

## Sequence and embryonic expression of the murine *Hox-3.5* gene

ANTÓNIO M. C. GEADA<sup>1</sup>, STEPHEN J. GAUNT<sup>2</sup>, MAY AZZAWI<sup>1</sup>, SEBASTIAN M. SHIMELD<sup>1</sup>, JONATHAN PEARCE<sup>2</sup> and PAUL T. SHARPE<sup>1,\*</sup>

<sup>1</sup>Molecular Embryology Laboratory, Department of Cell and Structural Biology, Stopford Building, University of Manchester, Manchester, M13 9PT, UK

<sup>2</sup>Department of Molecular Embryology, IAPGR, Babraham, Cambridge, CB2 4AT, UK

\*Author for correspondence

### Summary

The murine *Hox-3.5* gene has been mapped and linked genomically to a position 18 kb 3' of its most 5' locus neighbour, *Hox-3.4*, on chromosome 15. The sequence of the *Hox-3.5* cDNA, together with the position of the gene within the locus, show it to be a paralogue of *Hox-2.6*, *Hox-1.4* and *Hox-4.2*. The patterns of embryonic expression for the *Hox-3.5* gene are examined in terms of three rules, proposed to relate a Hox gene's expression pattern to its position within the locus. The anterior boundaries of *Hox-3.5* expression in the hindbrain and prevertebral column lie anterior to those of

*Hox-3.4* and all other, more 5'-located *Hox-3* genes. Within the hindbrain, the *Hox-3.5* boundary is seen to lie posterior to that of its paralogue, *Hox-2.6*, by a distance equal to about the length of one rhombomere. Patterns of *Hox-3.5* expression within the oesophagus and spinal cord, but not the testis, are similar to those of other *Hox-3* genes, *Hox-3.3* and *Hox-3.4*.

Key words: Hox, *Hox-3.5*, homeobox, rhombomeres, development.

### Introduction

Based on comparisons of homeobox gene sequences, genomic organisation and spatial domains of embryonic expression, both mammalian and *Drosophila* class I homeobox genes are believed to be descendants of a single homeobox gene cluster that was present in a common, ancient ancestor (Duboule and Dollé, 1989; Graham et al., 1989). Mouse class I, or Hox, genes (38 described so far) are organised as four clusters of between nine and eleven genes on four different chromosomal locations. *Hox-1* on chromosome 6, *Hox-2* on chromosome 11, *Hox-3* on chromosome 15 and *Hox-4* on chromosome 2 (Hart et al., 1985; Bucan et al., 1986; Breier et al., 1988; Featherstone et al., 1988). The spacing of genes in each cluster and their homeobox sequence shows that genes occupying the same relative position in each locus also have similar sequence characteristics, showing them to belong to the same sub-family or paralogous group.

All Hox genes are transcribed in the same direction, and the more 3' a gene the more anterior the embryonic expression (Gaunt et al., 1988; Duboule and Dollé, 1989; Graham et al., 1988). Paralogous genes, especially those 3' in location, often display similar anterior boundaries of expression in the spinal cord and prevertebral column, and individual genes within each cluster often show similar tissue specificities in expression (reviewed by Gaunt, 1991). The complex anterior-posterior and dorsal-ventral array of Hox gene expression patterns produced in the early embryo

can provide the basis of specifying positional co-ordinates to cells, possibly by a Hox code (Kessel and Gruss, 1991). Evidence for the developmental importance of Hox genes and confirmation of the significance of the anterior boundaries of expression has come from targeted gene deletion studies with the *Hox-1.5* (Chisaka and Capecchi, 1991) and *Hox-1.6* (Lufkin et al., 1991; Chisaka et al., 1992) genes.

The expression boundaries of the more anteriorly expressed genes have attracted much attention, since the boundaries correspond to the segmental boundaries of rhombomeres in the hindbrain. Paralogues *Hox-2.6*, *Hox-1.4* and *Hox-4.2* all have anterior limits of expression at the rhombomere 6/7 boundary (Hunt et al., 1991). *Hox-2.7*, *Hox-1.5* and *Hox-4.1* all have boundaries at rhombomere 4/5 (Hunt et al., 1991). *Hox-2.8* and *Hox-1.11* have boundaries at rhombomere 2/3. *Hox-1.6* and *Hox-2.9* have boundaries at rhombomere 3/4, although *Hox-2.9* is expressed exclusively in rhombomere 4 (Murphy and Hill, 1991). The significance of these rhombomere expression patterns is not yet clear, but the effects of the deletion of *Hox-1.5* (Chisaka and Capecchi, 1991) and *Hox-1.6* (Lufkin et al., 1991; Chisaka et al., 1992) suggest that the Hox genes may specify different sets of neural crest cells emanating at different rhombomeric levels (Hunt and Krumlauf, 1991).

Here we describe the identification of the *Hox-3.5* gene as the next homeobox gene 3' of *Hox-3.4*. Location and homeobox sequence show *Hox-3.5* to be a member of the *Hox-2.6*, *Hox-1.4* and *Hox-4.2* subfamily. *Hox-3.5* appears to mark the 3' extent of the *Hox-3* locus since chromosome

walking more 3 in both mouse and humans has, so far, failed to identify other *Hox-3* members (E. Boncinelli, personal communication). The expression boundary of *Hox-3.5* in the hindbrain appears to be more posterior than that of its paralogues.

## Materials and methods

### Isolation of *Hox 3.5* genomic and cDNA clones

A chromosome walk was carried out on a mouse genomic lambda (Charon 35) and cosmid libraries by screening with a *Hox-3.4* (*EcoRI-XbaI* fragment) probe (Gaunt et al., 1990). Clones were subsequently mapped and flanking regions were used for further screening. The 5' region of *Hox-3.5* was identified by hybridization to a human *HOX3E* (cp19) probe (Simeone et al., 1988) at high stringency.

The cDNA was isolated by screening a day-8.5 mouse embryonic cDNA library with the *HOX3E* and genomic *Hox-3.5* probes at high stringency.

### DNA sequencing

The *BamHI-EcoRI* fragment of genomic clone 31-1 was subcloned into pSP72 (Promega) and sequenced in both directions using a T7 DNA polymerase sequencing system (Promega). Restriction fragments of the cDNA clone 4-1 were subcloned into pSP72 and sequenced as above, using the dideoxy method described by Sanger et al. (1977).

### RNAase protection assay

The *BamHI-EcoRI* fragment of clone 31-1 (Fig. 4B) was subcloned into pSP72. <sup>32</sup>P-labelled antisense RNA probes were generated using the riboprobe system (Promega) and T7 RNA polymerase. Protection assays were performed as described by Zinn et al. (1983) using 40 µg of day-11.5 embryo total RNA and tRNA (Boehringer) as a negative control. Total RNA was prepared using the guanidinium method of Sambrook et al. (1989). The assay was run alongside a sequencing reaction as a marker.

### Primer extension assay

One picomole of the 30 mer oligonucleotide 5'-TAGCGACCCTGTAAAGTTACTTTTCACCATG-3' complementary to the sequence +86 to +116 was end-labelled with [<sup>32</sup>P]ATP and hybridized with 50 µg of liver RNA, kidney RNA and total mouse 13.5-day embryonic RNA. The primer extension assay was performed according to the method of Sambrook et al. (1989). *E. coli* tRNA (50 µg, Boehringer) was used as a negative control.

### In situ hybridization

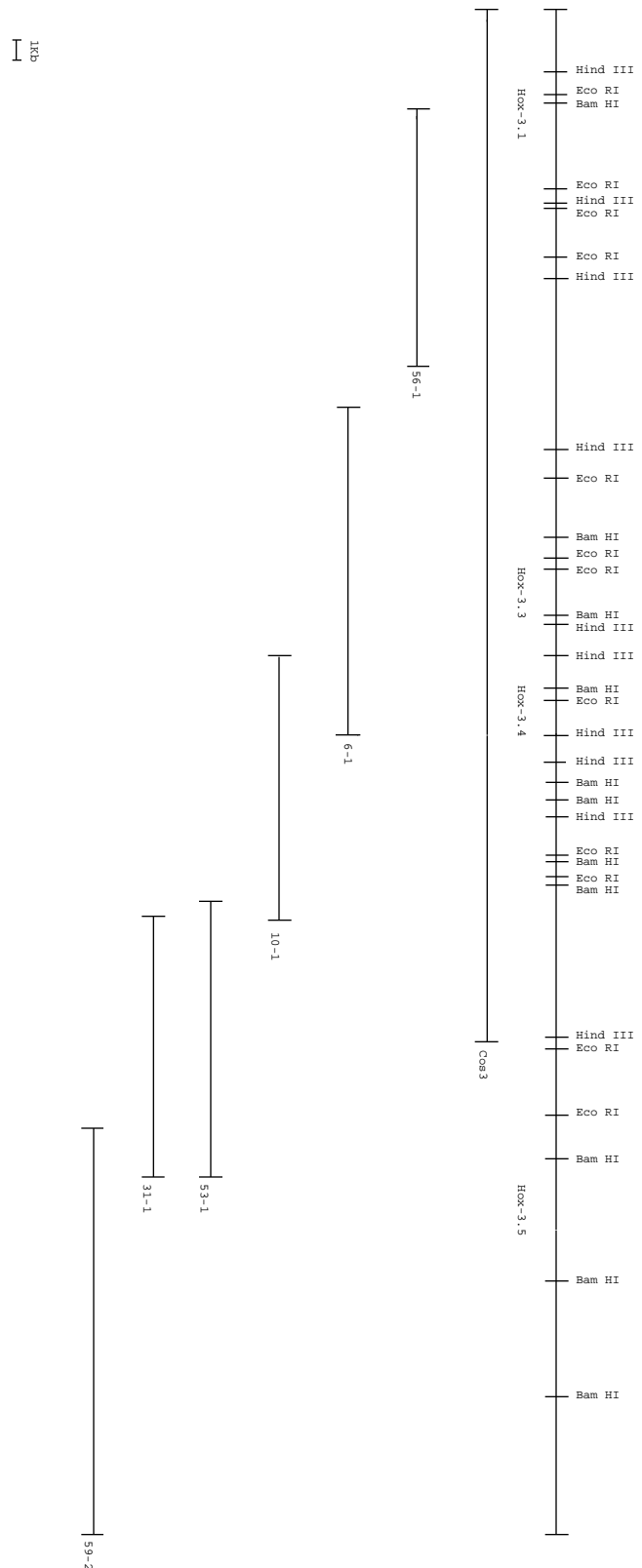
In situ hybridization to sections of mouse embryos was carried out as previously described by Sharpe et al. (1988). The full length *Hox-3.5* cDNA clone 4-1 was subcloned into pSP72 and cut with *XhoI* (nucleotide position 655 in Fig. 2) prior to preparation of antisense <sup>35</sup>S-labelled riboprobes.

## Results

### Genomic location and organisation of *Hox-3.5*

A genomic clone containing the *Hox-3.3* and *Hox-3.4* transcription units (Sharpe et al., 1988) was used as the start of a chromosomal walk in a Charon 35 genomic library. A homeobox-containing clone was identified by hybridisation

and the homeobox mapped to 18 kb 3' of *Hox-3.4*. Initial sequencing of the hybridising fragment confirmed that it



**Fig. 1.** Map of *Hox-3* cluster from *Hox-3.1* to *Hox-3.5* and clones isolated during walks. Clone Cos3 was isolated from a cosmid library (data not shown). Black boxes represent positions of the homeoboxes. Clone 59-2 was only mapped with *BamHI*.



**A**

```

Antp      RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRMRKWKKEN
Dfd       P--Q-TA---H-I-----Y-----T-V-S-----D-
Hox-1.4   P--S-TA---Q-V-----T---S---V-----DH
Hox-4.2   P--S-TA---Q-V-----T---P-----DH
Hox-2.6   P--S-TA---Q-V-----Y-----V-----S-----DH
HOX3E     P--S-AA---Q-V-----Y-----S---S-----DH
Hox-3.5   P--S-TA---Q-V-----Y-----S---S-----DH
    
```

**B**

```

1 39
Hox-3.5   miMSSYLMSDSNYIDPKFPPCEE YSQNSYIPE HSPEYYGR TR
HOX3E     -----
Hox-1.4   -T---F-IN---E-----F--FAPHGGPGGGD GAVGG-PGY P
Hox-4.2   -A---MVN-K-V----- -L-GG-LG- QGAD---GG AQ
Hox-2.6   -A---F-IN---V----- -SD-L-SD---G--AGGQRR

61
Hox-3.5   ESGFQHHHQELYPPPPP RPSYP
HOX3E     -----
Hox-1.4   RPQSAP- LPA-N-HAARQPPAYYAPRAR E----
Hox-4.2   GAD--PS G---R-DFGEQPFGGGGPSPGALPARGHGQE--G-
Hox-2.6   -----PE AAFGRRA-CTVQRYA

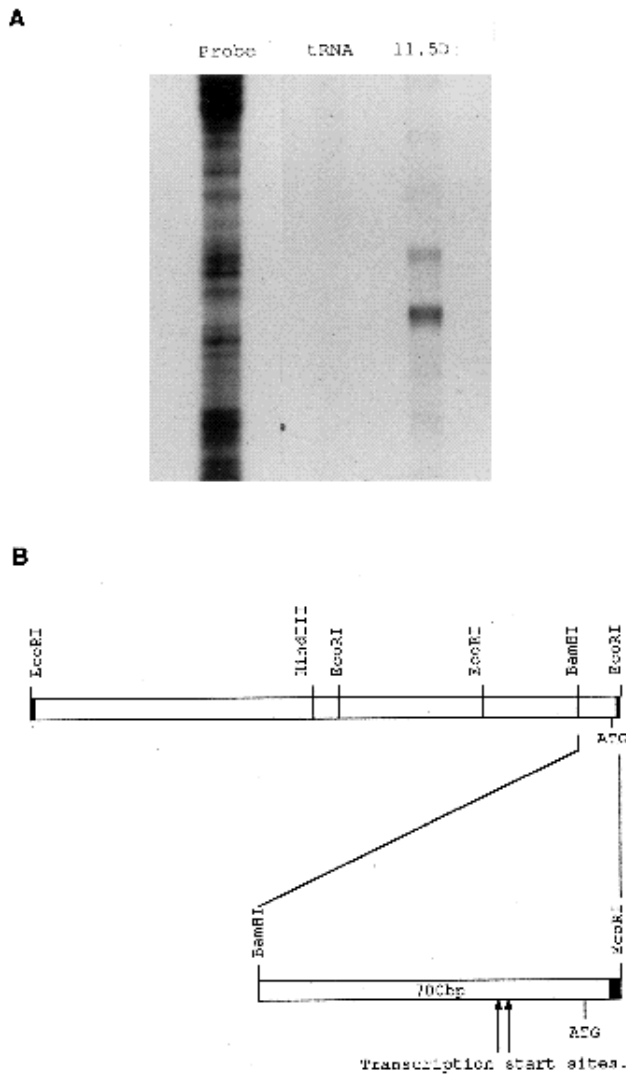
95
Hox-3.5   ERQY SCT SLQPGNSRGGHGAQAGHHHPEKSQSL
HOX3E     -----
Hox-1.4   GGL-PAPAAACPYA-RGA-PAR-EQ-PAP-AHPSPAPQ-PVPLGH
Hox-4.2   GSH-GAPGERCP -PPPAPL-
Hox-2.6   ACRDPGPPPPPPP-PPPPPP-LSPRAPVQ-TAGAL-

137
Hox-3.5   CEPAPLSGTSASPSAPPACSQ PAPDHPSSA ASLQP IVYPWM
HOX3E     -----A-----
Hox-1.4   -A-G-TTPAV-TG GSA---PL LLA-QGPAGPKG-E- V-----
Hox-4.2   GAR----PTG-KQ-PPGT-L---AV-----
Hox-2.6   P--GQR-EAVS-S -P--P-A- NPLHPSP-HS-C-E- V-----

182
Hox-3.5   KKIHVSTVNPNYNGGEPKRSRTAYTRQOVLELEKEFHYNRYLTRR
HOX3E     -----A-----
Hox-1.4   -----A-SS-----F-----
Hox-4.2   --V--NSA---T-----F-----
Hox-2.6   R-V-----A-----

227
Hox-3.5   RRIEIAHSLCLSERQIKIWFQNRMRKWKKDHRLPDTKVRSAFPAG
HOX3E     -----
Hox-1.4   -----T-----V-----K-----M--SNT-S
Hox-4.2   -----T---P-----K-----G--S S
Hox-2.6   --V---A-----K-----I--GGT--

262
Hox-3.5   AAPSTLSAATPGTSEDHSQSATPPEQORAEDITRL*
HOX3E     -----*
Hox-1.4   -PAGPPGK-QTHSPHP-PHPLPGASTPIPSS-*
Hox-4.2   SSSCSS---PGQHLQPMADHHTDLTTL*
Hox-2.6   --GGPPGRPNG-PPAL*
    
```

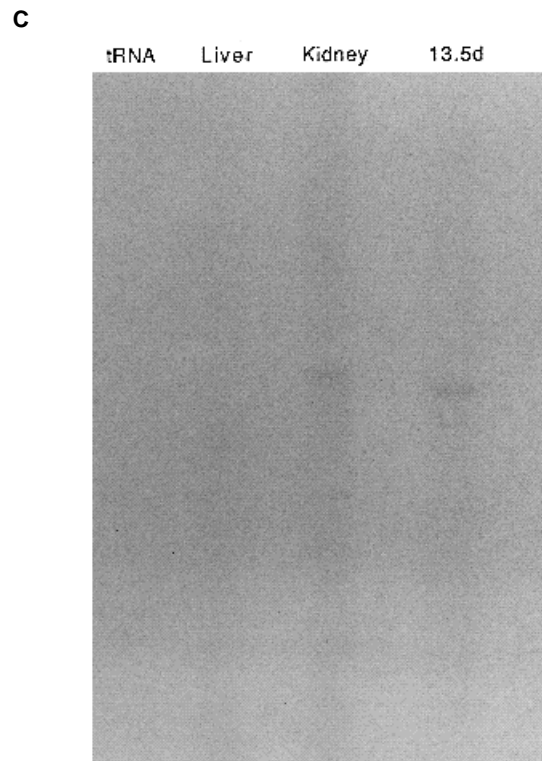


contained a homeobox and was named *Hox-3.5*. Overlapping phage and cosmid clones 5 to *Hox-3.3* were also isolated to physically link *Hox-3.3*, *Hox-3.4* and *Hox-3.5* to *Hox-3.1* (Fig. 1).

#### *Hox-3.5* sequence

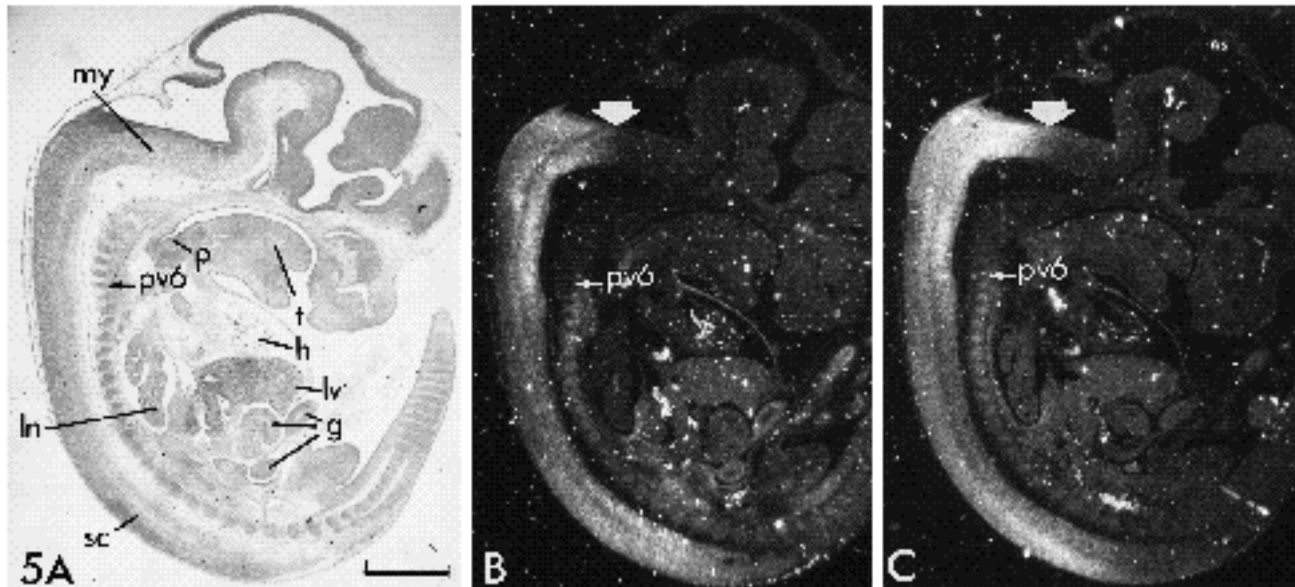
cDNAs containing the *Hox-3.5* homeobox were isolated by screening a day-8.5 embryonic cDNA library with a probe from the human cp19 gene (*HOX3E*, Simeone et al., 1988) and from a *Hox-3.5* homeobox-containing genomic fragment. A 1.4 kb cDNA was isolated and sequenced in full

**Fig. 3.** (A) A comparison of the amino acid sequences of the *Hox-3.5* homeobox and its paralogues, to the *Drosophila Dfd* and *Antp* genes, showing a similar pattern of differences between *Hox-3.5*, its paralogues and *Drosophila Dfd* in relation to the *Drosophila Antp* gene. (B) Comparison of the *Hox-3.5* predicted protein sequence with its human homologue and its paralogues in the mouse. The pentapeptide and the homeobox regions are boxed. Hyphens represent identical amino acids and spaces have been inserted to produce best possible alignment. There are two main regions of conservation between the proteins: the first 22 amino acids of the protein; the pentapeptide region through to eight amino acids after the homeobox region.



**Fig. 4.** (A) RNAase protection assay of the *BamHI-EcoRI* fragment of clone 31-1. Lane 1, free probe. Lane 2, Control lane, probe hybridized to tRNA. Lane 3, probe hybridized to day-11.5 embryo total RNA showing a main protected fragment of 205 bp and a minor protected fragment of 224 bp. (B) Map of genomic clone 31-1, showing the adaptor sites (black boxes). The *BamHI-EcoRI* fragment of the clone is enlarged showing the positions of the transcription start sites and the location of the first codon. (C) Primer extension assay of the 5' end of the *Hox-3.5*. The oligonucleotide complementary to the sequence +86 to +116 (Fig. 2) was end labelled and annealed to 50 µg of total embryo mRNA. After extension with reverse transcriptase, the extended products were separated by electrophoresis. A band of 115 bp was detected in the embryo and kidney RNA lanes. Lane 1, *E. coli* tRNA. Lane 2, Liver RNA. Lane 3, Kidney RNA. Lane 4, day-13.5 mouse embryo RNA.

(Fig. 2). The homeobox sequence of the cDNA was identical to the genomically derived sequence of the *Hox-3.5* homeobox. The homeobox sequence was also identical to the MAB87 PCR product recently identified by Murtha et al. (1991) and assumed to be the *Hox-3 Deformed (Dfd)* cognate. Comparison of the full cDNA sequence with the genomic sequence showed the cDNA to be derived from two exons separated by a 500 bp intron with the splice site being situated 35 bp 5' of the homeobox (Fig. 2). Assuming that the first inframe ATG is the initiator codon, the two exons code for an open reading frame of 264 amino acids to the inframe stop codon 47 amino acids downstream of the homeobox. Although in the majority of cases the first AUG is the initiation codon, in a few exceptions (Kozac, 1987) the first initiation codon may be ignored and translation initiated at the following AUG. The main requirement for this is the presence of an adenine nucleotide at



**Fig. 5.** (B) *Hox-3.4* and (C) *Hox-3.5* transcripts detected by in situ hybridization to nearby parasagittal sections of a 12.5-day mouse embryo. (A) Bright-field, (B,C) dark-field illumination. my, myelencephalon; ln, lung; sc, spinal cord; g, gut; lv, liver; h, heart; t, tongue; p, ventral pharynx; pv6, prevertebra 6. Large arrows in B and C show the anterior limits of expression in the myelencephalon. Bar, 1.0 mm.

position  $-3$  in respect to the primary initiation codon. As can be seen from Fig. 2, there is an adenine  $-3$  bp from the second initiation codon but not from the first codon. Initiation of translation from this second Met codon would make the first three amino acids of *Hox-3.5*, in common with many other homeobox genes, Met Ser Ser and therefore coding for a sequence of 262 amino acids. This putative *Hox-3.5* protein product is, in common with many other homeobox-containing proteins, proline (13.2%) and serine (11.7%) rich.

The position of the *Hox-3.5* gene in the *Hox-3* cluster (Fig. 1) suggests that it should be a member of the *Hox-1.4*, *Hox-2.6* and *Hox-4.2* paralogue group. This prediction is confirmed by a comparison of the *Hox-3.5* and the *Antennapedia* (*Antp*) homeodomains (Fig. 3A). The pattern of differences between these genes is characteristic of the *Dfd*-like *Hox-1.4*, *Hox-2.6* and *Hox-4.2* paralogue group (Scott et al., 1989).

The *Hox-3.5* homeodomain shows greatest homology (98%) with the human *HOX3E* gene. Although amino acid sequence homology of *Hox-3.5* to its human homologue *HOX3E* extends throughout the predicted protein sequence, this is not true of the paralogue members where amino acid conservation is restricted to two main areas (Fig. 3B). The first of these is the N terminus of the protein extending from three amino acids upstream of the translation start site to the first 22 amino acids. The second area of conservation begins at the pentapeptide and extends through to eight amino acids after the homeodomain. The *Hox-3.5* transcript has a large 3 untranslated sequence of at least 626 bp. Since the cDNA did not contain any obvious polyadenylation signals, the full extent of the 3' sequence is not determined.

The transcription start site was mapped by two separate methods. The first, RNAase protection assay identified two sites which were mapped approximately to positions 170

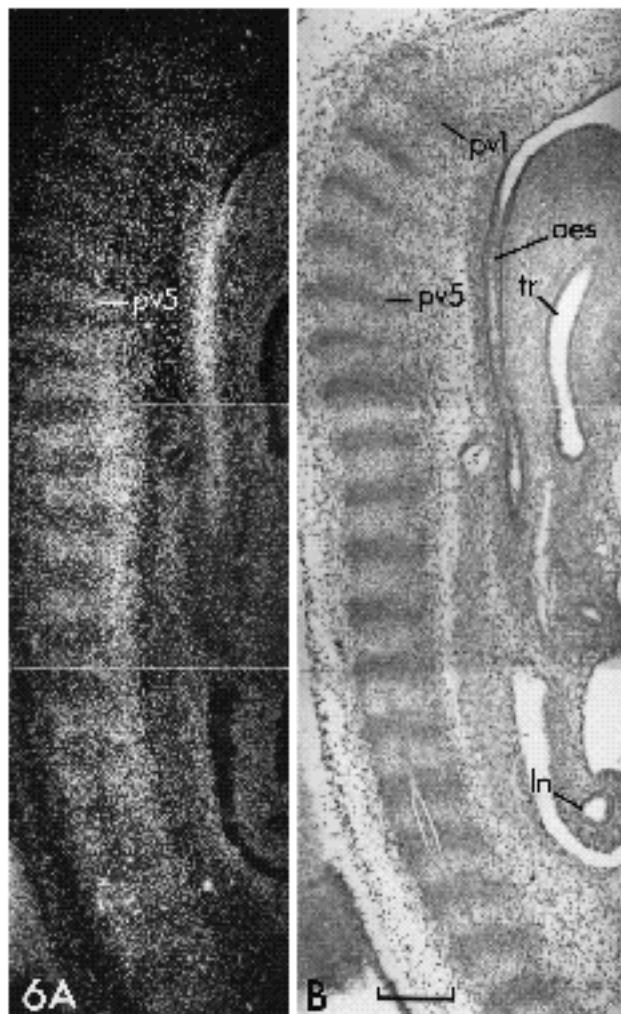
and 189 bp upstream of the translation start site (Fig. 4A). The second, primer extension assay, accurately mapped a single transcription start site at 197 bp upstream of the translation start site (Fig. 4C). This 197 bp start position probably corresponds to the 189 bp position mapped approximately by RNAase protection. The 170 bp RNAase protected fragment not detected by primer extension may be an alternative start site only detected with the more sensitive assay, or may be a degradation product. In common with many homeobox genes (Smale and Baltimore, 1989) no obvious TATA or CAAT boxes were identified 5' of the start sites.

#### *Embryonic expression of Hox-3.5*

The expression pattern for *Hox-3.5*, detected by in situ hybridization, is considered in terms of three rules that have been formulated to relate a Hox gene's expression to its cluster location (reviewed by Gaunt, 1991).

Consistent with Rule one (Hox genes and their expression domains are collinear), the anterior boundary of *Hox-3.5* expression (Fig. 5C) was found to lie anterior to that of *Hox-3.4* (Fig. 5B), both in the 12.5-day hindbrain and prevertebral column. Within the prevertebral column (Fig. 6), *Hox-3.5* transcripts increased in abundance over prevertebrae 5-7 (pv 5-7), were most abundant over pv 7-13/14, and then declined more posteriorly. No evidence for *Hox-3.5* labelling above background was detected in pv 1-3, but several sections showed low levels of transcripts apparently present in ventral parts of pv 4. The anterior boundary of *Hox-3.5* expression was therefore located anterior, by one prevertebra to that found earlier for *Hox-3.4* (Gaunt et al., 1990).

Regarding Rule two (anterior boundaries of expression may be conserved between paralogous genes, especially those 3' in location), Fig. 7 compares, at 10.5 days, the



**Fig. 6.** *Hox-3.5* expression within the 12.5-day prevertebral column and oesophagus. (A) Dark-field. (B) bright-field illumination. pv1,5, prevertebrae 1 and 5; oes, oesophagus; tr, trachea; ln, lung. Bar, 0.2 mm.

hindbrain boundary for *Hox-3.5* with that of its paralogue, *Hox-2.6*. The anterior boundary of *Hox-3.5* expression was found to lie posterior to that of *Hox-2.6*. The *Hox-2.6* boundary (Fig. 7F,H) clearly coincided with a rhombomere junction (the anterior boundary of rhombomere 7; Wilkinson et al., 1989). In contrast, the *Hox-3.5* boundary (Fig. 7B,D) did not coincide with a morphologically visible rhombomere junction, but it was located posterior to the position of the *Hox-2.6* boundary by a distance equal to about the length of one rhombomere. An apparently similar result for the position of the *Hox-3.5* boundary in the hindbrain was also observed at 9.5 days (Fig. 8).

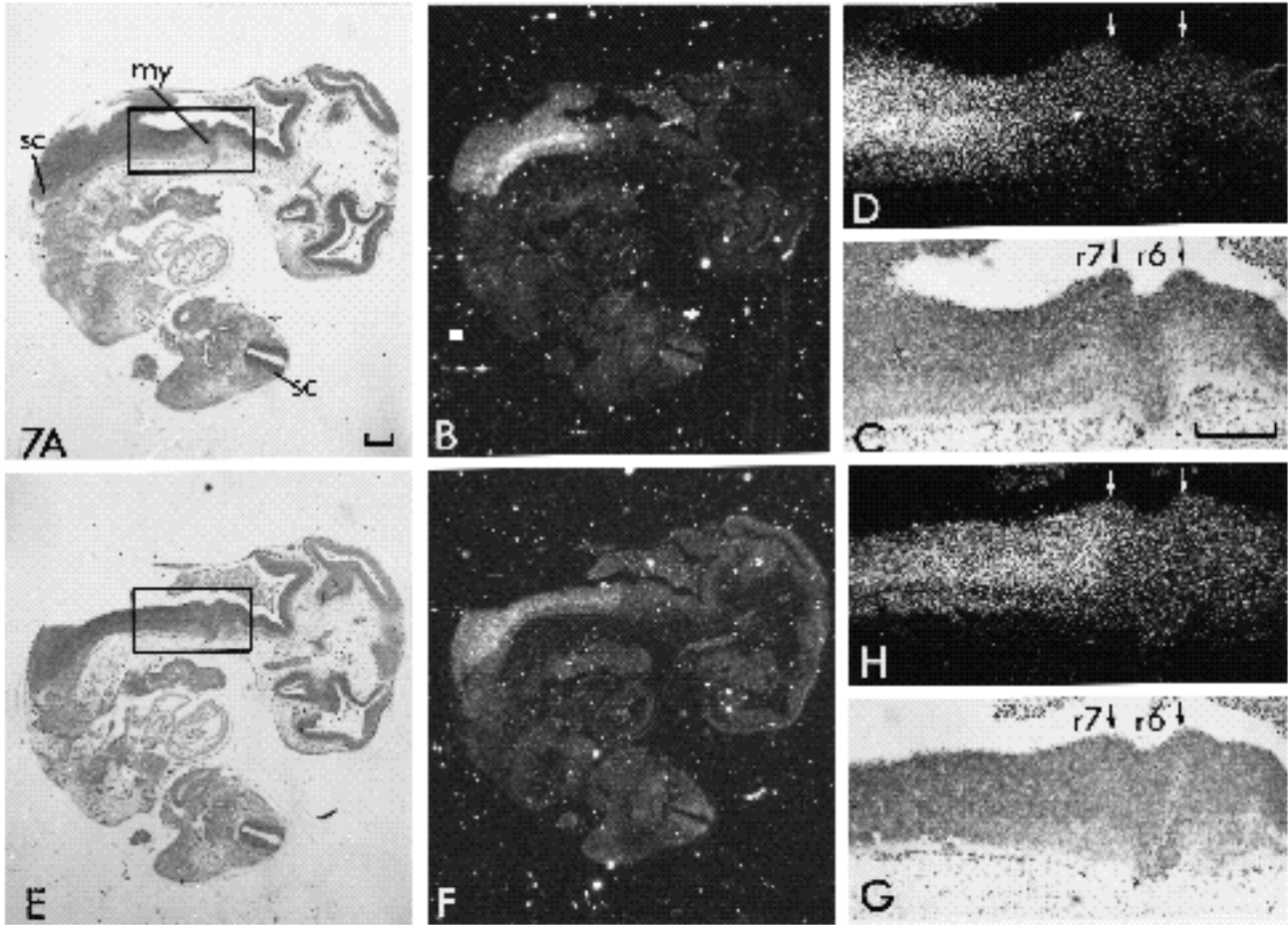
Regarding Rule three (the tissue specificity in expression of a Hox gene may vary according to its chromosomal cluster), *Hox-3.5* transcripts were, like *Hox-3.4* (Gaunt et al., 1990), abundant in the 12.5-day oesophagus, and detected at only very low levels in the trachea and lung (Fig. 6). So far, expression in the oesophagus has been detected only for *Hox-3.4* and *Hox-3.5*, and has not been observed for any *Hox-1*, *Hox-2* or *Hox-4* genes (Gaunt et al., 1988, 1989). Fig. 9 compares expression of *Hox-3.5* with that of

its paralogues, *Hox-2.6* and *Hox-1.4*, as detected on nearby transverse sections through the 12.5-day cervical spinal cord. The patterns for *Hox-2.6* and *Hox-1.4* are as previously described (Gaunt et al., 1990), and appear to be characteristic for *Hox-2* and *Hox-1* genes, respectively. The pattern for *Hox-3.5* most resembles that described earlier for *Hox-3.4* and *Hox-3.3* (Gaunt et al., 1990), with transcripts being abundant both ventrally and centrally within the mantle layer. Medially located transcripts were, however, found to be more abundant for *Hox-3.5* (Fig. 8B) than have been noted earlier for other *Hox-3* genes (Gaunt et al., 1990). *Hox-3.5* transcripts were detected within the 12.5-day metanephric kidney but not within the adjacent testis (Fig. 10B). This differs from findings made for *Hox-3.3* and *Hox-3.4*, whose transcripts are abundant within the testis, but is similar to results obtained for *Hox-3.1* (Gaunt et al., 1990).

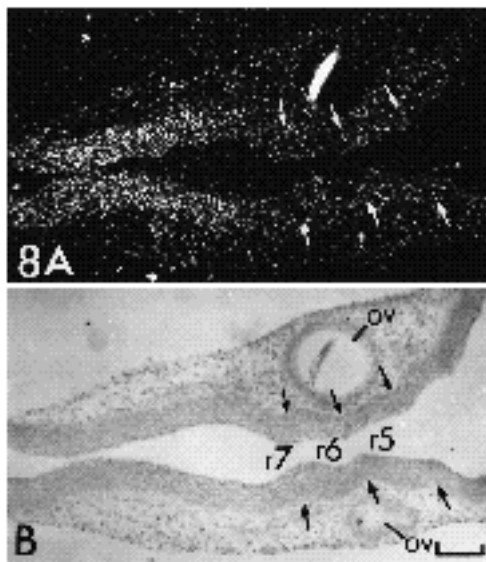
## Discussion

The *Hox-3.5* gene is located 18 kb 3' of *Hox-3.4* on chromosome 15. Sequencing of genomic and cDNA fragments show that transcripts isolated from a day-8.5 embryonic cDNA library are derived from two exons separated by a 500 bp intron. No other *Hox-3.5* transcripts were isolated from this library. Simeone et al. (1988) reported that the human homologue of *Hox-3.5*, *HOX3E*, is differentially spliced in placental tissue by insertion of a 5' exon located 12 kb 5' to *HOX3C*. Although we have observed differential splicing with this 5' exon in *Hox-3.3* transcripts (unpublished data), we have repeatedly failed to detect similarly spliced transcripts in embryos of either *Hox-3.4* or *Hox-3.5*. In common with other class I homeobox genes, the 3' exon codes for the conserved 61 amino acid homeodomain which has been shown to have a helix-turn-helix structure (Qian et al., 1989) and which is able to bind to DNA (Mihara and Kaiser, 1988) and act as a transcription factor (Kuziora and McGinnis, 1991). The 5' exon contains two regions of conservation, one at either extreme seen in many Hox genes. The first spans the first 22 amino acids and contains the triplet Met Ser Ser which is a common translation start sequence in many homeobox-containing genes. The second region of conservation begins nine amino acids upstream of the splice site and contains another typical feature of these genes, the pentapeptide. The region between these two areas of conservation in the 5' exon is unconserved and is proline and serine rich. It has been shown that proline-rich areas are unable to produce a helical structure.

Hox genes are arranged in the mouse along four clusters *Hox-1* (chromosome 6), *Hox-2* (chromosome 11), *Hox-3* (chromosome 15) and *Hox-4* (chromosome 2) and it has been shown that highest homologies are to be found in genes having the same location along these clusters (i.e. paralogue groups). Hence, *Hox-3.5* has highest homology to genes *Hox-1.4* (Galliot et al., 1989), *Hox-2.6* (Graham et al., 1988) and *Hox-4.2* (Featherstone et al., 1988; Duboule et al., 1990). *Hox-3.5* also has similar characteristic amino acid differences in the homeodomain from the archetypal *Antennapedia* sequence as *Hox-1.4*, *Hox-2.6* and



**Fig. 7.** (A-D) *Hox-3.5*, and (E-H) *Hox-2.6* transcript boundaries within the 10.5-day hindbrain. The transcripts from these two paralogous genes are compared on nearby sagittal sections. (A,C,E,G) Bright-field, (B,D,F,H) dark-field illumination. The higher power fields shown in C,D and G,H are shown in boxes on A and E respectively. my, myelencephalon; sc, spinal cord; r6,7, rhombomere 6 and 7; arrows, junctions between rhombomeres. Rhombomeres 6 and 7 were identified according to the known position of the anterior boundary for *Hox-2.6* (at the junction of rhombomeres 6 and 7; Wilkinson et al., 1989). Bars, 0.25 mm.

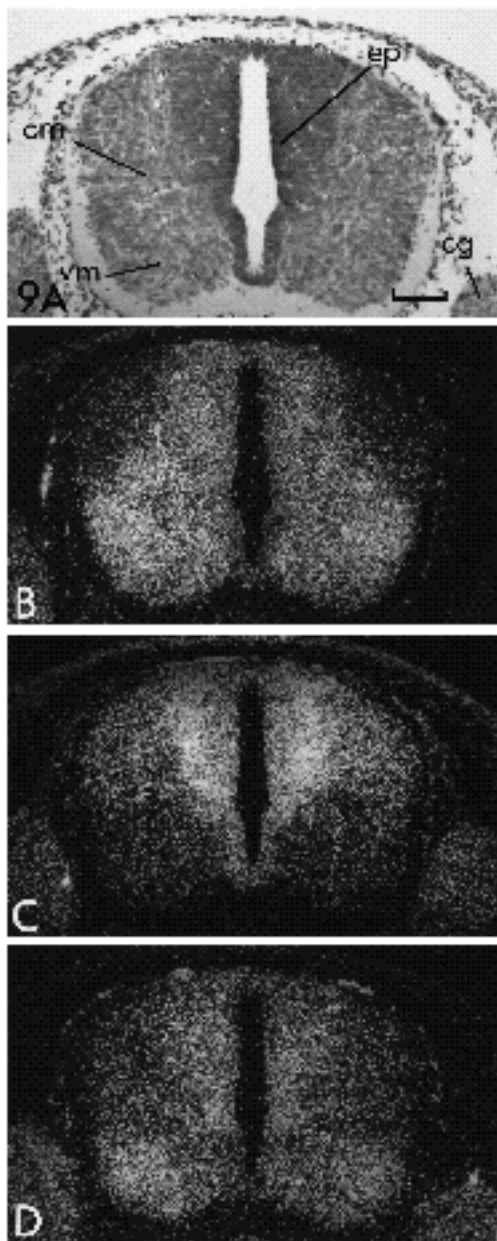


*Hox-4.2*. This conservation between paralogues appears to be due to a series of duplications and divergences of an ancestral homeobox cluster (Graham et al., 1989; Kappen et al., 1989). The *Drosophila* cognate of *Hox-3.5* and its paralogues is the *Deformed* (*Dfd*) gene, which also has a high degree of conservation in the homeodomain sequence, and hence the *Hox-3.5* paralogue group has also been named the *Deformed*-like subgroup. The human cognate of *Hox-3.5*, *HOX3E* (Simeone et al., 1988) is highly homologous in the entirety of its protein coding region, and this homology extends for approximately 200 bp in both 5' and 3' untranslated regions of the gene.

The transcription start site mapped, in common with many other homeobox genes (Smale and Baltimore, 1989), is not associated with a common basic promoter motif such as TATA or CAAG or a G/C rich area (G/C content in the preceding 400 bp is only 64%).

**Fig. 8.** *Hox-3.5* transcript boundary located on a coronal section of the 9.5-day hindbrain. (A) Dark-field, (B) bright-field illumination. ov, otic vesicle; arrows, junctions between rhombomeres; r5, 6 and 7, rhombomeres 5, 6 and 7. Bar, 0.1 mm.

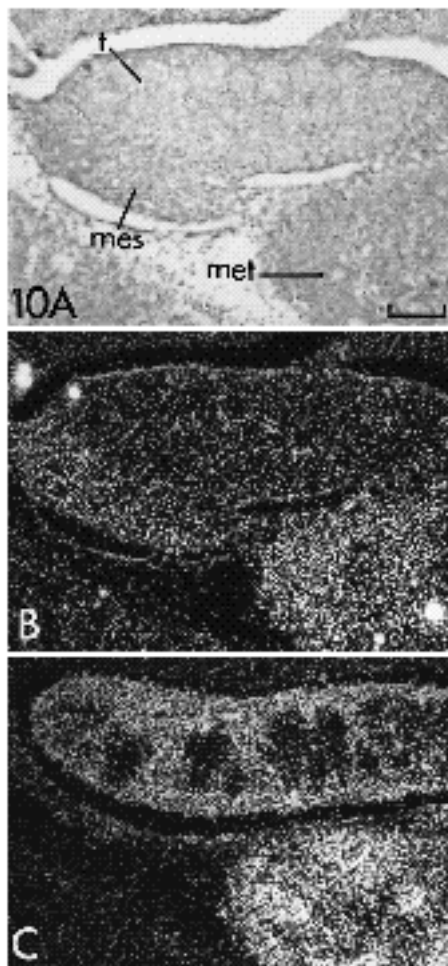




**Fig. 9.** (B) *Hox-3.5*, (C) *Hox-2.6*, (D) *Hox-1.4* expression detected on transverse sections through the mid-cervical region of the 12.5-day spinal cord. Transcripts from these three paralogous genes are compared here on adjacent sections. (A) Bright-field, (B-D) dark-field illumination. ep, ependymal layer; cm, vm, central and ventral regions of the mantle layer; cg, cervical ganglion. Bar, 0.1 mm.

The distribution of *Hox-3.5* transcripts, detected by in situ hybridization, was found to be generally in-keeping with three rules that have earlier been suggested (Gaunt, 1991) to relate a Hox gene's expression pattern to its cluster location. However, two points, considered below, merit further discussion.

First, Hunt et al. (1991) recently assumed, though did not demonstrate, that *Hox-3.5*, like other paralogues in its sub-family (*Hox-2.6*, *Hox-4.2* and *Hox-1.4*) has its anterior boundary of expression at the rostral limits of rhombomere



**Fig. 10.** (B) *Hox-3.5*, and (C) *Hox-4.2* expression within the 12.5-day testis and kidney. Transcripts from these two paralogous genes are compared on nearby sections from the same embryo. (A) Bright-field, (B,C) dark-field illumination. met, mes, metanephric and mesonephric kidneys; t, testis. Bar, 0.1 mm.

7. This assumption was based on a view that paralogous genes display identical boundaries of expression within the hindbrain. In apparent contradiction of this conclusion, we now show that the anterior boundary of *Hox-3.5* transcripts is seen to lie posterior to that of *Hox-2.6* by a distance equal to about the length of one rhombomere. This boundary remains constant between 9.5 and 10.5 days suggesting that *Hox-3.5* expression does not start more anteriorly and regress as has been reported for *Hox-1.6* (Murphy and Hill, 1991). As one possible explanation for this finding *Hox-3.5* may have, during the course of evolution, acquired a more posterior position for its boundary than that observed by its three paralogues. It may be of significance in this respect that *Hox-3.5* appears to mark the 3-most extent of the *Hox-3* cluster. Thus, more anterior genes, present in the other clusters, are not present in *Hox-3* and have either failed to be duplicated, been lost, or been relocated during evolution. Whatever the case, *Hox-3* genes do not contribute in the development of the hindbrain region to the same extent as the *Hox-1*, *Hox-2* or *Hox-4* genes. As a second possibility, however, *Hox-3.5* expression may be down-regulated in rhombomere 7, and the reduced levels of *Hox-3.5* tran-

scripts within this rhombomere (as noted for *Hox-1.4*; Hunt et al., 1991) may not be detected in our in situ hybridizations.

Another unexpected finding was the absence of *Hox-3.5* transcripts within the 12.5-day testis. This differs from findings made earlier for *Hox-3.3* and *Hox-3.4* (Gaunt et al., 1990), but is similar to the finding for *Hox-3.1* (Gaunt et al., 1990). Expression in the testis is apparently limited only to *Hox-3* and *Hox-4* genes. It is of interest that *Hox-3.5* and *Hox-3.1* both have paralogues within the *Hox-4* cluster that show strong expression within the testis (Izpisua-Belmonte et al., 1990). In contrast, *Hox-3.3* and *Hox-3.4* do not have corresponding paralogues within the *Hox-4* cluster. These observations suggest the interesting possibility that paralogous genes from different clusters may be coordinately regulated in their patterns of tissue-specific expression.

We thank Edoardo Boncinelli for the gift of the cp19 probe and for communicating unpublished details of the extent of the human *HOX-3* locus. We also thank Wolf Reik for his assistance in the isolation of the cosmid clone, Cos3, and the MRC for financial support.

## References

- Breier, G., Dressler, R. and Gruss, P. (1988). Primary structure and developmental expression pattern of *Hox-3.1*, a member of the murine *Hox-3* homeobox gene cluster. *EMBO J.* **7**, 1329-1336.
- Bucan, M., Yang-Feng, T., Colberg-Poley, A. M., Wolgemuth, D. J., Guenet, J. L., Franke, U. and Lehrach, H. (1986). Genetic and cytogenetic localisation of the homeo box containing genes on mouse chromosome 6 and human chromosome 7. *EMBO J.* **5**, 2899-2905.
- Chisaka, O. and Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from target disruption of the mouse homeobox gene *Hox-1.5*. *Nature* **350**, 473-479.
- Chisaka, O., Musci, T. S. and Capecchi, M. R. (1992). Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene *Hox-1.6*. *Nature* **355**, 516-520.
- Duboule, D., Boncinelli, E., De Robertis, E., Featherstone, M., Lonai, P., Oliver, G. and Ruddle, F. H. (1990). An update of mouse and human HOX gene nomenclature. *Genomics* **7**, 458-459.
- Duboule, D. and Dollé, P. (1989). The structural and functional organisation of the murine Hox gene family resembles that of the *Drosophila* homeotic genes. *EMBO J.* **8**, 1497-1505.
- Featherstone, M. S., Baron, A., Gaunt, S. J., Mattei, M. G. and Duboule, D. (1988). *Hox-5.1* defines a homeobox-containing gene locus on the mouse chromosome 2. *Proc. Natn. Acad. Sci. U.S.A.* **85**, 4760-4764.
- Galliot, B., Dollé, P., Vigneron, M., Featherstone, M. S., Baron, A. and Duboule, D. (1989). The mouse *Hox-1.4* gene: primary structure, evidence for promoter activity and expression during development. *Development* **107**, 343-359.
- Gaunt, S. J. (1991). Expression patterns of mouse Hox genes: Clues to an understanding of developmental and evolutionary strategies. *BioEssays* **13**, 505-512.
- Gaunt, S. J., Coletta, P. L., Pratcheva, D. and Sharpe, P. T. (1990). Mouse *Hox-3.4*: homeobox sequence and embryonic expression patterns compared with other members of the Hox gene network. *Development* **109**, 329-339.
- 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**Supplement, 169-179.
- 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. M., Tuddenham, E. G. D. and Krumlauf, R. (1988). Characterisation of a murine homeobox gene, *Hox-2.6*, related to the *Drosophila Deformed* gene. *Genes Dev.* **2**, 1424-1438.
- Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. and Ruddle, F. H. (1985). Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell* **43**, 9-18.
- Hunt, P., Gullisano, 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.
- Hunt, P. and Krumlauf, R. (1991). Dicephering the Hox code: clues to patterning branchial regions of the head. *Cell* **66**, 1075-1078.
- Izpisua-Belmonte, J. C., Dollé, P., Renucci, A., Zappavigna, V., Falkenstein, H. and Duboule, D. (1990). Primary structure and embryonic expression pattern of the mouse *Hox-4.3* homeobox gene. *Development* **110**, 733-745.
- Kappen, C., Schughart, K. and Ruddle, F. H. (1989). Two steps in the evolution of *Antennapedia*-class vertebrate homeobox genes. *Proc. Natn. Acad. Sci. U.S.A.* **86**, 5459-5463.
- Kessel, M. and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alterations of Hox codes induced by retinoic acid. *Cell* **67**, 89-104.
- Kozac, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**, 8125-8148.
- Kuziora, M. A. and McGinnis, W. (1991). Altering the regulatory targets of the *Deformed* protein in *Drosophila* embryos by substituting the *Abdominal-B* homeodomain. *Mech. Dev.* **33**, 83-94.
- Lufkin, T., Dierich, A., LeMeur, 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.
- Mihara, H. and Kaiser, E. T. (1988). A chemically synthesised *Antennapedia* homeo domain binds to a specific DNA sequence. *Science* **242**, 925-927.
- Murphy, P. and Hill, R. E. (1991). Expression of the mouse *labial*-like homeobox-containing genes, *Hox-2.9* and *Hox-1.6*, during segmentation of the hindbrain. *Development* **111**, 61-74.
- Murtha, M. T., Leckman, J. F. and Ruddle, F. H. (1991). Detection of homeobox genes in development and evolution. *Proc. Natn. Acad. Sci. U.S.A.* **88**, 10711-10715.
- Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J. and Wuthrich, K. (1989). The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59**, 573-580.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning, A Laboratory Manual*. Second Edition, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Sanger, F., Miklen, S. and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natn. Acad. Sci. U.S.A.* **74**, 5463-5467.
- Scott, M. P., Tamkun, J. W. and Hartzel, G. W. (1989). The structure and function of the homeodomain. *Bioch. Biophys. Acta* **989**, 25-48.
- Sharpe, P. T., Miller, J. R., Evans, E. P., Burtenshaw, M. D. and Gaunt, S. J. (1988). Isolation and expression of a new mouse homeobox gene. *Development* **102**, 397-407.
- 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. *Nucl. Acids Res.* **16**, 329-339.
- Smale, S. T. and Baltimore, D. (1989). The "initiator" as a transcription control element. *Cell* **57**, 103-113.
- 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.
- Zinn, K., Di Maio, D. and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human  $\beta$ -interferon gene. *Cell* **34**, 865-879.

*(Accepted 27 July 1992)*