1, egl-27* also has unique functions. *egl-27* mutants have pleiotropic phenotypes, including abnormal body morphology and postembryonic defects in the development of the male tail and the phasmids (Desai et al., 1988; Garriga et al., 1993; Trent et al., 1983; Herman et al., 1999; Table 1, Materials and methods, and data not shown). Our studies focused on the embryonic role of *egl-27* and showed that it is required for proper pattern of muscle and epidermal cells. Herman et al. (1999), who also cloned *egl-27*, studied other aspects of the *egl-27* phenotype. They found that *egl-27* mutants have defects in T cell polarity, which has been shown to be controlled by *lin-44*, a *wnt* family member (Herman and Horvitz, 1994; Herman et al., 1995). In addition, they showed that *egl-27* mutants have defects in cell migrations and divisions, some of which are controlled by HOX genes. Taken together, these results support the view that there is a widespread requirement for *egl-27* function in developmental patterning.

Functional redundancy of *egl-27* and *egr-1*

egl-27 has partial functional redundancy with a related *C. elegans* gene, *egr-1*. Inhibiting their activities together results in the abnormal positioning of cells without affecting tissue determination or differentiation. To our knowledge, no previously characterised mutants have this phenotype. As we have only studied *egr-1* using RNAi, an important goal for the future will be to identify mutations in the gene.

Although double mutants have global defects in development, several observations argue that *egl-27*/*egr-1* are not affecting cellular health or gene expression generally. First, tissues in double mutants, though not organised, appear healthy. We did not see increased cell deaths, widespread necrosis or lack of adhesion between cells during the time of normal embryogenesis. Second, cell lineages and the process of gastrulation appear to be normal based on the positions and numbers of cell types. Third, tissues in double mutants are well differentiated and functional where we can assay them (e.g., muscle twitching), and even late differentiation markers are expressed (e.g., the seam antigen in Fig. 7D). Therefore, double mutants are not blocked in development at an early stage and *egl-27/egr-1* are not needed for the transcription of terminal differentiation products. Rather *egl-27*/*egr-1* appear to have a specific role in positioning cells in the embryo at the time of major body patterning.

This phenotype of double mutants is consistent with a defect in patterning in all regions of the embryo. We therefore investigated whether the expression of a number of regional developmental genes (*vab-7*, HOX genes and *lin-44*) was affected in double mutants, but found that these are expressed in a normal number of cells. Although expression is normal, the activity of the HOX gene *mab-5*, a homeodomain transcription factor (Costa et al., 1988), is apparently altered, as a target (*hlh-8*) is not expressed in double mutants. This

suggests that MAB-5 requires EGL-27/EGR-1 in order to activate *hlh-8*. We propose that EGL-27/EGR-1 affect the ability of many transcription factors involved in patterning to function, and that cells fail to be organised because their patterning is absent. For example, *vab-7* patterns muscle and epidermal cells in the posterior of the embryo; in its absence, these cells still differentiate as muscle and epidermal tissue, but they are abnormally positioned resulting in disorganisation of the posterior end (Ahringer, 1996). Likewise, the *vab-3* gene, which encodes a Pax6 homologue, patterns anterior epidermal cells; in its absence, the anterior region is disorganised (Chisholm and Horvitz, 1995).

We also found that SCM, a nuclear marker of seam cells, fails to be expressed although seam cells are present and differentiated. One possibility is that SCM may be involved in seam cell patterning and may be regulated by EGL-27/EGR-1. Neither the function of SCM nor how SCM expression is controlled are yet known. Studying how *hlh-8* and SCM expression are activated should shed light on how *egl-27*/*egr-1* and regional patterning genes might cooperate.

Histone acetylation and deacetlylation in development

Many sequence-specific transcription factors function with multiprotein complexes that alter histone acetylation or remodel nucleosomes (reviewed in Kadonaga, 1998; Struhl, 1998). Histone hyperacetylation is associated with transcriptional activity, whereas histone hypoacetylation is correlated with transcriptionally silent chromatin and heterochromatin. Alterations in chromatin structure by ATP-dependent nucleosome remodelling complexes is thought to activate transcription.

Recently, functions for histone acetylation (HAT) complexes in development have been identified. For example, dCBP, a *Drosophila* homologue of the mammalian HAT proteins CBP and p300, is involved in signalling pathways important for pattern formation (Akimaru et al., 1997a,b; Waltzer and Bienz, 1998). In *C. elegans*, the CBP homologue *cbp-1* is required for the development of all non-neuronal tissues: in its absence, no mesodermal, epidermal or intestinal development occurs and most cells appear to differentiate into neurons (Shi and Mello, 1998).

Developmental roles for histone deacetylases (HDA) have also been identified. For example, one of a number of *C. elegans* histone deacetylases, *hda-1*, antagonizes the effect of *cbp-1* on intestinal differentiation (Shi and Mello, 1998). Further, two *C. elegans* RbAp46/48 homologues, histone-associated proteins which are found in HDA complexes (Taunton et al., 1996) are needed for development past the 100-cell stage, suggesting a shared general cellular function (Shi and Mello, 1998). Recently, the *C. elegans* gene *lin-53* was shown to encode one of these RbAp46/48 homologues (Lu and Horvitz, 1998). LIN-53 is involved in negatively regulating vulval cell fates promoted by the Ras pathway and is likely to act in a multiprotein complex containing the histone deacetlyase HDA1 and the retinoblastoma homolog LIN-35 (Lu and Horvitz, 1998).

In summary, chromatin remodelling complexes have important roles in development, but few of their targets are known. Finding these will help our understanding of how transcription factor activities are regulated to promote particular developmental outcomes.

Possible functions for *egl-27* and *egr-1*

Recently, the human protein MTA1, which is similar to EGL-27 and EGR-1 was found to be a component of a multiprotein complex called NURD, which has both ATP-dependent nucleosome remodelling and histone deacetylation activities (Xue et al., 1998). A similar complex containing an MTA1 related protein, MTA2, was identified by Zhang et al. (1998). Besides MTA1 the NURD complex contains the histone deacetylases HDAC1 and HDAC2, the two histone-binding proteins RbAp48/46, CHD4 (for chromodomain-helicase-DNA binding), and several unidentified proteins. Based on the similarity between EGL-27/EGR-1 and MTA1, and the phenotype of double mutants, we suggest that EGL-27 and EGR-1 are components of a multi-protein chromatin regulatory complex that is required for the functions of regional patterning proteins. Good candidates for interaction with an EGL-27 complex are VAB-7 and the HOX proteins. In the future, it will be important to identify proteins with which EGL-27 and EGR-1 interact, as these should shed light on their biochemical function.

How might an EGL-27-containing complex regulate the activity of these proteins? In one simple model, these transcription factors could associate with the complex and bring it to target genes. This could cause chromatin around the target genes to be altered, leading to changes in transcriptional activity. It is not yet known whether the NURD complex has transcriptional repressing or activating activities, or both; experiments in Xue et al. (1998) suggest a repressive function. However, loss of expression of *hlh-8* and SCM in double mutants suggests an activating function. Genetic and biochemical studies in *C. elegans* involving *egl-27* and *egr-1* should help both to identify further components and to understand the functions of the complex.

We thank Jose de Celis, Richard Durbin, Gos Micklem, Jordan Raff, Daniel St. Johnston and members of the laboratory for helpful comments on the manuscript, and Vivian Bardwell, Erik Miska and Tony Kouzarides for useful discussions. We are grateful to Cara Neades for providing excellent technical assistance. We also thank Howard Baylis, Alan Coulson, Andy Fire, Mike Herman, Jonathan Hodgkin, Yuji Kohara, Michel Labouesse and Cynthia Kenyon for strains, reporter constructs, cDNAs and antibodies. M. Herman, Q. Ch'ng, S. Hettenbach, T. R. Ratliff, C. Kenyon, and R. K. Herman kindly communicated results prior to publication. Some strains used in this study were obtained from the CGC, which is supported by the NIH-NCRR. F. S. was supported by an EMBO fellowship and a European Union Marie-Curie postdoctoral fellowship, A. G. B. is supported by Wellcome Trust grant 048880, and J. A. was supported by a Wellcome Trust Career Development Award (No. 045515/Z/95/Z/PMG/AH).

REFERENCES

Aasland, R. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem. Sci.* **21**, 87-88.

Ahringer, J. (1996). Posterior patterning by the *Caenorhabditis elegans even-skipped* homolog v*ab-7*. *Genes Dev.* **10**, 1120-1130.

Akimaru, H., Chen, Y., Dai, P., Hou, D.-X., Nonaka, M., Smolik, S. M., Armstrong, S., Goodman, R. H. and Ishii, S. (1997a). Drosophila CBP is a co-activator of *cubitus interruptus* in *hedgehog* signalling. *Nature* **386**, 735-738.

Akimaru, H., Hou, D.-X. and Ishii, S. (1997b). Drosophila CBP is required

for dorsal-dependent *twist* gene expression. *Nature Genetics* **17**, 211-214.

Albertson, D. G. (1984). Formation of the first cleavage spindle in nematode embryos. *Dev. Biol.* **101**, 61-72.

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

Bairoch, A., Bucher, P. and Hofmann, K. (1997). The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25**, 217-221

Bowerman, B. (1998). Maternal control of pattern formation in early *Caenorhabditis elegans* embryos. *Curr. Top. Dev. Biol.* **39**, 73-117.

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

Chisholm, A. D. and Horvitz, H. R. (1995). Patterning of the *Caenorhabditis elegans* head region by the *Pax-6* family member *vab-3*. *Nature* **377**, 52-55.

Clark, S. G., Chisholm, A. D. and Horvitz, H. R. (1993). Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*. *Cell* **74**, 43-55.

Costa, M., Weir, M., Coulson, A., Sulston, J. and Kenyon, C. (1988). Posterior pattern formation in *C. elegans* involves position-specific expression of a gene containing a homeobox. *Cell* **55**, 747-756.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. *Nature* **336**, 638-646.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). Potent and specific genetic interference by double- stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.

Fisher, A. L. and Caudy, M. (1998). Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. *Genes Dev.* **12**, 1931-40.

Francis, R. and Waterston, R. H. (1991). Muscle cell attachment in *Caenorhabditis elegans*. *J. Cell Biol.* **114**, 465-479.

Garriga, G., Desai, C. and Horvitz, R. H. (1993). Cell interaction control the direction of outgrowth, branching and fasciculation of the HSN axons of *Caenorhabditis elegans*. *Development* **117**, 1071-1087.

Harfe, B. D., Vaz Gomes, A., Kenyon, C., Liu, J., Krause, M. and Fire, A. (1998). Analysis of a *Caenorhabditis elegans* Twist homolog identifies conserved and divergent aspects of mesodermal patterning. *Genes Dev.* **12**, 2623-2635.

Herman, M. A. and Horvitz, H. R. (1994). The *Caenorhabditis elegans* gene *lin-44* controls the polarity of asymetric cell divisions. *Development* **120**, 1035-1047.

Herman, M. A., Vassilieva, L. L., Horvitz, H. R., Shaw, J. E. and Herman, R. K. (1995). The *C.elegans gene lin-44*, which controls the polarity of certain asymmetric cell divisions, encodes a Wnt protein and acts cell nonautonomously. *Cell* **83**, 101-110.

Herman, M. A., Ch'ng, Q., Hettenbach, S. M., Ratliff, T. R., Kenyon, C. and Herman, R. K. (1999). EGL-27 is similar to a metastasis-associated factor and controls cell polarity and cell migration in *C. elegans*. *Development***126**, 1055-1064.

Horiuchi, J., Silverman, N., Marcus, G. A. and Guarente, L. (1995). ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Mol. Cell. Biol.* **15**, 1203-1209.

Horlein, A. J., Naar, A. M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C. K., et al. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377**, 397-404.

Kadonaga, J. T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin modifying enzymes. *Cell* **92**, 307-313.

Kassavetis, G. A., Nguyen, S. T., Kobayashi, R., Kumar, A., Geiduschek, E. P. and Pisano, M. (1995). Cloning, expression, and function of TFC5, the gene encoding the B″ component of the Saccharomyces cerevisiae RNA polymerase III transcription factor TFIIIB. *Proc. Natl. Acad. Sci. USA* **92**, 9786-9790.

Kelly, W. G., Xu, S., Montgomery, M. K. and Fire, A. (1997). Distinct requirement for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227-238.

Krause, M., Fire, A., White Harrison, S., Priess, J. and Weintraub, H. (1990). CeMyoD accumulation defines the body wall muscle cell fate during *C. elegans* embryogenesis. *Cell* **63**, 907-919.

Labouesse, M., Hartwieg, E. and Horvitz, H. R. (1996). The *Caenorhabditis elegans* LIN-26 protein is required to specify and/or maintain all non neuronal ectodermal cell fates. *Development* **122**, 2579-2588.

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

Mann, R. S. and Affolter, M. (1998). Hox proteins meet more partners. *Curr. Opin. Gen Dev.* **8**, 423-429.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

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

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S. and Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* **79**, 639-648.

Omichinski, J. G., Glore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Apella, E., Stahl, S. J. and Gronenborn, A. M. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261**, 438-446.

Peterson, C. L. and Tamkun, J. W. (1995). The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci.* **20**, 143-146.

Priess, J. R. and Thomson, J. N. (1987). Cellular interactions in early *C.elegans* embryo. *Cell* **48**, 241-250.

Schnabel, R. (1991). Cellular interaction involved in the determination of the early *C. elegans* embryo. *Mech. Dev.* **34**, 85-99.

Seki, N., Ohira, M., Nagase, T., Ishikawa, K., Miyajima, N., Nakajima, D., Nomura, N. and Ohara, O. (1997). Characterization of cDNA clones in size-fractionated cDNA libraries from human brain. *DNA Res.* **4**, 345-349.

Shi, Y. and Mello, C. (1998). A CBP/p300 homolog specifies multiple differentiation pathways in *Caenorhabditis elegans*. *Genes Dev.* **12**, 943-955.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12**, 599-606.

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

Taunton, J., Hassig, C. A. and Schreiber, S. L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3. *Science* **272**, 408-411.

Terns, R. M., Kroll-Conner, P., Zhu, J., Chung, S. and Rothman, J. H. (1997). A deficiency screen for zygotic loci required for establishment and patterning of the epidermis in *Caenorhabditis elegans*. *Genetics* **146**, 185-206.

The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-2018.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-80.

Toh, Y., Pencil, S. D. and Nicolson, G. L. (1995). Analysis of the complete sequence of the novel metastasis-associated candidate gene, *mta1*, differentially expressed in mammary adenocarcinoma and breast cancer cell lines. *Gene* **159**, 97-104.

Trent, C., Tsung, N. and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode *Caenhorhabditis elegans*. *Genetics* **104**, 619-647.

Waltzer, L. and Bienz, M. (1998). Drosophila CBP represses the transcription factor TCF to antagonize Wingless signaling. *Nature* **395**, 521-525.

Wang, B. B., Muller-Immergluck, M. M., Austin, J., Robinson, N. T., Chisholm, A. and Kenyon, C. (1993). A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. *Cell* **74**, 29-42.

Wood, W. B., Ed. (1988). *The Nematode Caenorhabditis elegans.* New York: Cold Spring Harbour Laboratory.

Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J. and Wang, W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Molecular Cell* **2**, 851-861.

Zhang, Y., LeRoy, G., Seelig, H.-P., Lane, W. S. and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a protein complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95**, 279-289.

Zorio, D. A., Cheng, N. N., Blumenthal, T. and Spieth, J. (1994). Operons as a common form of chromosomal organization in *C. elegans*. *Nature* **372**, 270-2.