Development 116, 1001-1010 (1992)
Printed in Great Britain © The Company of Biologists Limited 1992

Coordinate embryonic expression of three zebrafish *engrailed* genes

MARC EKKER*, JEREMY WEGNER, MARIE ANDRÉE AKIMENKO* and MONTE WESTERFIELD

Institute of Neuroscience, University of Oregon, Eugene OR 97403, USA

*Present address: Loeb Institute for Medical Research, Ottawa Civic Hospital, 1053 Carling Avenue, Ottawa, ONT, Canada K1Y 4E9

Summary

We have identified three genes, expressed in zebrafish embryos, that are members of the *engrailed* gene family. On the basis of sequence comparisons and analyses of their expression patterns, we suggest that two of these genes, *eng2* and *eng3*, are closely related to the *En-2* gene of other vertebrates. The third gene, *eng1*, is probably the zebrafish homolog of *En-1*. Subsets of cells at the developing junction between the midbrain and hindbrain express three different combinations of these genes, revealing a previously unknown complexity of this region of the CNS. Other cells, for example, jaw and myotomal muscle precursors, express two of the three genes in combinations which, in the myotomal muscles, change during development. Cells in the developing hindbrain and fins express only a single *engrailed* gene. We propose that the fates and patterning of these cells may be regulated by the coordinate expression of particular combinations of these closely related homeoproteins.

Key words: *Brachydanio rerio*, fin bud, gene expression, hindbrain, homeodomains, jaw muscles, midbrain, muscle pioneers, somites.

Introduction

Transcription factors, including homeoproteins, are known to regulate the development of cell fates and patterning in embryos (Ingham, 1988). By studying the structures and expression patterns of genes coding for homeoproteins, we can begin to learn about the genetic regulation of patterning during development.

In both vertebrates (Fjose et al., 1988; Joyner et al., 1985) and invertebrates (Fjose et al., 1985; Kornberg et al., 1985), *engrailed* is one of the best characterized homeoproteins. Mutational analyses in *Drosophila* have demonstrated that *engrailed* plays multiple roles during development. For example, *engrailed* functions in segment formation (Kornberg, 1981a,b; Nüsslein-Volhard and Wieschaus, 1980), the determination of segment identity (Garcia-Bellido, 1975; Lawrence and Morata, 1976), preblastoderm organization (Karr et al., 1985), compartment formation (Morata and Lawrence, 1975; Kronberg, 1981a,b; Lawrence and Struhl, 1982) and neurogenesis (DiNardo et al., 1985; Brower, 1986). Analyses of *engrailed* expression in vertebrates also supports the notion of multiple roles for this type of homeoprotein, and most vertebrates analyzed to date express two *engrailed* genes in diverse and complicated patterns (reviewed by Joyner and Hanks, 1992).

In a previous study, we analyzed *engrailed* expression in developing and adult zebrafish and found a remarkable complexity of Engrailed proteins (Hatta et al., 1991). Many cell types are reproducibly recognized by two antibodies generated against the Engrailed protein, but other cells are recognized by only one or the other. Moreover, our analysis showed that the antibodies recognize proteins of several different sizes, suggesting that there may be more than two *engrailed* genes in zebrafish, although only two *engrailed* genes had been previously identified (Fjose et al., 1988; Holland and Williams, 1990).

In this report, we provide direct evidence that the zebrafish genome contains at least three distinct *engrailed* genes. The expression pattern of one of these genes, *eng1*, is reminiscent of that of the mouse *En-1* gene (Davis and Joyner, 1988; Davidson et al., 1988), whereas the structures and expression patterns of the other two genes, *eng2* (originally called En-2, Fjose et al., 1988) and *eng3* (designated *En-1* by Holland and Williams, 1990) are more similar to those of the *En-2* gene from mouse (Joyner and Martin, 1987; Davis et al., 1988; Davis and Joyner, 1988). From analysis of the gene products translated in vitro, we show that the two antibodies used in the previous studies recognize different combinations of the Engrailed proteins, thus explaining the complexity of their labeling patterns in embryos. Using gene-specific probes for in situ hybridization, we have discovered that a region of the CNS, the presumptive border between the midbrain and hindbrain, is subdivided into three discrete regions of gene expression before overt signs of morphological differentiation have occurred. Our studies also reveal that a special class of

muscle cells in the myotomes, the muscle pioneers, express two of the *engrailed* genes as they start to differentiate but express only one a few hours later. Thus, cells in diverse regions of the embryo precisely express particular combinations of the three *engrailed* genes, and these combinations change during differentiation of individual cells.

Materials and methods

Animals

Embryos from the Oregon AB line were maintained using standard methods (Westerfield, 1989) and were staged at 28.5°C according to hours (h) and days (d) postfertilization.

cDNA and genomic libraries

Two λZAP II cDNA libraries prepared from 9-16 h and 20-28 h zebrafish mRNA, kindly provided by D. J. Grunwald, were screened using as a probe, the 905 bp *Eco*RI fragment amplified from 20-28 h zebrafish cDNA by the polymerase chain reaction (PCR; see below: Cloning of *eng3* and Fig. 3).

DNA sequencing

Restriction fragments of the zebrafish *eng1*, *eng2* and *eng3* cDNA clones were isolated and subcloned into Bluescript phagemids (Stratagene). Single-strand templates were prepared from these phagemids according to the manufacturer's instructions. DNA sequencing was performed by the dideoxy-termination method using Sequenase (USB, Inc.), according to the manufacturer's directions.

In situ hybridization

Zebrafish embryos were fixed in 4% paraformaldehyde in PBS (140 mM NaCl, 3 mM KCl, 10 mM sodium phosphate, pH 7.2) overnight at 4°C followed by two 5 minute washes in PBS. For hybridization on sections, the embryos were embedded in 1.5% agar in 30% sucrose, sectioned on a cryostat and mounted on gelatin-subbed slides. Sections (16 μm) were fixed for 5 minutes in 4% paraformaldehyde in PBS, washed twice for 5 minutes in PBS, and dehydrated through an ethanol series. They were kept over desiccant at 4°C for up to six months.

In situ hybridization was performed as previously described for sections (Wilkinson et al., 1987; Sassoon et al., 1988; Akimenko, unpublished data) and for whole-mount embryos (Püschel et al., 1992). Similar results were obtained with the two methods although the sensitivity using sections was slightly greater whereas the resolution was better with the whole mounts. The data presented were obtained using both methods.

The following cDNA fragments were used to synthesize probes for in situ hybridization: *eng1*, the 520 *Eco*RI-*Bam*HI fragment (Fig. 1); *eng2*, the 249 bp *Dra*I fragment (Fig. 2); *eng3*, the 905 bp *Eco*RI fragment obtained by PCR amplification (Fig. 3). Antisense riboprobes were synthesized with T7 RNA polymerase in the presence of ^{35}S-UTP for hybridization on sections or in the presence of digoxigenin-labeled UTP for hybridization of whole mounts. These probes were chosen, on the basis of sequence comparisons (Figs 1-3) to minimize cross hybridization among the three genes. Southern and northern blots (not shown) and the unique in situ hybridization patterns (e.g. Table 1) confirmed the specificity of these probes.

Western analysis

Transcripts (1 μg) of the three *engrailed* genes, synthesized with an mRNA capping kit (Stratagene), were translated using a rabbit reticulocyte lysate (Promega). One third of each translation reaction, mixed with 10% glycerol, 3.5% SDS, 2% β-mercaptoethanol, and 0.0001% pyronin, was boiled for 5 minutes. Insoluble particles were then removed by centrifugation and the translation products were separated from other proteins in the lysate by electrophoresis on a 12% polyacrylamide gel and transferred to PVDF blotting paper (Millipore) according to Towbin et al. (1979). The blots were presoaked in blotto (5% skim milk in PBS) for more than an hour and then incubated in primary antibody (either the 4D9 monoclonal antibody, Patel et al., 1989, or the α*Enhb-1* polyclonal antibody, Davis et al., 1991) in blotto at 4°C overnight. They were then washed in PBS, incubated with the appropriate alkaline phosphatase secondary antibody (Biorad) in blotto for 1 hour, and washed in PBS. The immunoreactive bands were visualized by staining for alkaline phosphatase activity with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (Sigma).

Results

Structural analyses of three engrailed *genes*

Cloning of eng1 *and* eng2

The *eng1* and *eng2* cDNAs were obtained from suspensions of three λgt-11 phage clones kindly provided by Scott Smiley. One of the phage had an insert of 1.8 kb. It encodes a protein of 231 amino acids with a homeodomain similar to that of other Engrailed-like gene products. We have named this gene *eng1*. The coding region of the *eng1* cDNA clone is preceded by an untranslated region of 243 bp and is followed by a 3′-untranslated region of 872 bp with a polyadenylation signal at position 1780 (Fig. 1).

The other two phage contained inserts of 2.6 kb corresponding to the transcript of the *eng2* gene whose homeobox is identical to that of the zebrafish gene previously designated *En-2* (Fjose et al., 1988; see Njølstad et al., 1990, for gene name conventions). The sequence of one of these clones (Fig. 2) has an open reading frame of 795 bp that encodes a homeoprotein of 265 amino acids. The open reading frame is preceded by an untranslated sequence of 42 bp and is followed by a 3′-untranslated region of 1770 bp which ends with a polyadenylation signal at position 2600.

Cloning of eng3

We isolated cDNAs for the zebrafish *eng3* gene using a probe made by PCR amplification. We synthesized an oligonucleotide PCR primer, oriented facing the 5′-end of the *eng3* mRNA, based on part of the zebrafish *eng3* homeobox sequence (designated *En-1* by Holland and Williams, 1990; positions 893 to 912 in Fig. 3). We used DNA from a λ ZAP II cDNA library made from 20-28 h zebrafish embryos (provided by D. J. Grunwald) for PCR amplification with the *eng3* primer and a second primer corresponding to a sequence from the vector (SK primer, Stratagene). We obtained an amplified fragment of 905 bp and sequenced it to confirm that it corresponded to an *eng3* cDNA. We then used this fragment to screen cDNA libraries made from 9-16 h and 20-28 h zebrafish embryos. We screened 300,000 plaques and isolated two positive clones that corresponded to the *eng3* gene.

The 5′-end of the longer cDNA clone is 63 bp downstream from the 5′-end of the PCR-amplified fragment. The remaining sequences of the clones that overlap the PCR

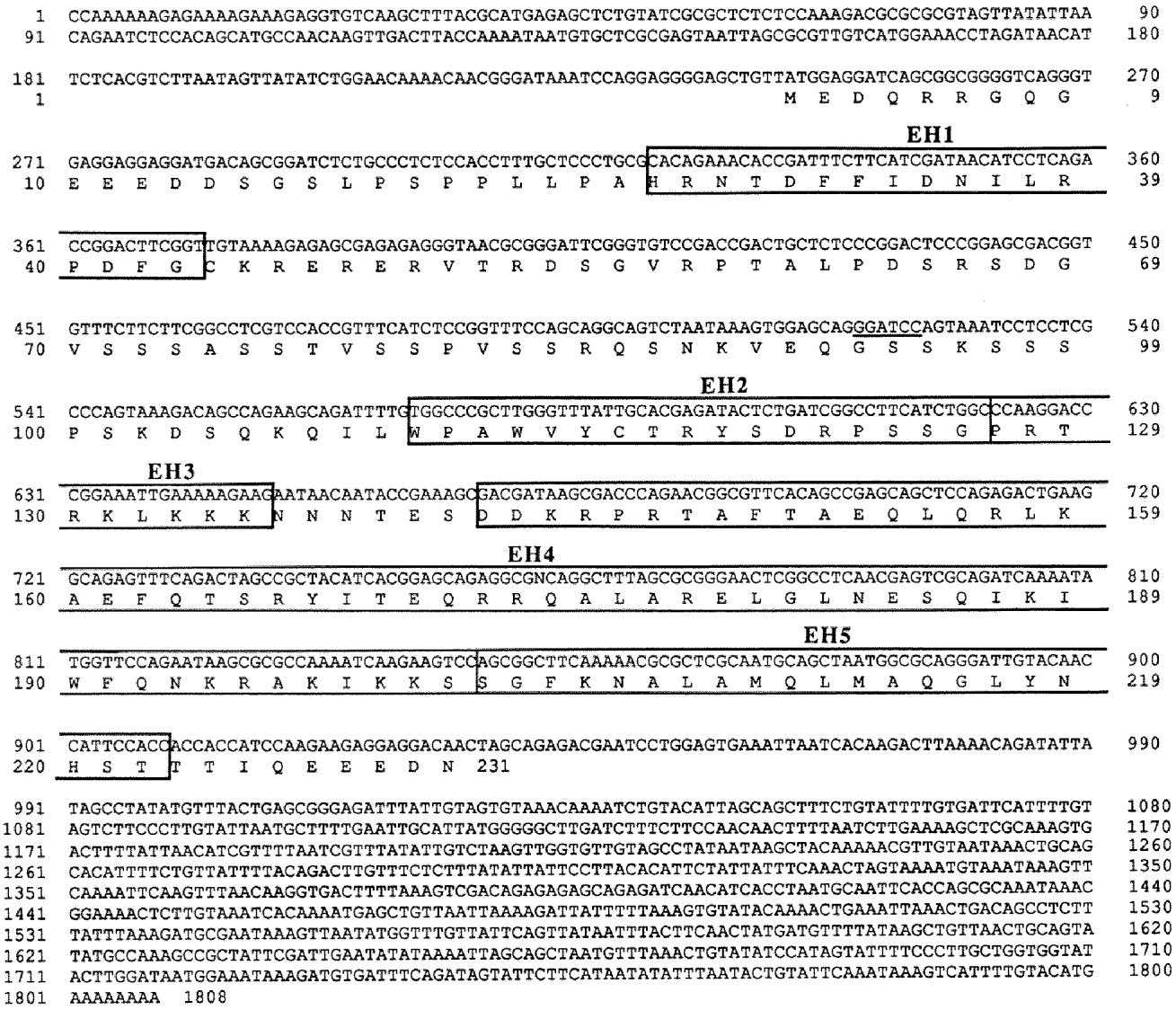

Fig. 1. Sequence of the zebrafish *eng1* gene. The nucleotide and the predicted amino acid sequences of the *eng1* gene are shown. The conserved regions (EH1-5) are enclosed in boxes. The homeobox is contained within EH4. The probe used for in situ hybridization was made from the sequences between the *Eco*RI cloning site immediately upstream of nucleotide 1 and the *Bam*HI site (underlined) at position 520.

fragment are 100% colinear with each other and with the fragment. The combined sequences from the cDNA clones and the PCR fragment are shown in Fig. 3 and predict that the *eng3* protein is 261 amino acids. The coding region is preceded by an untranslated region of 337 bp and is followed by a 3′-untranslated region of 803 bp with a polyadenylation signal at position 1794.

Amino acid sequence similarities of vertebrate engrailed *gene products*

The homeodomains predicted from the three zebrafish *engrailed* genes show a high degree of sequence similarity (> 90%) to the engrailed homeodomains of other vertebrate species (Fig. 4). Furthermore, the three zebrafish Engrailed proteins contain the additional four domains conserved among all known Engrailed proteins (Joyner and Hanks, 1992). The first domain, EH1, is located in the N-terminal region of the protein whereas the other four domains, EH2-5, are found in a 111-113 amino acid stretch that includes the homeodomain (EH4) (Fig. 4).

Sequences outside the homeodomains of the zebrafish Engrailed proteins are also related to the corresponding regions of their homologues in other species. This is illustrated in Fig. 4A which shows an alignment of the amino acid sequences of the Engrailed gene products from several species. The two proteins that show the highest degree of sequence identity (69%) are zebrafish Eng2 and Eng3. These two proteins are 79% similar, considering conservative amino acid substitutions. Both the zebrafish Eng2 and Eng3 proteins have more amino acid sequence identities with the En-2 proteins of mouse (63% for Eng2 and 67% for Eng3; Logan et al., unpublished data) and *Xenopus*

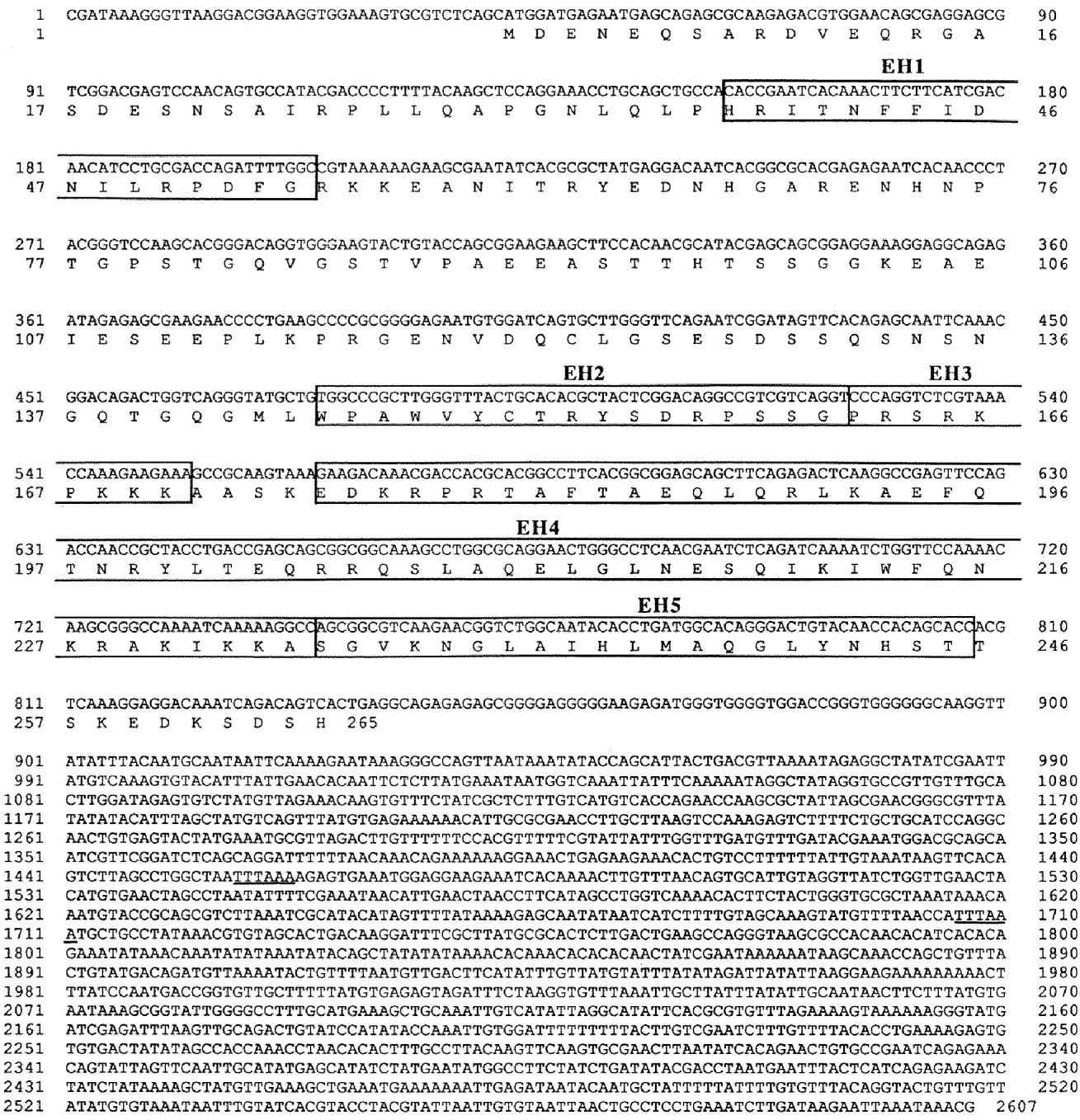

Fig. 2. Sequence of the zebrafish *eng2* gene. The nucleotide and the predicted amino acid sequences of the *eng2* gene are shown. The conserved regions (EH1-5) are enclosed in boxes. The homeobox is contained within EH4. The probe used for in situ hybridization was made from the sequences between the *Dra*I restriction sites at positions 1457 and 1706 (underlined).

(67% for Eng2; 66% for Eng3) than they do with the mouse En-1 protein (52% for Eng2; 54% for Eng3). This greater resemblance to the En-2 proteins reflects a better conservation of the sequences outside the highly conserved EH1-5 domains. This interpretation was further strengthened when we extended our comparisons to the 3′-untranslated regions of the engrailed cDNAs (Fig. 4B). We found a stretch of about 45 bp that is well conserved among the zebrafish *eng2, eng3* and *Xenopus En-2* cDNAs. A second stretch of about 20 bp is conserved between *Xenopus En-2* and *eng3*. These conserved regions are absent from the zebrafish *eng1* sequence (Fig. 1) and from the mouse and chicken *En-1* sequences (Logan et al., unpublished data).

The zebrafish engrailed *proteins*

The three *engrailed* cDNAs encode proteins recognized by the anti-engrailed antibodies. We expressed the proteins in vitro and analyzed them on western blots (Fig. 5). The 4D9

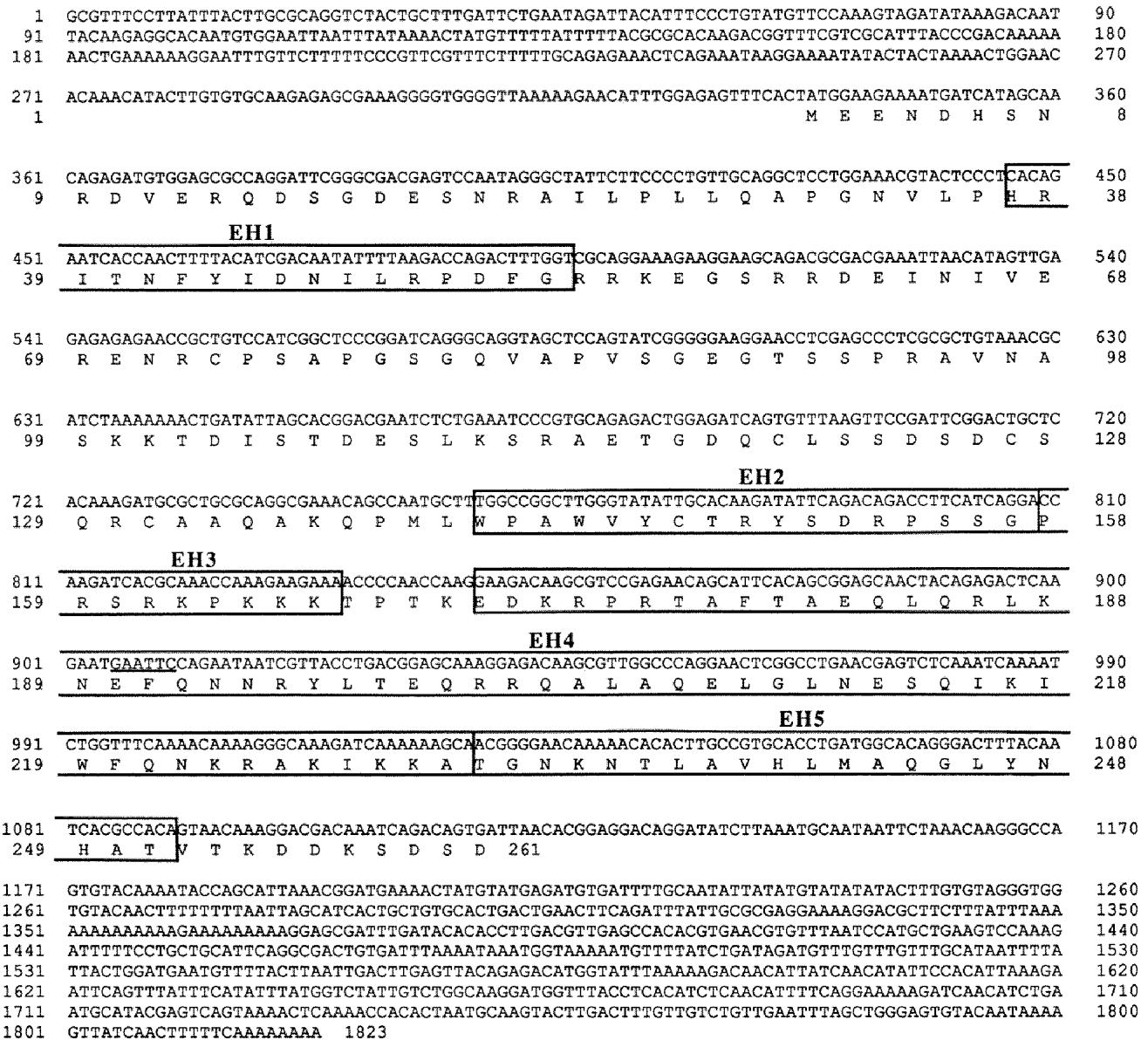

Fig. 3. Sequence of the zebrafish *eng3* gene. The nucleotide and the predicted amino acid sequences of the *eng3* gene are shown. The conserved regions (EH1-5) are enclosed in boxes. The homeobox is contained within EH4. The probe used for in situ hybridization was made from the sequences between the *Eco*RI cloning site immediately upstream of nucleotide 1 and the *Eco*RI site at position 905 (underlined).

monoclonal antibody recognized all three gene products (4D9 lanes 1-3, Fig. 5). In contrast, the α*Enhb-1* polyclonal antibody primarily recognized only the Eng2 and Eng3 gene products. The apparent relative molecular masses of the bands recognized by the 4D9 antibody corresponded to 32, 43 and 38×10^3, for Eng1, Eng2, and Eng3, respectively, and are slightly higher than those predicted from the deduced amino acid sequences (25, 29, and 28×10^3 M_r, respectively).

The in vitro translation products of the three *engrailed* genes can be correlated with the bands we previously detected on western blots of embryo extracts (Hatta et al., 1991). The 4D9 antibody recognized three bands in proteins from solubilized embryos at 39, 47, and 41×10^3 M_r. Because these bands are similar to those of the in vitro translation products (Fig. 5), they probably correspond to the products of the endogenous *eng1*, *eng2*, and *eng3* genes, respectively. The heavier apparent relative molecular masses in vivo could indicate post-translational modifications.

Analyses of engrailed *gene expression*

Midbrain-hindbrain border

Before formation of the border. At 12 h, *engrailed* expression in the CNS is restricted to a stripe of cells across the brain primordium at the presumptive border between the midbrain and hindbrain, as previously reported (Njølstad and Fjose, 1988; Patel et al., 1989; Hatta et al., 1991).

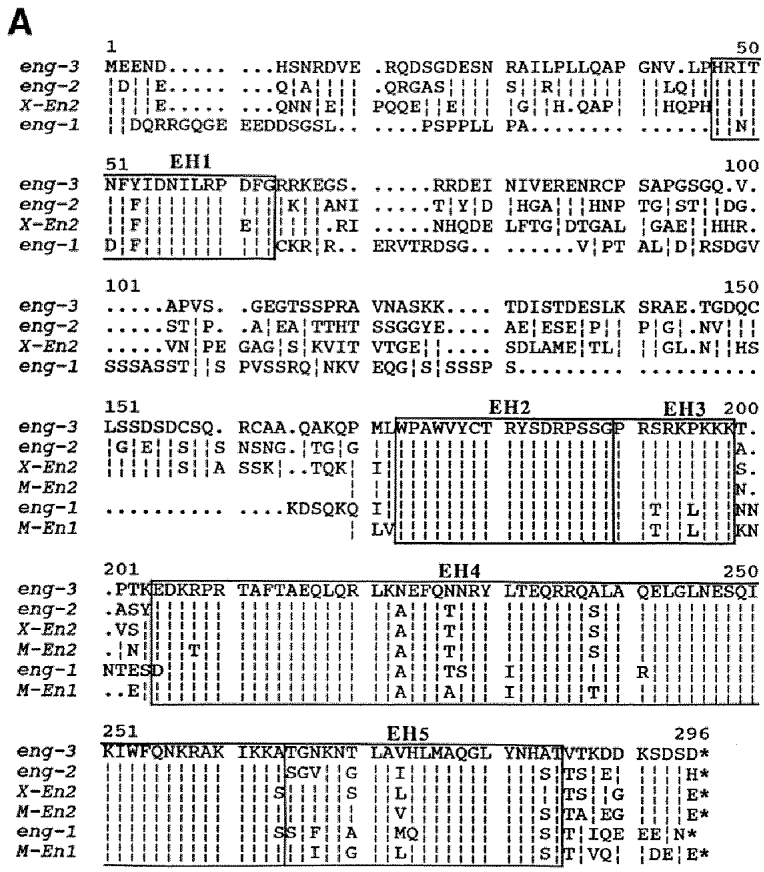

Fig. 4. Comparisons of *engrailed* genes. (A) The amino acid sequences encoded by the *engrailed* genes are aligned. We have also included in this analysis the mouse En-1 and En-2 proteins (Joyner and Martin, 1987) and the product of the *Xenopus* 2.7 kb *X-En2* cDNA (Hemmati-Brivanlou et al., 1991). The conserved regions (EH1-5) are enclosed in boxes. The homeobox is contained within EH4. Amino acid identities are indicated by vertical bars, spaces introduced for alignment are indicated by dots. * indicates the end of the open reading frame. (B) Alignments of conserved sequences in the 3′-untranslated regions of the zebrafish *eng2*, *eng3* and *Xenopus X-En2* genes. The sequences from positions 14 to 67 are relatively well conserved among all three genes. The zebrafish *eng3* and *Xenopus X-En2* genes share an additional sequence of 20 bp (box) which is absent from the *eng1* gene.

All three probes hybridize at this stage and the pattern for each is distinct (Fig. 6A-C).

The *eng1* transcripts appear in a dorsal stripe of cells. The stripe is 8-10 cells wide at its dorsal aspect and narrows as it extends towards the ventral surface of the brain (Fig. 6A). The *eng2* and *eng3* transcripts occupy respectively larger regions that include the cells expressing *eng1* (Fig. 6B,C, respectively). The *eng3* transcripts occupy a wedge-shaped region extending from the dorsal to the ventral surface of the brain, with the broader end at the dorsal aspect. The wedge is more than two dozen cells wide at the dorsal surface and tapers to about half as many cells near the ventral surface. Expression is heavy and probably includes most cells. The *eng2* wedge is slightly smaller than that of *eng3* and, like *eng3*, extends completely between the dorsal and ventral surfaces.

After formation of the border. Later, when the furrow that forms the midbrain-hindbrain border is visible, transcripts from each gene still occupy distinct, but overlapping regions (Fig. 6E-H). The *eng1* transcripts appear over the most restricted subset of cells expressing the other two genes. The stripe of cells expressing *eng1* is located precisely where the furrow between the midbrain and hindbrain has formed and includes approximately 8-10 cells lining the posterior and ventral walls of the furrow (Fig. 6E). There is little or no expression of *eng1* by cells anterior to the furrow.

The *eng2* transcripts appear in the cells that express *eng1* and in additional neighboring cells (Fig. 6F). The stripe of cells expressing *eng2* transcripts extends across the width and depth of the brain in this region and narrows slightly at its ventral extent. Anteriorly, *eng2* tran-

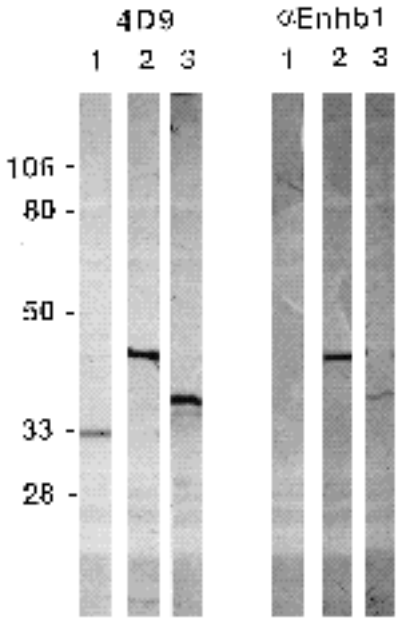

Fig. 5. The zebrafish *engrailed* genes encode proteins recognized by the 4D9 and α*Enhb-1* antibodies. The proteins, translated in vitro, from the *eng1* (1), *eng2* (2), or *eng3* (3) genes were analyzed by gel electrophoresis and probed with either the 4D9 monoclonal (left) or α*Enhb-1* polyclonal (right) antibodies. The apparent relative molecular masses ($\times 10^{-3}$) are indicated on the left.

scripts extend into the posterior wall of the midbrain. Posteriorly, cells throughout the presumptive cerebellum express *eng2*.

The *eng3* transcripts are still the most broadly distributed, extending from the posterior part of the midbrain through the presumptive cerebellum to a position posterior to both the *eng1*- and *eng2*-expression domains (Fig. 6G,H). Expression is heavy and uniform and still includes the entire width and dorsal-ventral extent of the brain in this region, suggesting that most, if not all, cells express *eng3*.

Muscle pioneers

A subset of muscle cells in the myotomes expresses two of the *engrailed* genes, *eng1* and *eng2*, during somitogenesis, although *eng2* is only transiently expressed (Fig. 7). Somites are added at a rate of approximately 2 per hour (Hanneman and Westerfield, 1989) and both *engrailed* transcripts first appear two to four somites anterior to the youngest somite, correlating well with the previously observed onset of Engrailed protein synthesis (Hatta et al., 1991). By 16 h, each of the posterior somites that hybridizes with the *eng1* and *eng2* probes has a patch of hybridization in the middle of the myotome close to its anterior margin. Developmentally older somites, located in more anterior segments, have longitudinally oriented bands of *eng1* and *eng2* transcripts centered in the middle of each myotome (Fig. 7A,B). The relative sizes and intensities of these hybridizing patches are similar with the two probes.

The patches of transcripts are localized over 2-6 cells at the medial surface of the myotome facing the notochord. These *eng1*- and *eng2*-expressing cells include the muscle pioneers that we identify on the bases of their early development, shapes and positions (Felsenfeld et al., 1991) and that are known to produce Engrailed proteins (Hatta et al., 1991). By 24 h, most of the somites have formed and *eng2* transcripts have disappeared from the anterior myotomes (Fig. 7D), although they are still localized over muscle pioneers in the younger, posterior myotomes. At 32 h, *eng2* transcripts are undetectable in the myotomes by in situ hybridization, although *eng1* transcripts are still abundant. The *eng1* transcripts persist in the muscles beyond this stage (Fig. 7E) and throughout development of the septum that forms at this location and divides the myotome into the dorsal and ventral muscles (Westerfield et al., 1986).

Hindbrain and jaw

Small clusters of cells in the hindbrain and in the mesenchyme just posterior to the eye express the *eng2* and *eng3* genes. In the hindbrain, *eng3* transcripts appear in discrete patches localized over clusters of about a dozen cells each. Hybridization is weak at 24 h, but by 32 h, segmentally iterated clusters of cells in the ventral hindbrain hybridize with the *eng3* probe (Figs 6H, 8). The clusters appear only in the three anterior rhombomeres. They are bilaterally symmetrical and approximately centered in each rhombomere.

By 32 h, a loosely associated cluster of mesenchymal cells posterior to the eye hybridizes with the *eng2* and *eng3* probes (Fig. 9), but not with the probe for *eng1* transcripts. The size and position of this cluster and the morphology of the cells suggests that these *eng2*- and *eng3*-expressing cells are probably the precursors of jaw muscles that were shown in a previous study (Hatta et al., 1990) to contain Engrailed immunoreactivity.

Fins

Epidermal cells in the pectoral fin buds express *eng1* but not the other two *engrailed* genes (Fig. 10). Cells on the ventral-anterior surface of the buds have a strong hybridization signal by 32 h, as do epidermal cells in the body wall anterior to the buds.

The epidermal cells in the fin bud that express *eng1* contribute to the ventral surface of the fin (Hatta et al., 1991). They are later separated from dorsal cells of the fin by the formation of a ridge, equivalent to the apical ectodermal ridge of other vertebrates (Wood, 1982).

Discussion

Structural relationships among the engrailed *genes*

Our results demonstrate that the zebrafish genome contains at least three *engrailed* genes and that these genes are expressed in distinct patterns during embryonic development. Other vertebrate species including hagfish (Holland and Williams, 1990), *Xenopus* (Hemmati-Brivanlou et al., 1991; Holland and Williams, 1990), chicken (Davis et al., 1991), mouse (Joyner et al., 1985; Joyner and Martin, 1987) and human (Logan et al., 1989) are known to have two *engrailed* genes. Our structural analyses suggest relationships among the three zebrafish genes and the two *engrailed* genes of other vertebrates. Amino acid and nucleotide sequence comparisons (Figs 1-4) suggest that the zebrafish *eng2* and *eng3* genes are more closely related to each other and to the *En-2* genes of other vertebrates than to *En-1*. Similarly, the zebrafish *eng1* gene is less similar to *En-2* than are *eng2* and *eng3*, and like *En-1*, it is missing the 3′-untranslated sequences present in the *eng2*, *eng3* and *En-2* genes. Thus, we suggest that the zebrafish *eng1* is probably related to *En-1* and that the *eng2* and *eng3* genes may

Table 1. *Coordinate expression of three* engrailed *genes*

	Riboprobes			Antibodies	
	eng1	eng2	eng3	αEnhb-1	4D9
midbrain/ hindbrain border	z,m	z,m	z	z,m	z
anterior hindbrain	m	–	z	z	z
jaw	–	z	z	z,m	z
somites: myotome/	z	z	–	z	z
dermatome	m	–	–	m	–
fin/limb bud	z,m	–	–	m	z
posterior hindbrain and spinal cord	m	–	–	z,m	–

(Riboprobes) Tissues are listed in which transcripts of each of the three *engrailed* genes have (z) or have not (−) been detected in zebrafish embryos. Tissues expressing the mouse *En-1* or *En-2* genes (Joyner and Hanks, 1992) are indicated by (m).

(Antibodies) Zebrafish (z; Hatta et al., 1991) and mouse (m; Davis et al., 1991) tissues labeled by the 4D9 monoclonal antibody or the αEnhb-1 polyclonal antibody are indicated. The α-Enhb-1 antibody recognizes both the En-1 and En-2 proteins whereas the 4D9 antibody does not recognize Engrailed proteins in mice (Davis et al., 1991).

be the products of a duplication of an ancestral *En-2* gene. Full-length sequences of the *engrailed* genes from additional species will help resolve this issue. Similarly, it will be important to learn whether there are more than two *engrailed* genes in other species.

Development of some regions of the embryo, like the junction between the midbrain and hindbrain, involves coordinate expression of several *engrailed* genes. Specific deletion of the *En-2* homeobox in mice produces only a very subtle change in the development of this region of the central nervous system, which has led to the suggestion that there may be functional redundancy of the two *engrailed* genes (Joyner et al., 1991). Our results, however, demonstrate that some structures express a single *engrailed* gene, like *eng1* in the pectoral fin buds (Fig. 10) or *eng3* in the hindbrain (Fig. 8). Thus, although these genes are structurally related, the specificity of their expression patterns (Table 1) suggests that they subserve distinct functions during embryonic development.

The complexity of *engrailed* expression, even at the junction between the midbrain and hindbrain, is consistent with this view. Although the expression domains of the three genes overlap, each gene is expressed in a distinct region (Fig. 6). In a previous study, we noted that the *eng2*-expressing region is centered between the midbrain and hindbrain expression domains of the *pax6* gene (Püschel et al., 1992). The *eng2*-expressing cells are flanked by cells which express neither gene. Our present analyses suggest that these flanking cells are probably the cells that express *eng3* (Fig. 6). Thus, these four genes define a remarkably complex subdivision of this CNS region. This distinct pattern of gene expression is apparent at approximately the time Engrailed proteins first appear (Hatta et al., 1991) Thus, expression of the three *engrailed* genes appears to be coordinately regulated from the outset.

The eng3 *gene product and identities of hindbrain cells and jaw muscles*

Hatta et al. (1990) suggested that Engrailed expression in the LAP and DO jaw muscle precursors may be involved in determining their identities, including their innervation by specific hindbrain motoneurons. They also reported that cells in the vicinity of the motoneurons, but not the motoneurons themselves, express Engrailed. In the present study, we have demonstrated that these two sets of cells express the same *engrailed* gene, *eng3*, and that the precursors of one or both jaw muscles additionally express *eng2* (Fig. 9). The Eng3 gene product may, thus, be involved in determining the identity of the cells with which the motoneurons interact in both the central nervous system and in the periphery. For example, the Eng3 product may regulate expression of factors which are recognized by the motoneurons and which are involved in specifying their synaptic connections. Alternatively, the expression of *eng3* alone by the anterior hindbrain cells may relate to the specification of a particular brain structure.

In a previous study (Hatta et al., 1991), we detected Engrailed immunoreactivity in the posterior hindbrain and the spinal cord with a polyclonal antibody, α*Enhb-1* (Davis et al., 1991) but not with the 4D9 monoclonal antibody. In the present study, we were unable to observe hybridization in these regions of the CNS with any of the specific *engrailed* probes. This discrepancy may be due to a lower sensitivity of the in situ hybridization method, compared to antibody staining. The CNS cells recognized by the α*Enhb-1* antibody are separated from one another. Hybridization signals in individual cells might be difficult to discern, although *En-1* expression has been detected in the spinal cords of mice with this technique (Davidson et al., 1988; Davis and Joyner, 1988). However, the α*Enhb-1* polyclonal antibody may recognize molecules in addition to the three Engrailed proteins. Our western analysis (Fig. 5) demonstrates that 4D9 recognizes all three Engrailed gene products and that α*Enhb-1* recognizes the in vitro translation products of the *eng2* and *eng3* genes. We also know from our previous work (Hatta et al., 1991) that α*Enhb-1* may recognize additional proteins in embryo extracts. Thus, the spinal cord and posterior hindbrain cells labeled by α*Enhb-1* may express a fourth Engrailed gene product or other proteins that differ from Engrailed but that have common antibody binding epitopes.

We know from our previous work that cells on the ventral surface of the pectoral fin bud produce Engrailed proteins that 4D9, but not the α*Enhb-1* antibody, recognizes. Our structural analyses of the three *engrailed* genes may provide an explanation for this observation; the fins express high levels of primarily the Eng1 gene product (Fig. 10), and α*Enhb-1* recognizes Eng1 poorly, if at all (Fig. 5).

The engrailed *gene products and borders in the brain and myotomes*

In a previous study (Hatta et al., 1991), we noted that, in two independent systems, expression of the *engrailed* genes is associated with subdivisions of organ primordia. At both the border between the midbrain and hindbrain and the border between the dorsal and ventral muscles in the myotomes, cells express Engrailed where the furrow that

later separates these regions appears. Our findings of the present study extend this notion; expression of one of the three genes, *eng1*, precisely defines the future locations of these furrows (Figs 6, 7) and could be important in specifying their positions. Because the 4D9 monoclonal antibody recognizes the products of all three genes (Fig. 5), we could not previously distinguish the more restricted expression domain of *eng1* from those of the other two *engrailed* genes.

At the future midbrain-hindbrain border, which is later defined by the furrow, a very narrow stripe, about a half dozen cells wide, expresses *eng1* before morphological differentiation of the furrow (Fig. 6A). Later, one side of the furrow is lined by a narrow band of cells that continue to express *eng1* (Fig. 6E). Although these cells at the border between the midbrain and hindbrain also express the *eng2* and *eng3* genes, other cells located farther from the furrow also express these two genes. Thus, *eng1* is a likely candidate to be involved in establishing the midbrain-hindbrain border whereas expression of the *eng2* and *eng3* genes may specify other aspects of development in this region of the CNS. Alternatively, the combinatory expression of these genes may specify position. In adult mice, cells in a number of motor nuclei within the pons and cells in the substantia nigra coordinately express *En-1* and *En-2*, while cells in the granule cell layer of the cerebellum express *En-2* alone (Davis et al., 1988; Davis and Joyner, 1988).

In the myotomes, a subset of muscle cells, the muscle pioneers (Felsenfeld et al., 1991), expresses *eng1*, and *eng2* transiently, in the location where a furrow later forms (Fig. 7). The furrow eventually develops into the horizontal myoseptum, a connective tissue structure that separates the dorsal and ventral muscles of the myotomes (Westerfield et al., 1986). The muscle pioneers are thought to participate in the formation of the horizontal myoseptum, based on observations of their morphogenesis (Hatta et al., 1991) and their absence in *spt* mutants, which also lack horizontal myosepta (Kimmel et al., 1989). As in the CNS, expression of *eng1* by cells in the future location of the furrow begins before overt morphological differentiation of the furrow, consistent with this gene's potential role in specification of that position.

We wish to thank Walter Gehring, in whose lab this work began, Ruth Bremiller for technical assistance, Alex Joyner for sharing unpublished data, and Kohei Hatta and Charles Kimmel for critical reading of the manuscript. Supported by the NIH NS21132 and HD22486 and the McKnight Foundation. M. E. was supported by a Centennial Fellowship from the MRC of Canada. M.-A. A. acknowledges support from the Institut Pasteur de Paris, the French Ministere de la Recherche et de la Technologie and NATO.

References

Brower, D. (1986). *Engrailed* gene expression in *Drosophila* imaginal discs. *EMBO J.* **5**, 2649-2656.

Davidson, D., Graham, E., Sime, C. and Hill, R. (1988). A gene with sequence similarity to *Drosophila engrailed* is expressed during the development of the neural tube and vertebrae in the mouse. *Development* **104**, 305-316.

Davis, C. A., Holmyard, D. P., Millen, K. J. and Joyner, A. L. (1991). Examining pattern formation in mouse, chicken and frog embryos with an *En*-specific antiserum. *Development* **111**, 287-298.

Davis, C. A., Noble-Topham, S. E., Rossant, J., and Joyner, A. L. (1988). Expression of the homeobox containing gene *En-2* delineates a specific region of the developing mouse brain. *Genes Dev.* **2**, 361-371.

Davis, C. A. and Joyner, A. L. (1988). Expression patterns of the homeobox-containing genes *En-1* and *En-2* and the proto-oncogene *int-1* diverge during mouse development. *Genes Dev.* **2**, 1736-1744.

DiNardo, S., Kuner, J. M., Theis, J., and O'Farrell, P. H. (1985). Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear *engrailed* protein. *Cell* **43**, 59-69.

Felsenfeld, A. L., Curry, M. and Kimmel, C. B. (1991). The *fub-1* mutation blocks initial myofibril formation in zebrafish muscle pioneer cells. *Dev. Biol.* **148**, 23-30.

Fjose, A., Eiken, H. G., Njølstad, P. R., Molven, A. and Hordvik, I. (1988). A zebrafish engrailed-like homeobox sequence expressed during embryogenesis. *FEBS Lett.* **181**, 173-182.

Fjose, A., McGinnis, W. J., and Gehring, W. J. (1985). Isolation of a homoeobox-containing gene from the *engrailed* region of *Drosophila* and the spatial distribution of its transcripts. *Nature* **313**, 284-289.

Garcia-Bellido, A. (1975). Genetic control of wing disc development in *Drosophila*. In *Cell Patterning*, pp. 161-182. Amsterdam: Elsevier.

Hanneman, E. and Westerfield, M. (1989). Early expression of acetylcholinesterase activity in functionally distinct neurons of the zebrafish. *J. Comp. Neurol.* **284**, 350-361.

Hatta, K., Schilling, T. F., BreMiller, R. A., and Kimmel, C. B. (1990). Specification of jaw muscle identity in zebrafish: Correlation with *Engrailed*-homeoprotein expression. *Science* **250**, 802-805.

Hatta, K., Bremiller, R., Westerfield, M. and Kimmel, C. B. (1991). Diversity of expression of *engrailed*-like antigens in zebrafish. *Development* **112,** 821-832.

Hemmati-Brivanlou, A., de la Torre, J. R., Holt, C. and Harland, R. M. (1991). Cephalic expression and molecular characterization of *Xenopus En-2*. *Development* **111**, 715-724.

Holland, P. W. H. and Williams, N. A. (1990). Conservation of *engrailed*-like homeobox sequences during vertebrate evolution. *FEBS Lett.* **277**, 250-252.

Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in Drosophila. *Nature* **335**, 25-34.

Joyner, A. L. and Hanks, M. (1992). The *engrailed* genes: Do they play a conserved role in maintaining developmental domains? *Sem. Dev. Biol.* (in press).

Joyner, A. L. and Martin, G. R. (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Genes Dev.* **1**, 29-38.

Joyner, A., Herrup, K., Auerbach, A., Davis, C. A. and Rossant, J. (1991). Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the *En-2* homeobox. *Science* **252**, 1239-1243.

Joyner, A., Kornberg, T., Coleman, K. G., Cox, D. and Martin, G. R. (1985). Expression during embryogenesis of a mouse gene with sequence homology to the Drosophila engrailed gene. *Cell* **43**, 29-37.

Joyner, A. L. and Martin, G. R. (1987). En-1 and En-2, two mouse genes with sequence homology to the Drosophila engrailed gene: expression during embryogenesis. *Genes Dev.* **1**, 29-38.

Karr, T. L., Ali, Z., Drees, B., and Kornberg, T. (1985). The *engrailed* locus of *D. melanogaster* provides an essential zygotic function in precellular embryos. *Cell* **43**, 591-601.

Kimmel, C. B., Kane, D. A., Walker, C., Warga, R. M., and Rothman, M. B. (1989). A mutation that changes cell movement and cell fate in the zebrafish embryo. *Nature* **337**, 358-362.

Kornberg, T. (1981a). Compartments in the abdomen of *Drosophila* and the role of the *engrailed* locus. *Dev. Biol.* **86**, 363-372.

Kornberg, T. (1981b). *Engrailed*: A gene controlling compartment and segment formation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **78**, 1095-1099.

Kornberg, T., Siden, I., O'Farrell, P. and Simon, M. (1985). The *engrailed* locus of *Drosophila*: *in situ* localization of transcripts reveals compartment-specific expression. *Cell* **40**, 45-53.

Lawrence, P. A. and Morata, G. (1976). Compartments in the wing of *Drosophila*: A study of the *engrailed* gene. *Dev. Biol.* **50**, 321-337.

Lawrence, P. A. and Struhl, G. (1982). Further studies of the *engrailed* phenotype in *Drosophila*. *EMBO J.* **1**, 827-833.

Logan, C., Hanks, M. C., Noble-Topham, S., Nalllainathan, D., Provart, N. J. and Joyner, A. L. (1992). Cloning and sequence comparison of the

mouse, human and chicken *engrailed* genes reveal potential functional domains and regulatory regions. *Devel. Gen.* **13**, in press.

Logan, C., Willard, H. F., Rommens, J. M. and Joyner, A. L. (1989). Chromosomal localization of the human homeobox-containing genes, EN1 and EN2. *Genomics* **4**, 206-209.

Morata, G. and Lawrence, P. A. (1975). Control of compartment development by the *engrailed* gene of *Drosophila*. *Nature* **255**, 614-617.

Njølstad, P. R. and Fjose, A. (1988). *In situ* hybridization patterns of zebrafish homeobox genes homologous to *Hox-2.1* and *En-2* of mouse. *Biochem. Biophys. Res. Comm.* **157**, 426-432.

Njølstad, P. R., Molve, A., Apold, J. and Fjose, A. (1990). The zebrafish homeobox gene *hox*-2.2: transcription unit, potential regulatory regions and *in situ* localization of transcripts. *EMBO J.* **9**, 515-524.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.

Patel, N. H., Martin-Blanco, E., Coleman, K. G., Pool, S. J., Ellis, M. C., Kornberg, T., and Goodman, C. S. (1989). Expression of *engrailed* protein in arthropods, annelids, and chordates. *Cell* **58**, 955-968.

Püschel, A., Gruss, P. and Westerfield, M. (1992). Sequence and expression pattern of *pax-6* are highly conserved between zebrafish and mice. *Development* **114**, 643-651.

Sassoon, D. A., Garner, I. and Buckingham, M. (1988). Transcripts of α-cardiac and α-skeletal actins are early marker for myogenesis in the mouse embryo. *Development* **104**, 155-164.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354.

Westerfield, M. (1989). *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (*Brachydanio rerio*).* Eugene, OR, USA: University of Oregon Press.

Westerfield, M., McMurray, J. V., and Eisen, J. S. (1986). Identified motoneurons and their innervation of axial muscles in the zebrafish. *J. Neurosci.* **6**, 2267-2277.

Wilkinson, D. G., Bailes, J. A., Champion, J. E. and McMahon, A. P. (1987). A molecular analysis of mouse development from 8 to 10 days *post coitum* detects changes only in embryonic globin expression. *Development* **99**, 493-500.

Wood, A. (1982). Early pectoral fin development and morphogenesis of the apical ectodermal ridge in the killifish, *Aphyosemion scheeli*. *Anat. Rec.* **204**, 349-356.

(Accepted 27 August 1992)

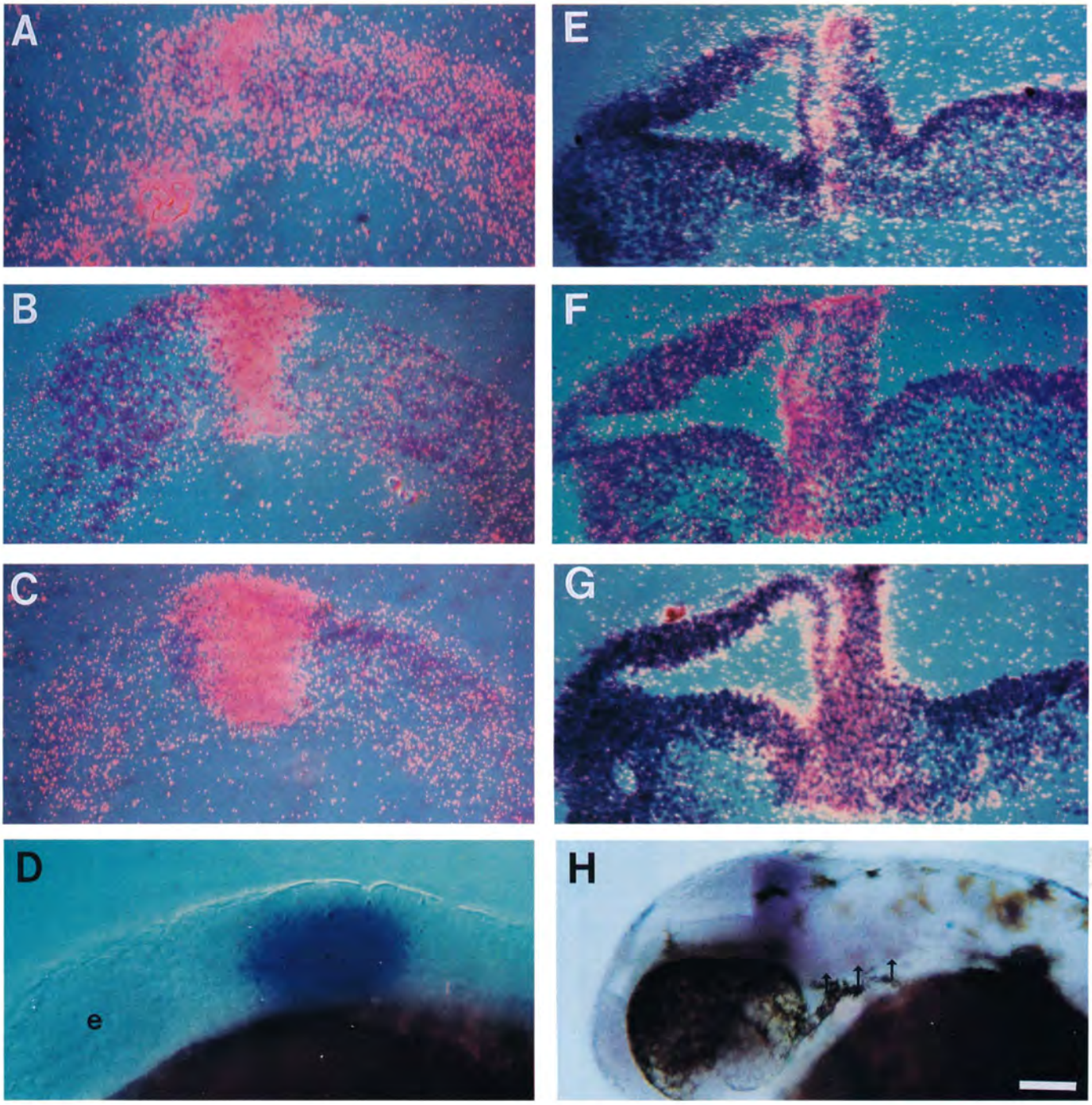

Fig. 6. Cells at the junction between the midbrain and hindbrain express all three *engrailed* genes. Parasagittal sections through the heads of 12 h (A-C) and 32 h (E-G) embryos and 22 h (D) and 32 h (H) whole-mount embryos were hybridized with probes for the *eng1* (A,E), *eng2* (B,F) or *eng3* (C,D,G,H) genes. A cartoon of the engrailed expression domains at the midbrain/hindbrain junction is shown in (I). The domain expressing only *eng3* is shown in blue, the domain expressing both *eng2* and *eng3* is shown in red, and the domain expressing all three is shown in green. Anterior is oriented to the left and dorsal to the top in this and subsequent figures. AC and E-G are double exposures with bright field (blue) and darkfield (red) illumination. The red signal in the lower part of A is due to the yolk, rather than specific hybridization, as shown in control sections without hybridization. Arrows in H indicate hindbrain cells expressing *eng3* (see Fig. 8). Scale bar, 80 μm in A-D, 50 μm in E-G and I, 100 μm in H.

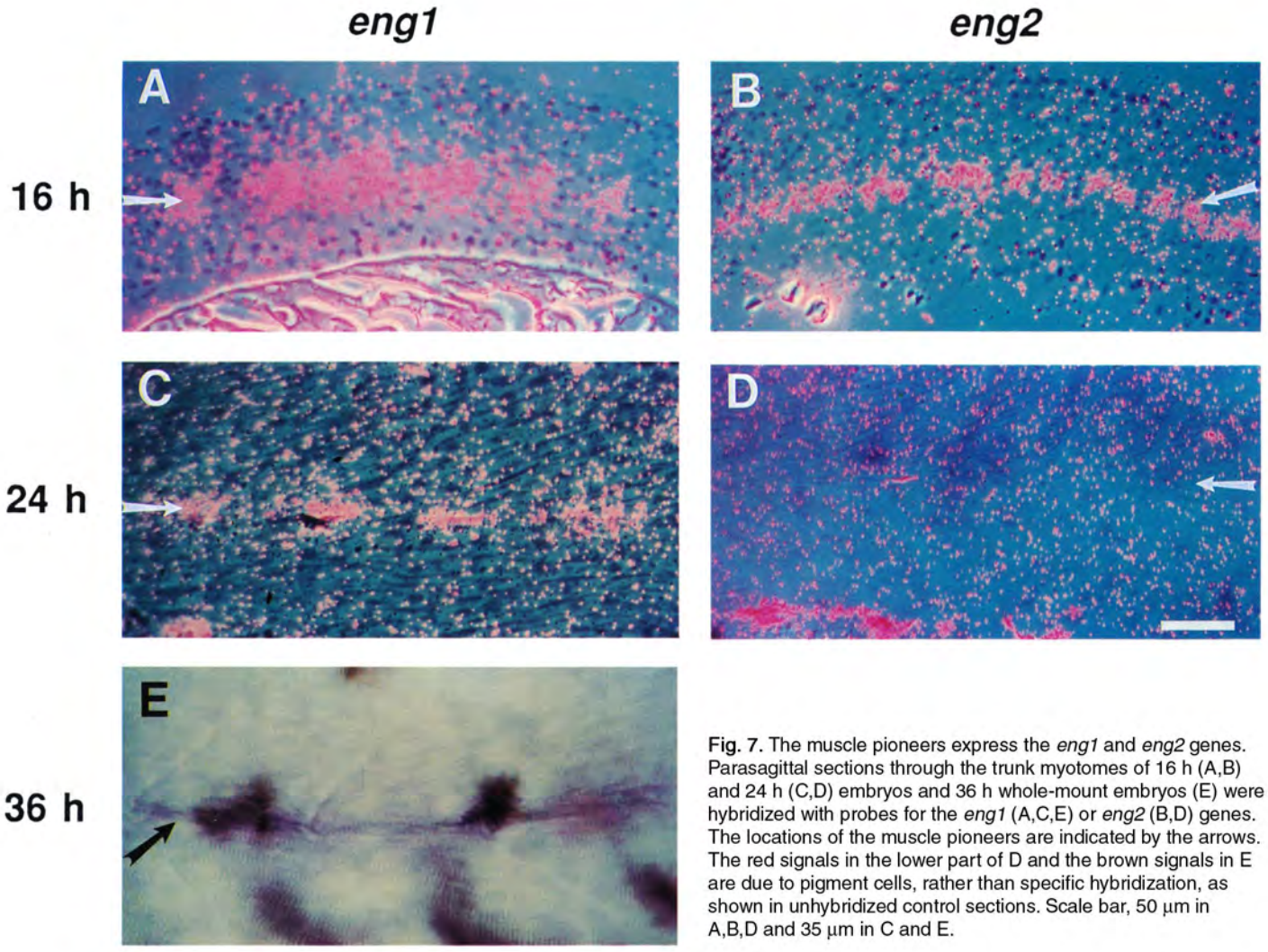

Fig. 7. The muscle pioneers express the *eng1* and *eng2* genes. Parasagittal sections through the trunk myotomes of 16 h (A,B) and 24 h (C,D) embryos and 36 h whole-mount embryos (E) were hybridized with probes for the *eng1* (A,C,E) or *eng2* (B,D) genes. The locations of the muscle pioneers are indicated by the arrows. The red signals in the lower part of D and the brown signals in E are due to pigment cells, rather than specific hybridization, as shown in unhybridized control sections. Scale bar, 50 μm in A,B,D and 35 μm in C and E.

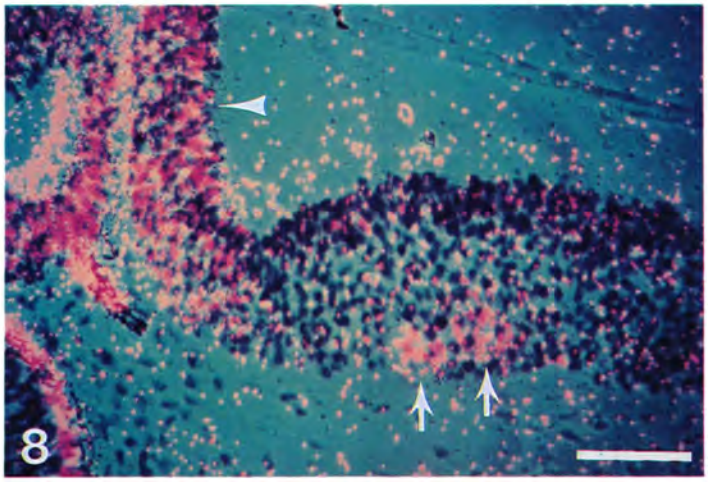

Fig. 8. Cells in the first three hindbrain rhombomeres express the *eng3* gene. A parasagittal section through the hindbrain of a 32 h embryo was hybridized with a probe for the *eng3* gene. Small clusters of cells (arrows) in the ventral parts of the second and third rhomobomeres hybridize. A cluster of cells in the first rhombomere (contained in a different section, not shown, and as illustrated by the whole-mount embryo in Fig. 6 H) also hybridized. The hybridization at the junction between the midbrain and hindbrain is also apparent (arrowhead). The signal in the lower left is caused by the pigment epithelium of the retina and is not due to specific hybridization as shown in unhybridized control sections. Scale bar, 50 μm.

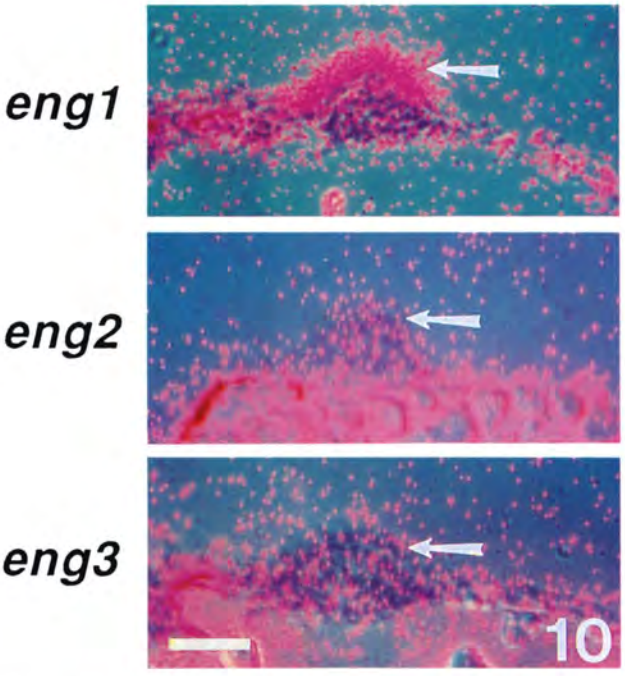

Fig. 10. Ectodermal cells in the pectoral fins express the eng1 gene. Parasagittal sections through the distal portion of the pectoral fin buds of 32 h embryos were hybridized with probes for the *eng1* (top), *eng2* (middle) or *eng3* (bottom) genes. The lateral ectoderm in the ventral part of the bud (arrows) hybridized with only the eng1 probe. The red signals in the lower part of B and C are due to the yolk, rather than specific hybridization, as shown in unhybridized control sections. Scale bar, 50 μm.

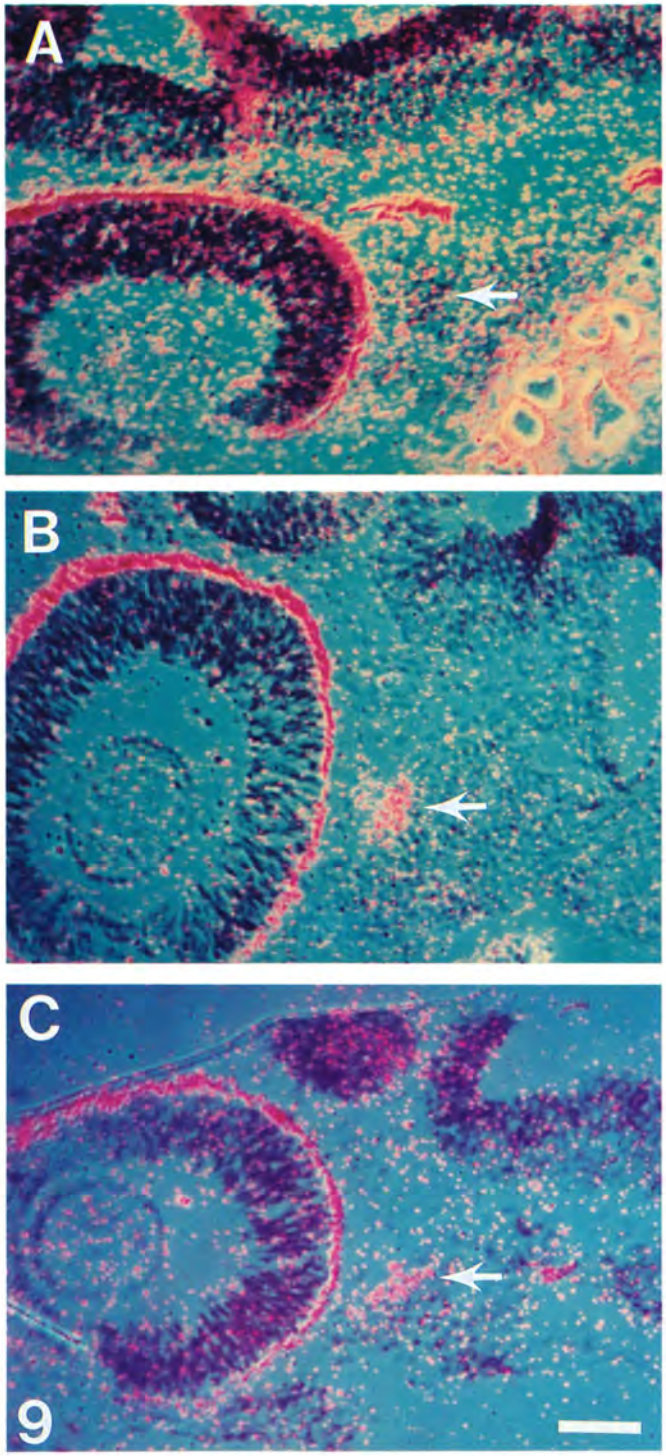

Fig. 9. Jaw muscle precursors express the *eng2* and *eng3* genes. Parasagittal sections of 32 h embryos were hybridized with probes for the *eng1* (A), *eng2* (B), or *eng3* (C) genes. Precursors of the jaw muscles (arrows) are indicated. The signal around the eye in each panel is caused by the pigment epithelium of the retina rather than to specific hybridization as shown in unhybridized control sections. Scale bar, 50 μm.