Int. J. Dev. Biol. 36: 229-237 (1992)

Original Article

The zebrafish homeobox gene *hox[zf-114]:* primary structure, expression pattern and evolutionary aspects

ANDERS MOLVEN[1]*, IVAR HORDVIK[1], PÅL RASMUS NJØLSTAD[1], MARIJKE VAN GHELUE[2] and ANDERS FJOSE[2]

[1]*Laboratory of Biotechnology, University of Bergen, Bergen and*
[2]*Institute for Medical Biology, University of Tromsø, Breivika, Tromsø, Norway*

ABSTRACT It is gradually becoming accepted that vertebrate homeobox genes, like their counterparts in *Drosophila*, are crucial for normal development of the embryo. Most vertebrate homeoboxes reported so far are related to the *Drosophila Antennapedia* (*Antp*) sequence, and here we describe *hox*[*zf-114*], a novel *Antp*-like homeobox gene from the zebrafish. The sequence of the *hox*[*zf-114*] homeodomain indicates that this gene could be a member of a subfamily defined by the mouse *Hox-1.5/-2.7/-4.1* genes. However, the evolutionary origin of *hox*[*zf-114*] is unclear and, based on the putative protein sequence, we conclude that it is not directly homologous to *Hox-1.5, Hox-2.7* or *Hox-4.1*, or to other known mammalian homeobox genes. Nevertheless, as revealed by *in situ* hybridization, *hox*[*zf-114*] exhibits a spatial expression pattern typical for vertebrate *Antp*-like homeobox genes. Transcripts are detected in the posterior hindbrain, where a sharp anterior border of expression is observed, and throughout the spinal cord. The *hox*[*zf-114*] gene is also active in a region that gives rise to the pectoral fins. These findings suggest a role for *hox*[*zf-114*] in anteroposterior patterning of the neural tube and in pectoral fin development.

KEY WORDS: *Brachydanio rerio, embryogenesis, homeobox gene, neural development, zebrafish*

Introduction

Molecular studies on animal embryogenesis have been greatly facilitated by the discovery of the homeobox. This conserved gene element was originally identified in the fruit fly *Drosophila melanogaster* (McGinnis *et al.*, 1984a; Scott and Weiner, 1984), but is probably present in all metazoans. In *Drosophila*, the homeobox is characteristically found in many of the genes regulating pattern formation and differentiation during early development (reviewed in Ingham, 1988; Hayashi and Scott, 1990). Similarly, most of the analyzed vertebrate homeobox genes have temporal and spatial expression patterns that suggest a regulatory function in embryogenesis (reviewed in Holland and Hogan, 1988; Kessel and Gruss, 1990). This function is likely to be mediated through the homeodomain, the precisely defined protein region encoded by the homeobox (Gehring *et al.*, 1990). The homeodomain consists of some 60 amino acids and can bind to DNA in a sequence-specific manner (Desplan *et al.*, 1988; Müller *et al.*, 1988). Thus, it is now generally accepted that homeodomain-containing proteins act as transcription factors (Levine and Hoey, 1988).

Since the homeobox is evolutionarily conserved, similar sequences can be isolated from vertebrates by using probes from, for example, *Drosophila*. Thereby one can identify vertebrate genes potentially having control functions in embryogenesis without first isolating specific developmental mutants. In the mouse, nearly 40 different genes related to the *Drosophila Antp* sequence have been identified. These so-called *Hox* genes are organized into four large gene clusters and, within each *Hox* cluster, individual genes map in the same linear order as their most closely related *Drosophila* homeotic genes (Boncinelli *et al.*, 1988; Graham *et al.*, 1989). As in the fruit fly, the expression domain of each *Hox* gene is correlated with its position in the cluster, i.e., the more 3' a gene is located in the complex, the more anterior its transcripts extend in the embryo (Gaunt *et al.*, 1988; Duboule and Dollé, 1989; Graham *et al.*, 1989). However, the homology between *Hox* genes and *Drosophila* homeotic genes is even more intriguing: when transfected into *Drosophila* embryos, vertebrate *Hox* genes can substitute for at least some of the functions of their counterpart fly genes (Malicki *et al.*, 1990; McGinnis *et al.*, 1990). Moreover, it has recently been shown by mutation analyses that *Hox* genes, like their relatives in *Drosophila*, are essential for normal development of the embryo (Chisaka and Capecchi, 1991; Lufkin *et al.*, 1991).

Abbreviations used in this paper: Antp, *Antennapedia* gene; Bp, base pairs; CNS, central nervous system; kb, kilobase pairs.

***Address for reprints**: Laboratory of Biotechnology, HIB, University of Bergen, N-5020 Bergen, Norway. FAX: 47-5-31.49.52.

0214-6282/92/$03.00
© UBC Press
Printed in Spain

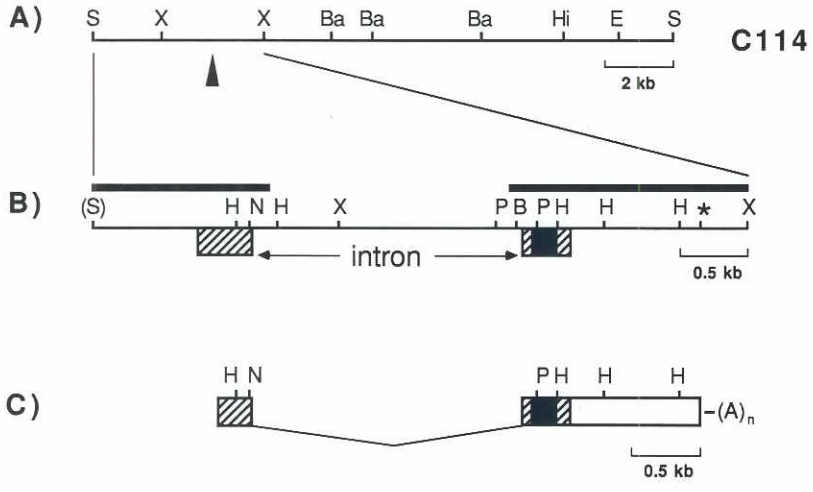


Fig. 1. Organization of the *hox[zf-114]* gene. (A) *Restriction map of the genomic lambda clone C114 containing* hox[zf-114]. *The homeobox-containing XbaI-fragment is indicated by an arrowhead. The SalI sites are derived from the phage.* **(B)** *Enlargement of the genomic region where* hox[zf-114] *is located. The hatched boxes symbolize the open reading frame, of which the homeobox is painted black. The horizontal bars mark the region that has been sequenced and that is listed in Fig. 2. The positions of the poly(A)-signal (indicated by an asterisk) and the intron were deduced by analysis of the cDNA clone.* **(C)** *Structure of the* hox[zf-114] *cDNA clone aligned with* ***B****. Non-coding 3' sequence is indicated by a white box and* $(A)_n$ *denotes the poly(A)-tail. Restriction enzymes:* B, *BglII*; Ba, *BamHI*; E, *EcoRI*; H, *HincII*; Hi, *HindIII*; N, *NcoI*; P, *PvuII*; S, *SalI*; X, *XbaI*.

Despite the striking similarities described above and the rapid progress of molecular biology, we still have only limited information about the developmental function of vertebrate *Hox* genes. Due to the great complexity of the mammalian embryo it will probably be necessary to exploit other systems before a profound understanding of the genetic control of vertebrate development can be reached. Considerable progress has been made with the zebrafish *(Brachydanio rerio)*, which has a number of experimental advantages (Kimmel, 1989). We have initiated a study of zebrafish homeobox genes and have previously reported the cloning and characterization of seven such sequences (Eiken *et al.*, 1987; Fjose *et al.*, 1988a; Njølstad *et al.*, 1988a, 1988b, 1990). For some of the genes, definite mouse homologues could be assigned (Njølstad *et al.*, 1988b, 1990). In these cases, both the gene structure and the expression pattern were highly conserved considering that the fish and mammalian lines of evolution have been separated for more than 300 million years (Benton, 1990).

Here we describe the lambda clone C114, isolated from a zebrafish genomic library during a screening for homeobox genes (Eiken *et al.*, 1987). We present the sequence of the homeobox found in C114 and the molecular structure of the corresponding gene, *hox[zf-114]*. We also show by *in situ* hybridization that this gene seems to be expressed in a restricted anteroposterior domain of the central nervous system (CNS) and in the region where the pectoral fins will form. Finally, we discuss possible evolutionary relations between *hox[zf-114]* and other vertebrate *Hox* genes.

Results

Structural analysis of the lambda clone C114

The screening of a zebrafish genomic library for homeobox sequences (see Materials and Methods) resulted in 30 independent clones. These clones could be classified into 10 groups on the basis of their cross-hybridization patterns with various homeobox probes (not shown). The lambda clone C114 (Fig. 1A) was the single member of one of these groups. The hybridization studies indicated that a 0.56 kb *Sau*3A subclone, which is located within the 3.0 kb *XbaI* fragment shown in Fig. 1A, contained a homeobox and this was confirmed by subcloning, followed by sequence analysis. The corresponding gene has provisionally been named *hox[zf-114]* in accordance with the proposed nomenclature for zebrafish genes (Njølstad *et al.*, 1990).

Next, to determine the organization of the *hox[zf-114]* gene, we screened a cDNA library with a probe from C114 (see Materials and Methods). Only one positive clone, having a size of 1.55 kb, was obtained (Fig. 1C). Sequence analysis of this cDNA indicated that it was not full length at the 5' end (see below), but a canonical poly(A)-signal and a poly(A)-tail were present at the 3' terminus. Hybridization experiments together with a comparison of genomic and cDNA restriction maps revealed the orientation and structure of the *hox[zf-114]* gene depicted in Fig. 1B.

Characterization of the hox[zf-114] *gene and predicted protein*

The nucleotide sequence of *hox[zf-114]*, based on the analysis of both genomic and cDNA subclones, is shown in Fig. 2. At a position 47 bp upstream of the homeobox (nucleotide number 2386, Fig. 2), the cDNA and genomic sequences diverge, the latter continuing into a consensus acceptor site for mRNA splicing. The 5'-most 203 bp of the cDNA clone are colinear with the genomic sequence from position 385 (adjacent to a consensus donor splice site) and upstream. Thus, an intron is present in the *hox[zf-114]* gene, and its size was estimated at 2.0 kb.

In the cDNA clone, the open reading frame containing the homeobox region extends to the ultimate 5' end of the clone (corresponds to position 183 in Fig. 2). The only ATG codons in this frame are located very close to the splice site and would result in homeodomain proteins of unusually small sizes. We therefore conclude that the cDNA clone is not full-length. Consistent with this, analysis of the genomic sequence shows that the open reading frame continues approximately 180 bp upstream of the first nucleotide of the cDNA clone. In the genomic extension, there are two possible initiation codons. The 5'-most we have given the position number 1, the other is located 84 bp downstream (Fig. 2). If we assume that the longest open reading frame is the one translated *in vivo*, the corresponding *hox[zf-114]* protein has a length of 250 amino acids. The homeodomain is contained in the C-terminal part of the protein. Hence, it is encoded by the second exon, a common feature of the *Hox* genes. Common is also the

```
-777  GATCCTTTTACTTGTGTTGCAATTTCCTTAACATTAAAGAAACACATTTTTACAAAGCACGATGATAGAGATAGACAATGTAGCCAATAT
-687  GCTATTATTACTAAAGTATTTTATTCGCAATAAAACATCGCTAAAATGTACATCACTATTTACGTAAGATTAAAACCGGTTAAAATACGC
-597  AATCGAAGAACATTGACAAGCGTAAGGAACTGTTGTTAGATCAGCAGAGTGTGAAATGAATAGTTATTTAACCATAAAGCTAAACAAGTC
-507  GATTATTTTTAAATACCTACCACAAGAGGACTGAAAAGTCAGTCAAAGATTTAAACATTGTTTTATAAAACATTATCACGGAAATTAAGT
-417  CTGTATTCTTCAATCCATTTAAGCAAAACTAAGACGACGCGATCCCAGACATGTATTTTCTCTGTCCAAAATGGGTCTCTCTCTCTCCCT
-327  CCCACCTCACGCTCTCTGAACTCGTGTCAGCACGATGACAGTGTAATGTATGCGCCCTGCTCATGGTTTTTCGTCTGTCCACAGGAAAAT
-237  CACTCGAAATGGCAAAATAGTCGAAGTATGCGGATGAAAACGGAATCACCGTTTGGCTGCTTACATGCAGTTTACAACTCAAACACCGAG
-147  TGCGCAAGGCTATTGGAGGAGGCCACCACGTGACTACATTGTTTACTCTTGTTGGCAAACAATTAAAGACGCGAGTAAGTGTATTTTTTG

   1                                   ***                MetAsnAsnAsnSerPheHisGluTyrAsnPhe
 -57  GCTTTCCTACACCTCGAATGAACAATTATATATTTCGGGTGATTAGGACGATTAGAAATGAACAATAACTCGTTTCACGAATATAATTTT

  12  AsnSerGluTyrLeuAsnLysIleCysCysAsnLysSerTyrValCysTyrMetGlyGlnHisPheSerProCysAlaLeuGlnGluAsn
  34  AACAGTGAATACTTAAACAAAATATGTTGCAATAAATCTTATGTTTGTTATATGGGGCAACATTTCTCCCCCTGTGCTCTGCAGGAGAAC

  42  SerSerThrTrpLysArgHisAlaGluGluSerSerSerGluAsnAspLysGlyAsnValTrpAsnPheGlnAsnLeuSerAsnValGln
 124  AGTTCTACGTGGAAACGACATGCTGAAGAAAGTTCATCAGAAAATGATAAAGGCAACGTGTGGAACTTTCAAAACCTCTCGAATGTCCAA
                                                                 Start cDNA
  72  HisProTyrSerPheSerAspAspGlyAsnHisSerLeuAspProAlaAsnAlaLeuGluArgLysLysAlaCysGluLeuSerThrSer
 214  CATCCCTATTCATTCTCTGATGATGGCAACCACTCTTTGGACCCTGCGAATGCATTGGAAAGAAAAAAGGCATGTGAGCTGTCAACTTCC

 102  CysLeuSerThrMetLysTyrProTrpMetArgGluThrHisAlaProThrHisPheSerSerIleAsnAlaMetGluSerG
 304  TGCTTATCAACCATGAAGTACCCTTGGATGAGAGAGACACATGCACCCACGCACTTCAGCTCCATTAACGCCATGGAATCAGgtgagaga

            ***                   ***
 394  acacgactttgaagtaaattatgcatattgatttcatgactgtattgcttagaaatacccactaaatttaatcagtcagtcagtcagtca
                                                                                          ***
 484  gt-----------------------------<  INTRON (2.0 kb)  >----------------------------ctgggatgagag
                                                                                                ly
2298  ttctgttcactgtctaaaactgtgcagcagtctctgtttgacagatctggtatgcctaaacacttctgtatatctgcatattttctagGT

 130  AspSerLysTyrSerAsnGlyGluAlaValValArgAsnSerSerSerLysArgAlaArgValAlaPheThrSerSerGlnLeuLeuGlu
2388  GATTCTAAATACAGCAATGGTGAAGCAGTTGTACGGAATAGCAGCAGTAAACGGGCTCGTGTGGCGTTCACCTCCTCTCAGCTGCTGGAG

 160  LeuGluLysGluPheHisPheSerAlaTyrLeuCysArgAsnArgArgLeuGluMetAlaGluLeuLeuLysLeuThrAspArgGlnIle
2478  TTAGAGAAGGAGTTTCACTTCAGCGCTTACCTGTGCCGCAACCGGAGACTTGAAATGGCTGAGCTGCTGAAGCTCACCGACCGGCAGATT

 190  LysIleTrpPheGlnAsnArgArgMetLysTyrLysLysAspHisLysGluLysSerThrAlaLysSerSerTyrThrTyrLeuGlyThr
2568  AAAATCTGGTTCCAAAACCGTCGCATGAAGTATAAGAAGGACCACAAGGAAAAGTCAACGGCAAAGTCCTCTTACACCTACCTGGGAACA

 220  GluAsnGlnProLeuIleIleSerArgSerThrThrAspSerProValProLeuLysPheGlnAsnAsnTyrGluThrProSerMetAsn
2658  GAAAACCAGCCATTAATAATTAGCAGAAGCACAACAGATTCTCCGGTGCCATTGAAATTCCAGAATAACTATGAAACACCATCAATGAAC

 250  Trp***
2748  TGGTGAAATTACAGAAAAACTGAAAACCGGTGACTTTTGCTTTTAGAACTTATATAATGATAGTAATCTGCATCAACCATAAAGACTGAA
2838  AATGAATGAATGAATGAATGAATGTGTGAATGAATGAAATTTCTACAAAACATGTACATGATCTGTCACTGTTCAGCTATTTCTCCATTT
2928  GTAACTCTCTCTGACATGATTTATATTTAAAAACATTGATGATTAATTGCCTCTGGATGATGATTTATGTGCAAAGTTAACTTGGAAGAG
3018  TTCAATAGGCAAGCTCGCTCTGAACCAGATTACATGATATCAGCTAGTACATTTCAGAGTTAGTCTCTATTGTGCATAAACTGCTGTAGT
3108  AATATCTTTGCAAAGCACCAAGATTTATTGCAACCCTACTTGTGCAAGTTTATCTCTTGGGCACAGAAAATACAATTTTAAAAGGGCATT
3198  CAAGACTTTGAATGGTCATTAACATCCTCTCAACTTTTTAAGTGCACAACATATTTTTTATTACTCTCAAAATGATATGCTACTTTTCCC
3288  AAGGCATGAGCTAAAATGTACAACTTGCAGCGTTGTTCATTATGATGGTCACAATTGTTCCATGTTAAAGCCTATTGACCTGCTCTACTA
3378  TTAACAATGAATTTGAATGACTTTTTTTCTGCTTTACACTGAAAGTAGGTGCGTTTAAGTCTGAATTTACAGCATATCTACTGACAAAAT
3468  GAATGAAAGAGCAAAGCAGGATTTTCACAATGCTATATCAATGCATTTATAGTGTAATTACTTTACCCATTCATGTACACAAAAGTGTCA
3558  ACATGGCACCTGCAATGTAAAATATGTGCATATTTGTAATTTATACCATTTCAATTATAACCTGTCCATATTTGACAAAAAAACATGCAT
                                                                                  *
3648  TTTGAAAATAATAAGTGGTTGTTTTTTAAACTGATTTATTCCCTCTGGATAGTCAATAAATAATCATTGTAATTCATGTAACTGACAATA
3738  ATGTTGTTTTGTTTAGTTAAAACAGGTAATGGTGAGTCTAAATTTAGCTTAAAGGTGCTTCTGTGAATCTTACTTGATACTTTTTGAATT
3828  AGGCAATACTGTACTTTGCATTCTTAAAACATAACCTTACATATATGCCTTCCTCCGCCTCCGCCTACCAACCTCGGAACATTAATCATT
3918  TAGACAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAGCTTCACAAGTGCTCAACCCGCTGGGCCAAATGACAGGC
4008  TGTGCGCTTCTCTTCCAACACACTTCTCTCCAATACGGGAGCTGAAATACATCTAGA
```

Fig. 2. Genomic sequence (corresponding to Fig. 1B) and a translation of the protein coding region of *hox[zf-114]*. *The two possible initiation codons are underlined and numbering starts at the more upstream position. Asterisks mark stop codons delimiting the open reading frame and in-frame stop codons within the intron. The hexapeptide and homeobox sequences are boxed. A canonical TATA-box upstream of the putative initiation codon and the consensus donor and acceptor splice sites are in bold face letters. For the intron, only the 100 bp closest to the splice sites are listed. In the 3'untranslated region, ATTTA-stretches (see text) are underlined. The poly(A)-signal is shown in bold face and the asterisk indicates the site of poly(A)-addition. The sequence of* hox[zf-114] *appears in the EMBL nucleotide sequence database under the accession number X60095.*

presence of a short conserved sequence in the first exon, close to the splice site. It translates into a hexapeptide of unknown function and occurs in *hox*[*zf-114*], too (boxed in Fig. 2). The N-terminus of the predicted *hox*[*zf-114*] protein, however, differs significantly from typical N-termini of vertebrate *Hox* proteins (Schughart *et al.*, 1988).

Since we failed to obtain a full-length cDNA clone, conclusive data about the 5' structure of the *hox*[*zf-114*] gene is lacking. For example, additional upstream exons like those reported by Simeone *et al.* (1988) and Cho *et al.* (1988) cannot be ruled out. But we note that a canonical TATA-box, which might be used for transcription *in vivo*, occurs at position -444 (Fig. 2). Usage of this TATA-box would result in a processed transcript of about 2.2 kb, which is in good agreement with the major band (estimated at 2.3 kb) detected on a northern blot (see below). However, *in vitro* mapping of transcription start points and isolation of additional cDNA clones are necessary for clarification of how the 5' region of *hox*[*zf-114*] is organized.

In the 3' region of the gene, there is a poly(A)-signal at position 3702, the actual site for poly(A)-addition following 20 bp downstream. This implies that the *hox*[*zf-114*] transcript has an untranslated 3' tail of nearly 1 kb. Long, non-coding 3' regions are often observed in the mRNAs of *Hox* genes and could serve regu-

```
                                 20                  40                  61                    Reference:
      Antp  RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENK   Scott and Weiner (1984)
   Hox-2.5  SRKK-CP--K----------L--M----D--H-V-RL-N-S---V----------M--M-N   Bogarad et al. (1989)
   Hox-2.4  -R------S-----------L--P----K----VS---G-----V---------------N   Hart et al. (1987)
   Hox-2.3  ---------------------Y--------------T-----------------------   Hart et al. (1987)
   Hox-2.2  GR-------------------Y-------------------------------------S-   Hart et al. (1987)
   Hox-2.1  G--A-TA--------------------------------S-----------------D--   Krumlauf et al. (1987)
   Hox-2.6  P--S-TA---Q-V--------Y---------V--------S-----------------DH-   Graham et al. (1988)
 * Hox-2.7  S--A-TA--SA-LV------------C-P--V-M-NL-N-S--------------Y--DQ-   Graham et al. (1988)
   Hox-2.8  SR-L-TA--NT-L----------K--C-P--V---AL-D-----V-V--------H-RQTQ   Rubock et al. (1990)
   Hox-2.9  PGGL-TNF-TR-LT---------K--S-A--V---AT-E-N-T-V----------Q--RER   Rubock et al. (1990)

 * Hox-1.5  S----TA---P-LV------------M-P--V-M-NL-N----------------Y--DQ-   McGinnis et al. (1984b)
 * Hox-4.1  S--A-TA--SA-LV-----------FV-P--VQM-NL-N-S--------------Y--DQ-   Lonai et al. (1987)

hox[zf-114]  S--A-VAF-SS-L---------SA--C-N--L-M-EL-K--D-------------Y--DH-   this report
```

Fig. 3. Comparison of the *hox[zf-114]* homeodomain with the homeodomains of all known genes in the mouse *Hox-2* cluster (top) and paralogs of *Hox-2.7* (bottom). *The homeodomain of the* Drosophila Antp *gene is included as a common reference sequence. Hyphens represent amino acid identity as compared to* Antp. *The* hox[zf-114] *homeodomain is most related to that of* Hox-2.7 *and the* Hox-2.7 *paralogs* Hox-1.5 *and* Hox-4.1 *(marked with asterisks). See also Table 1.*

latory functions. In this connection, it might be relevant that the untranslated part of the *hox[zf-114]* transcript contains several AUUUA motifs (Fig. 2), which have been suggested to influence message stability (Shaw and Kamen, 1986). Such motifs are also found in 3' non-coding regions of some other mRNAs transiently expressed during animal embryogenesis (e.g. Graham *et al.*, 1988; Njølstad *et al.*, 1988b; Nakano *et al.*, 1989).

The hox[zf-114] *homeodomain is related to* Hox-2.7 *paralogs*

The *Hox* complexes are thought to have arisen by duplications of a single ancestral gene cluster (Boncinelli *et al.*, 1988; Graham *et al.*, 1989). This hypothesis is based upon the fact that a given *Hox* gene is more related to one specific gene (denoted *paralog*) in each of the other clusters than it is to any gene in its own cluster. It was therefore of interest to see whether *hox[zf-114]* can be assigned to any subfamily of *Hox* genes. A comparison with all known genes of the murine *Hox-2* complex reveals that the *hox[zf-114]* homeodomain displays the highest similarity to that of *Hox-2.7* (78.7%, Table 1). The similarity is confirmed by relating these two homeodomains to the *Antp* sequence (Fig. 3). In a set of paralogous genes, the corresponding homeodomains tend to have amino acid substitutions (as compared to *Antp*) in identical positions and also to have the same substituting amino acids in those positions. Such a relationship is seen between *hox[zf-114]* and *Hox-2.7*; of the 21 positions in the *hox[zf-114]* homeodomain that are different from the *Antp* sequence, 17 are also different in *Hox-2.7*. Of the latter, 10 are identical substitutions. The only other *Hox* genes that show a similar level of identity to *hox[zf-114]* are the paralogs of *Hox-2.7*, namely *Hox-1.5* and *Hox-4.1* (Table 1 and Fig. 3). We therefore conclude that the *hox[zf-114]* homeobox appears to be a member of the subfamily of homeobox sequences defined by *Hox-1.5*, *-2.7* and *-4.1*. However, *hox[zf-114]* is probably not the zebrafish homologue of any of these genes, and we will elaborate on this issue below.

Expression of hox[zf-114] *in the neural tube*

The zebrafish *hox* genes that we have studied are all expressed during embryogenesis (see, for example, Njølstad *et al.*, 1988a). They are characteristically turned on 11-12 hours after fertilization, around the time when somites start to form and differentiation of the CNS is initiated. We have previously reported the temporal distribution of *hox[zf-114]* transcripts in embryos (Fig. 3 in Fjose *et al.*, 1990), and the pattern is similar to that of the other zebrafish *hox* genes. On the northern blot, one major transcript of 2.3 kb was detected, and there were also weaker bands of 2.8, 1.9 and 1.5 kb (Fjose *et al.*, 1990). Multiple transcripts and complex expression patterns are in fact often observed for vertebrate *Hox* genes (Graham *et al.*, 1988; Condie *et al.*, 1990).

Having established that *hox[zf-114]* has a temporally restricted expression during embryogenesis, we investigated the spatial

TABLE 1

HOMEODOMAIN HOMOLOGIES

	Homology to *hox[zf-114]*	Homology to *Antp*
Antp	65.6%	-
Hox-2.5	59.0%	67.2%
Hox-2.4	55.7%	83.6%
Hox-2.3	63.9%	96.7%
Hox-2.2	62.3%	93.4%
Hox-2.1	68.9%	90.2%
Hox-2.6	67.2%	82.0%
* *Hox-2.7*	78.7%	68.9%
Hox-2.8	65.6%	63.9%
Hox-2.9	57.4%	59.0%
* *Hox-1.5*	75.4%	73.8%
* *Hox-4.1*	73.8%	65.6%
hox[zf-114]	-	65.6%

Amino acid homology (based on 61 residues) between the homeodomain of *hox[zf-114]* and those listed in Fig. 3. A comparison has also been performed to the *Antp* sequence. The *hox[zf-114]* homeodomain exhibits the closest relationship to the homeodomains encoded by *Hox-1.5*, *Hox-2.7* and *Hox-4.1* (marked with asterisks).

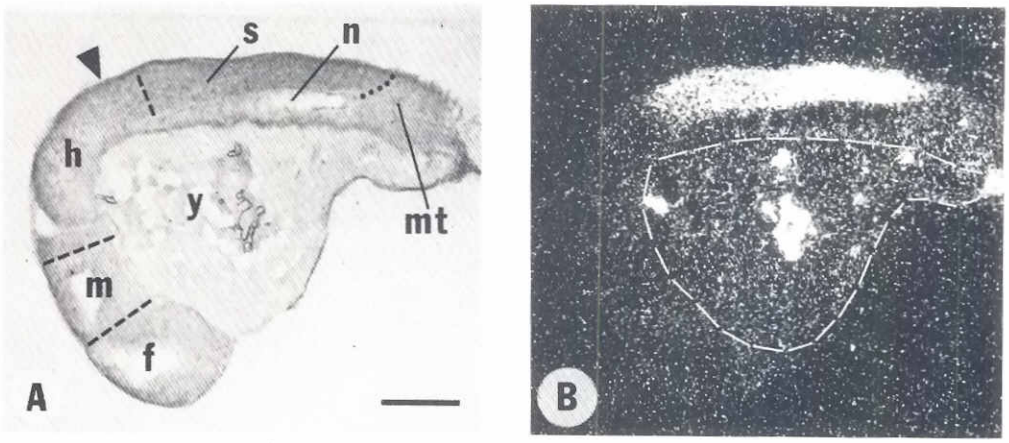


Fig. 4. A parasagittal section of a zebrafish embryo (22 h old) hybridized with a probe from the *hox*[*zf-114*] gene. (A) *The section in bright-field illumination. The arrowhead points to the anterior limit of* hox[zf-114] *expression in the posterior hindbrain. The borders between the main regions of the developing CNS are indicated by dashed lines. Abbreviations: f, forebrain; h, hindbrain; m, midbrain; mt, muscle tissue; n, notochord; s, spinal cord; y, yolk sac.* **(B)** *Dark-field view of A. Signals are present in the posterior part of the hindbrain and in the spinal cord. The apparent posterior limit of expression is simply due to spinal cord leaving the oblique section and muscle tissue entering it (at the stippled line in A). From other sagittal sections (not shown) and from the cross-sections in Fig. 5, it is clear that expression in the spinal cord continues into the tail region. Scale bar (A and B): 0.2 mm.*

distribution of its transcripts. Typical for the *Hox* genes is that each is active in a particular subdomain of the CNS (reviewed in Holland and Hogan, 1988; Kessel and Gruss, 1990), having a discrete anterior boundary of expression. From *in situ* hybridization studies performed on sagittal sections of 22 h zebrafish embryos, it is clear that the same holds for *hox*[*zf-114*] (Fig. 4). Transcripts are detected in the developing neural tube, and there is a sharp anterior border of expression in the hindbrain, posterior to the otic vesicle. This is confirmed by the horizontal section in Fig. 5A,B. More posterior sections (Fig. 5C-H) reveal that spinal expression continues into the tail region of the embryo. Apparently, *hox*[*zf-114*] hybridization signals are uniformly distributed across each transverse section of the spinal cord. In contrast, Graham *et al.* (1991) observed a dorsoventral restriction of expression in the neural tube for several murine *Hox-2* genes. This pattern was developmentally regulated, and further analyses of additional embryonic stages in the zebrafish will be necessary to see whether similar restrictions exist for *hox*[*zf-114*].

Expression of hox[zf-114] *in the pectoral fin buds*

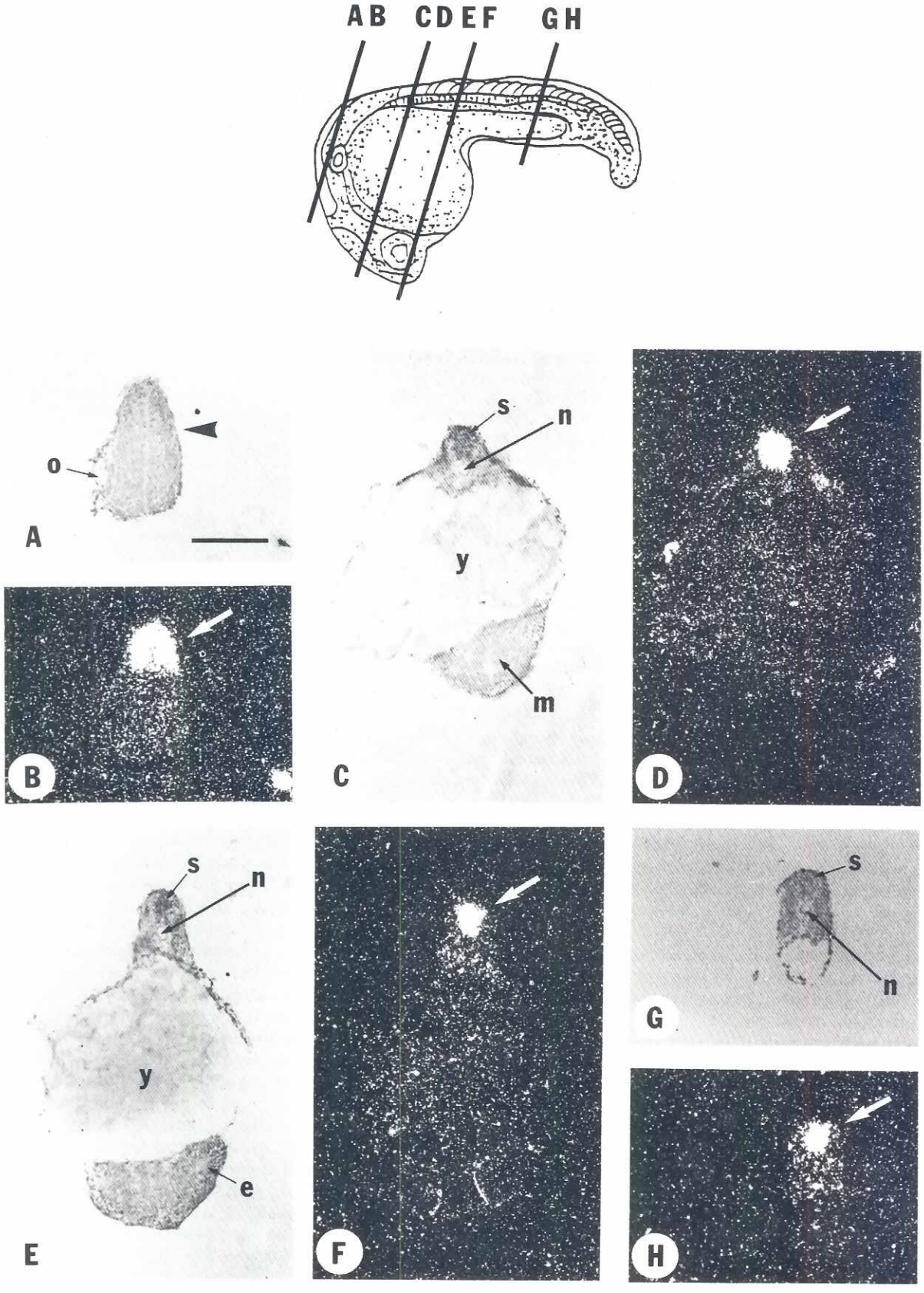
Many *Hox* genes are expressed in the limb buds of mouse and chicken embryos (see Tabin, 1991 for a review). Genes in the *Hox-4* cluster (previously named *Hox-5*) have been studied most extensively in this regard, and each *Hox-4* gene shows a unique pattern of limb bud expression, which correlates to its position within the complex. This finding suggests that vertebrate homeobox genes play a key role during limb development. In a previous report, we have shown that antigens related to the *Xenopus* homeoprotein XIHbox1 (probably the zebrafish *hox-3.3* gene product) are distributed in an anteroposterior gradient in fish pectoral fin buds (Molven *et al.*, 1990), which corresponds to the anterior limb buds of tetrapods. Fig. 6 reveals that in 22 h zebrafish embryos, *hox*[*zf-114*]-derived transcripts have accumulated in the region where the pectoral fin buds will appear 6-7 hours later. The section in Fig. 6 was cut slightly oblique and it should be noted that there is a difference in signal intensity between the two sides of the embryo. This indicates that *hox*[*zf-114*] could be expressed as a gradient in the pectoral fin buds, lending support to the hypothesis that this gene is involved in their specification.

Discussion

Classification of hox[zf-114]

Several lines of evidence suggest that the gene described in this report relates to the growing family of vertebrate *Hox* genes. For example, the genomic organization of *hox*[*zf-114*] points to this conclusion. The open reading frame of the gene spans two exons, with the homeobox located in the second. Close to the splice site, the first exon encodes a conserved hexapeptide. Also, a long untranslated 3' sequence is present in the mRNA. These are all features commonly found in the *Hox* genes. Moreover, the size of the predicted *hox*[*zf-114*] protein is in the range (220-300 amino acids) typical for *Hox* gene products.

Our proposed classification of *hox*[*zf-114*] is further strengthened by its temporal and spatial expression pattern. The gene is turned on shortly after gastrulation is finished, and transcripts are detected in the developing CNS, exhibiting a sharp anterior limit of expression. This is highly reminiscent of the pattern observed for the murine *Hox* genes. However, *Hox* transcripts are usually detected also in mesodermally-derived tissues of the mouse embryo, in domains posteriorly displaced from the anterior expression borders in the CNS (e.g., Gaunt *et al.*, 1988). We have obtained preliminary evidence that *hox*[*zf-114*] is expressed in somitic mesoderm (not shown), but further analyses are necessary for verification and for comparison with the CNS expression domain. Another observation linking *hox*[*zf-114*] to the murine *Hox* genes is that *hox*[*zf-114*] transcripts accumulate in the pectoral fin fields; many genes of the *Hox-3* and *Hox-4* complexes are active in the limb buds of mice (Oliver *et al.*, 1988; Dollé *et al.*, 1989). Since *hox*[*zf-114*] shares so many general properties with the vertebrate *Hox*


AB
CD
EF
GH
o
A
B
s
n
y
m
C
D
s
n
y
e
E
F
s
n
G
H

genes, it is tempting to speculate that *hox[zf-114]* is a member of a gene complex. By performing a chromosomal walk in both directions from the C114 clone it should be possible to verify or disprove this assumption.

The anterior expression border of hox[zf-114]

In mouse embryos, it has now been firmly established that the *Hox* genes are expressed in the embryonic CNS according to their positions in the gene clusters (see Introduction). It still remains to be proven, though, that this is the case also for the zebrafish. In light of the considerable conservation of some zebrafish *hox* genes as compared to their mammalian counterparts (Njølstad *et al.*, 1988b, 1990), it would not be surprising if the same correlation between anterior borders of expression and chromosomal localization were to apply to fish. An indication that it might be so is obtained by comparing the expression domain of *hox[zf-114]* with that previously reported for the putative *hox-3.3* gene product (Molven *et al.*, 1990). Based on analysis of the homeobox sequences, *hox[zf-114]* is likely to be positioned more 3' in a cluster than *hox-3.3*. Consistent with this, the anterior limit of *hox[zf-114]* expression reported here (posterior hindbrain) lies more anterior than the border of *hox-3.3* expression at the 22 h stage (spinal segment number 2).

Wilkinson *et al.* (1989) have performed a detailed investigation of the expression limits of murine *Hox-2* genes in the developing hindbrain. They found that *Hox-2.1*, *-2.6*, *-2.7* and *-2.8* transcripts are distributed in a segment-specific pattern with successive genes having borders at two-segment intervals. Since *hox[zf-114]* transcripts extend into the posterior hindbrain, it is pertinent to ask whether the anterior restriction coincides with a segment border. The resolution of our *in situ* hybridization experiment did not allow us to answer this question unambiguously, but if such a relationship exists for *hox[zf-114]*, the anterior border is at the 7th or 8th hindbrain segment (Ca1 or Ca2, see Hanneman *et al.*, 1988).

Relationships between hox[zf-114] *and other* Hox *genes*

Among the *Hox* sequences reported so far, the homeodomain of *hox[zf-114]* is most similar to that of *Hox-2.7* (Table 1 and Fig. 3). Nevertheless, when we compared the whole coding region of *Hox-2.7* (R. Krumlauf, personal communication) with the open reading frame of *hox[zf-114]* no obvious sequence similarity outside the homeodomains was observed. For example, the hexapeptide of *hox[zf-114]* (Lys-Tyr-Pro-Trp-Met-Arg) shares only 3 residues with the *Hox-2.7* hexapeptide. The sizes of the gene products are also different: the protein encoded by *Hox-2.7* is considerably larger than the predicted *hox[zf-114]* protein. We conclude that *hox[zf-114]* appears not to be the fish equivalent of *Hox-2.7*.

Could *hox[zf-114]* instead correspond to any of the *Hox-2.7* paralogs (i.e. *Hox-1.5*, *Hox-4.1* or a predicted, but yet unidentified, gene in the *Hox-3* cluster)? At least two observations make this idea less attractive. Firstly, the homeodomains of *Hox-1.5*, *Hox-2.7* and *Hox-4.1* have 89-95% mutual similarity, which is substantially higher than the 74-79% identity they exhibit to *hox[zf-114]*. And secondly, the hexapeptides of the three mouse paralogs are identical (P. Lonai, personal communication), but quite different from the *hox[zf-114]* hexapeptide (see above).

We have also compared the predicted *hox[zf-114]* protein with the sequences of the other *Hox* gene products reported thus far. Except for the homeodomain and the hexapeptide, no regions of obvious similarity were revealed (not shown). It is therefore difficult to draw conclusions about the evolutionary origin of *hox[zf-114]*, especially since we have no information about a possible cluster containing this gene.

However, a clue might come from the pattern of *hox[zf-114]* expression in the neural tube. Paralogous genes in the 3' part of the *Hox* complexes tend to have nearly identical anterior borders of expression in the embryonic CNS (Gaunt *et al.*, 1989; Hunt *et al.*, 1991). *Hox-2.7* transcripts are detected in all hindbrain segments from the 5th and caudally (Wilkinson *et al.*, 1989), while *hox[zf-114]* is active from the 7th or 8th segment and caudally. Assuming that there is a one-to-one relationship between the hindbrain segments of mammalian and fish embryos and that the segmental expression pattern of *Hox* genes in the hindbrain is highly conserved, this

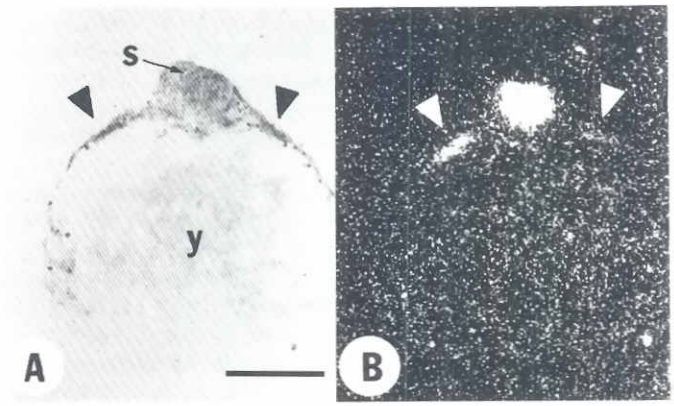


Fig. 6. The *hox[zf-114]* gene is expressed in the developing pectoral fins. (A) *A transverse section of a 22 h zebrafish embryo at the level of the pectoral fin fields (arrowheads) hybridized with the* hox[zf-114] *probe (bright-field illumination).* **(B)** *Dark-field view of A. The fin field signals are indicated by white arrowheads, but due to oblique sectioning the signal over the right side of the embryo is barely visible. Abbreviations: s, spinal cord; y, yolk sac. Scale bar (A and B): 0.2 mm.*

Fig. 5. Sections of a zebrafish embryo (22 h old) hybridized with a probe from the *hox[zf-114]* gene. *The positions of the sections are illustrated in the schematic drawing at the top of the figure. Note that hybridization signals (indicated by white arrows) are present both in the posterior hindbrain and in the spinal cord, and that there are no signals in the forebrain/midbrain regions.* **(A, B)** *A section through the hindbrain at the level of the otic vesicle. An arrowhead points to the anterior border of expression (posterior is to the top). This section is almost horizontally oriented with respect to the anteroposterior axis.* **(C, D)** *A transverse section through the cervical part of the spinal cord (upper) and the midbrain region (lower).* **(E, F)** *A transverse section through the thoracic part of the spinal cord (upper) and the eyes/forebrain (lower).* **(G, H)** *A transverse section through the tail region. Panels* ***A, C, E, G****: bright-field illumination. Panels* ***B, D, F, H****: dark-field illumination. Abbreviations: e, eye; m, midbrain; n, notochord; o, otic vesicle; s, spinal cord; y, yolk sac. Scale bar (A-H): 0.2 mm.*

indicates that *hox[zf-114]* relates to a slightly more 5' or «posterior» subfamily than *Hox-2.7* does. Support for this contention comes from analysis of the *hox[zf-114]* hexapeptide: it is most similar (5 of 6 residues) to that encoded by *Hox-1.3* and some other genes located in the middle of the *Hox-1* and *Hox-2* complexes (Schughart *et al.*, 1988).

Thus, whereas the homeodomain of *hox[zf-114]* hints at a connection to the *Hox-2.7* subfamily, the hexapeptide and anterior expression border suggest that *hox[zf-114]* is a member of one of the more 5' subfamilies. Interestingly, for the murine *Hox-1.3/-2.1/-3.4* group a paralog remains to be identified in the *Hox-4* cluster. If it exists, *hox[zf-114]* may be the zebrafish counterpart of that gene. A high sequence divergence compared to the other members of the subfamily could then explain why the murine *Hox-4* member has been difficult to trace by hybridization analysis.

Alternatively, it is not the hexapeptide and expression border, but the homeodomain that reflects the true ancestry of *hox[zf-114]*. We note that there is a much smaller number of nucleotide differences between the homeoboxes of *Hox-1.5*, *-2.7* and *-4.1*, than there is within genes of the other *Hox* subfamilies (Schughart *et al.*, 1989). This observation can be explained by gene conversions (see Lewin, 1990), and it is conceivable that such events took place after the evolutionary lines leading to fish and mammals had separated. Hence, *hox[zf-114]* might represent an ancient gene which in higher vertebrates was replaced because of gene conversion. The original function(s) of the *hox[zf-114]* ancestor could even have been taken over by more than one gene, as additional genes were recruited during the evolution of an increasingly complex body pattern. One should also keep in mind the possibility that *hox[zf-114]* itself evolved *after* the divergence of fishes and tetrapods (for example by a duplication event within a cluster), thus being a fish-specific gene. Careful mapping and analysis of all zebrafish *hox* complexes will clarify both this issue and other questions concerning the evolution of the vertebrate *Antp*-like genes.

Materials and Methods

Embryos

Zebrafish were maintained and bred essentially as described in Stuart *et al.* (1988). Developmental age is given as 'h', hours after fertilization at 28.5°C, which was the temperature of incubation.

Isolation of hox[zf-114]*-containing clones*

The construction of a zebrafish genomic library in the EMBL3 vector is described in Eiken *et al.*, 1987. The library was screened under low-stringency hybridization and washing conditions (McGinnis *et al.*, 1984a) with a mixture of homeobox probes from *Drosophila* (*Antp*: McGinnis *et al.*, 1984a; *Sex combs reduced*: Kuroiwa *et al.*, 1985; *engrailed*: Fjose *et al.*, 1985) and Atlantic salmon (S12-B: Fjose *et al.*, 1988b). The screening resulted in 30 independent clones, one of them was the lambda clone C114 which contained the *hox[zf-114]* gene reported here.

After confirming by sequence analysis that a homeobox was present in C114, a zebrafish cDNA library (Njølstad *et al.*, 1988b) was screened under high-stringency hybridization and washing conditions (Sambrook *et al.*, 1989) with a 0.56 kb *Sau*3A homeobox-containing fragment from C114. A single, positive clone was isolated.

The restriction fragments used as probes in the above experiments were all purified on agarose gels and labeled with [α-^{32}P]dCTP using nick-translation.

Subcloning and sequencing

Restriction fragments from C114 and from the cDNA clone were subcloned into various pGEM-vectors (Promega Biotec). Double-stranded DNA was sequenced by the chain termination method using the Sequenase system (USB Corporation) and oligonucleotide primers complementary to the plasmid SP6 and T7 promoters. Where necessary, specific sequencing primers were employed or sets of overlapping clones were generated by digestion with exonuclease III ('Erase-A-Base System', Promega Biotec).

Sectioning and in situ *hybridization*

The embryos were fixed and sectioned as described in Molven *et al.* (1991). The *hox[zf-114]* cDNA fragment (Fig. 1C) was labeled with [α-^{35}S]dATP, and *in situ* hybridization and autoradiography were carried out essentially as described in Njølstad *et al.* (1990). The exposure time was 2 weeks.

Acknowledgments

We thank Michele McDowell for making cryostat sections of zebrafish embryos and Robb Krumlauf and Peter Lonai for communicating data prior to publication. We are also indebted to the Department of Medical Genetics, Haukeland Hospital for providing us with necessary facilities, to Rein Aasland for assistance with the computer analyses and to Chuck Kimmel and Monte Westerfield for valuable comments on the manuscript. A.M., P.R.N. and A.F. were supported by grants and fellowships from the Norwegian research councils NAVF and NFFR. Financial support was also obtained from the Nansen Foundation.

References

BENTON, M.J. (1990). Phylogeny of the major tetrapod groups: morphological data and divergence dates. *J. Mol. Evol. 30:* 409-424.

BOGARAD, L.D., UTSET, M.F., AWGULEWITSCH, A., MIKI, T., HART, C.P. and RUDDLE, F.H. (1989). The developmental expression pattern of a new murine homeobox gene: *Hox-2.5. Dev. Biol. 133:* 537-549.

BONCINELLI, E., SOMMA, R., ACAMPORA, D., PANNESE, M., D'ESPOSITO, M., FAIELLA, A. and SIMEONE, A. (1988). Organization of human homeobox genes. *Hum. Reprod. 3:* 880-886.

CHISAKA, O. and CAPECCHI, M.R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *Hox-1.5. Nature 350:* 473-479.

CHO, K.W.Y., GOETZ, J., WRIGHT, C.V.E., FRITZ, A., HARDWICKE, J. and DE ROBERTIS, E.M. (1988). Differential utilization of the same reading frame in a *Xenopus* homeobox gene encodes two related proteins sharing the same DNA-binding specificity. *EMBO J. 7:* 2139-2149.

CONDIE, B.G., HEMMATI-BRIVANLOU, A. and HARLAND, R.M. (1990). Most of the homeobox-containing Xhox 36 transcripts in early *Xenopus* embryos cannot encode a homeodomain protein. *Mol. Cell. Biol. 10:* 3376-3385.

DESPLAN, C., THEIS, J. and O'FARRELL, P.H. (1988). The sequence specificity of homeodomain-DNA interaction. *Cell 54:* 1081-1090.

DOLLÉ, P., IZPISUA-BELMONTE, J.-C., FALKENSTEIN, H., RENUCCI, A. and DUBOULE, D. (1989). Coordinate expression of the murine *Hox-5* complex homoeobox-containing genes during limb pattern formation. *Nature 342:* 767-772.

DUBOULE, D. and DOLLÉ, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. *EMBO J. 8:* 1497-1505.

EIKEN, H.G., NJØLSTAD, P.R., MOLVEN, A. and FJOSE, A. (1987). A zebrafish homeobox-containing gene with embryonic transcription. *Biochem. Biophys. Res. Commun. 149:* 1165-1171.

FJOSE, A., EIKEN, H.G., NJØLSTAD, P.R., MOLVEN, A. and HORDVIK, I. (1988a). A zebrafish engrailed-like homeobox sequence expressed during embryogenesis. *FEBS Lett. 231:* 355-360.

FJOSE, A., EIKEN, H.G., NJØLSTAD, P.R., MOLVEN, A. and HORDVIK, I. (1990). Homeobox sequences of Atlantic salmon (*Salmo salar*) and zebrafish (*Brachydanio rerio*). *Aquaculture 85:* 51-60.

FJOSE, A., McGINNIS, W.J. and GEHRING, W.J. (1985). Isolation of a homoeobox-containing gene from the *engrailed* locus of *Drosophila* and the spatial distribution of its transcripts. *Nature 313:* 284-289.

FJOSE, A., MOLVEN, A. and EIKEN, H. G. (1988b). Molecular cloning and characterization of homeobox-containing genes from Atlantic salmon. *Gene 62:* 141-152.

GAUNT, S.J., KRUMLAUF, R. and DUBOULE, D. (1989). Mouse homeo-genes within a subfamily, *Hox-1.4*, *-2.6* and *-5.1*, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. *Development 107:* 131-141.

GAUNT, S.J., SHARPE, P.T. and DUBOULE, D. (1988). Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. *Development 104 (Suppl.):* 169-179.

GEHRING, W.J., MÜLLER, M., AFFOLTER, M., PERCIVAL-SMITH, A., BILLETER, M., QIAN, Y.Q., OTTING, G. and WÜTHRICH, K. (1990). The structure of the homeodomain and its functional implications. *Trends Genet. 6:* 323-329.

GRAHAM, A., MADEN, M. and KRUMLAUF, R. (1991). The murine *Hox-2* genes display spatially and temporally dynamic dorsoventral patterns of expression during central nervous system development. *Development 112:* 255-264.

GRAHAM, A., PAPALOPULU, N. and KRUMLAUF, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell 57:* 367-378.

GRAHAM, A., PAPALOPULU, N., LORIMER, J., McVEY, J.H., TUDDENHAM, E.G.D. and KRUMLAUF, R. (1988). Characterization of a murine homeobox gene, *Hox-2.6*, related to the *Drosophila Deformed* gene. *Genes Dev. 2:* 1424-1438.

HANNEMAN, E., TREVARROW, B., METCALFE, W.K., KIMMEL, C.B. and WESTERFIELD, M. (1988). Segmental pattern of development in the hindbrain and spinal cord of the zebrafish embryo. *Development 103:* 49-58.

HART, C.P., FAINSOD, A. and RUDDLE, F.H. (1987). Sequence analysis of the murine *Hox-2.2, -2.3* and *-2.4* homeoboxes; evolutionary and structural comparisons. *Genomics 1:* 182-185.

HAYASHI, S. and SCOTT, M. P. (1990). What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell 63:* 883-894.

HOLLAND, P.W.H. and HOGAN, B.L.M. (1988). Expression of homeo box genes during mouse development: a review. *Genes Dev. 2:* 773-782.

HUNT, P., GULISANO, M., COOK, M., SHAM, M., FAIELLA, A., WILKINSON, D., BONCINELLI, E. and KRUMLAUF, R. (1991). A distinct *Hox* code for the branchial region of the vertebrate head. *Nature 353:* 861-864.

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

KESSEL, M. and GRUSS, P. (1990). Murine developmental control genes. *Science 249:* 374-379.

KIMMEL, C.B. (1989). Genetics and early development of zebrafish. *Trends Genet. 5:* 283-288.

KRUMLAUF, R., HOLLAND, P.W.H., McVEY, J.H. and HOGAN, B.L.M. (1987). Developmental and spatial expression of the mouse homeobox gene *Hox-2.1. Development 99:* 603-617.

KUROIWA, A., KLOTER, U., BAUMGARTNER, P. and GEHRING, W.J. (1985). Cloning of the homeotic *Sex combs reduced* gene in *Drosophila* and *in situ* localization of its transcripts. *EMBO J. 4:* 3757-3764.

LEVINE, M. and HOEY, T. (1988). Homeobox proteins as sequence-specific transcription factors. *Cell 55:* 537-540.

LEWIN, B. (1990). *Genes IV.* Oxford University Press, Oxford, pp. 515-516.

LONAI, P., ARMAN, E., CZOSNEK, H., RUDDLE, F.H. and BLATT, C. (1987). New murine homeoboxes: structure, chromosomal assignment, and differential expression in adult erythropoiesis. *DNA 6:* 409-418.

LUFKIN, T., DIERICH, A., LEMEUR, M., MARK, M. and CHAMBON, P. (1991). Disruption of the *Hox-1.6* homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell 66:* 1105-1119.

McGINNIS, N., KUZIORA, M.A. and McGINNIS, W. (1990). Human *Hox-4.2* and *Drosophila Deformed* encode similar regulatory specificities in *Drosophila* embryos and larvae. *Cell 63:* 969-976.

McGINNIS, W., HART, C.P., GEHRING, W.J. and RUDDLE, F.H. (1984b). Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes in *Drosophila. Cell 38:* 675-680.

McGINNIS, W., LEVINE, M.S., HAFEN, E., KUROIWA, A. and GEHRING, W.J. (1984a). A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature 308:* 428-433.

MALICKI, J., SCHUGHART, K. and McGINNIS, W. (1990). Mouse *Hox-2.2* specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell 63:* 961-967.

MOLVEN, A., NJØLSTAD, P R. and FJOSE, A. (1991). Genomic structure and restricted neural expression of the zebrafish *wnt-1* (*int-1*) gene. *EMBO J. 10:* 799-807.

MOLVEN, A., WRIGHT, C.V.E., BREMILLER, R., DE ROBERTIS, E.M. and KIMMEL, C.B. (1990). Expression of a homeobox gene product in normal and mutant zebrafish embryos: evolution of the tetrapod body plan. *Development 109:* 279-288.

MÜLLER, M., AFFOLTER, M., LEUPIN, W., OTTING, G., WÜTHRICH, K. and GEHRING, W.J. (1988). Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain. *EMBO J. 7:* 4299-4304.

NAKANO, Y., GUERRERO, I., HIDALGO, A., TAYLOR, A., WHITTLE, J.R.S. and INGHAM, P.W. (1989). A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene *patched. Nature 341:* 508-513.

NJØLSTAD, P.R., MOLVEN, 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.

NJØLSTAD, P.R., MOLVEN, A., EIKEN, H.G. and FJOSE, A. (1988a). Structure and neural expression of a zebrafish homeobox sequence. *Gene 73:* 33-46.

NJØLSTAD, P.R., MOLVEN, A., HORDVIK, I., APOLD, J. and FJOSE, A. (1988b). Primary structure, developmentally regulated expression and potential duplication of the zebrafish homeobox gene ZF-21. *Nucleic Acids Res. 16:* 9097-9111.

OLIVER, G., WRIGHT, C.V.E., HARDWICKE, J. and DE ROBERTIS, E.M. (1988). A gradient of homeodomain protein in developing forelimbs of *Xenopus* and mouse embryos. *Cell 55:* 1017-1024.

RUBOCK, M.J., LARIN, Z., COOK, M., PAPALOPULU, N., KRUMLAUF, R. and LEHRACH, H. (1990). A yeast artificial chromosome containing the mouse homeobox cluster *Hox-2. Proc. Natl. Acad. Sci. USA 87:* 4751-4755.

SAMBROOK, J., FRITSCH, E.F. and MANIATIS, T. (1989). *Molecular Cloning. A Laboratory Manual* (Second ed.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

SCHUGHART, K., KAPPEN, C. and RUDDLE, F.H. (1989). Duplication of large genomic regions during the evolution of vertebrate homeobox genes. *Proc. Natl. Acad. Sci. USA 86:* 7067-7071.

SCHUGHART, K., UTSET, M.F., AWGULEWITSCH, A. and RUDDLE, F.H. (1988). Structure and expression of *Hox-2.2*, a murine homeobox-containing gene. *Proc. Natl. Acad. Sci. USA 85:* 5582-5586.

SCOTT, M.P. and WEINER, A.J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of *Drosophila. Proc. Natl. Acad. Sci. USA 81:* 4115-4119.

SHAW, G. and KAMEN, R. (1986). A conserved AU-sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell 46:* 659-667.

SIMEONE, A., PANNESE, M., ACAMPORA, D., D'ESPOSITO, M. and BONCINELLI, E. (1988). At least three human homeoboxes on chromosome 12 belong to the same transcription unit. *Nucleic Acids Res. 16:* 5379-5390.

STUART, G.W., McMURRAY, J.V. and WESTERFIELD, M. (1988). Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. *Development 103:* 403-412.

TABIN, C.J. (1991). Retinoids, homeoboxes, and growth factors: toward molecular models for limb development. *Cell 66:* 199-217.

WILKINSON, D.G., BHATT, S., COOK, M., BONCINELLI, E. and KRUMLAUF, R. (1989). Segmental expression of *Hox-2* homeobox-containing genes in the developing mouse hindbrain. *Nature 341:* 405-409.

Accepted for publication: February 1992