Development 102, 397–407 (1988)
Printed in Great Britain © The Company of Biologists Limited 1988

Isolation and expression of a new mouse homeobox gene

P. T. SHARPE[1,*], J. R. MILLER[1], E. P. EVANS[2], M. D. BURTENSHAW[2] and S. J. GAUNT[1]

[1]*Department of Molecular Embryology, AFRC Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK*
[2]*Sir William Dunn School of Pathology, South Parks Road, Oxford, UK*

* Present address and address for correspondence: Department of Cell & Structural Biology, Stopford Building, University of Manchester, Manchester M13 9PT, UK

Summary

A homeobox-containing clone has been isolated from an adult mouse kidney cDNA library and shown by DNA sequence analysis to be a new isolate, *Hox-6.1*†. A genomic clone containing *Hox-6.1* has been isolated and found to contain another putative homeobox sequence (*Hox-6.2*), within 7 kb of *Hox-6.1*. *In situ* hybridization of mouse metaphase chromosomes shows this Hox-6 locus to be located on chromosome 14 (14E2).

***Hox-6.1* has been studied in detail and the predicted protein sequence of the homeobox is 100 % homologous to the *Xenopus Xebl* (formally *AC1*) homeobox and the human *c8* homeobox (Carrasco *et al.* 1984; Boncinelli *et al.* 1985; Simeone *et al.* 1987). Southern blotting shows that the DNA sequence encoding *Hox-6.1* is single copy.**

Expression of *Hox-6.1* has been studied in adult tissues and embryos by RNase protection assays, Northern blotting analysis and *in situ* hybridization. RNase protection assays show that *Hox-6.1* transcripts are present in embryos between days 9½ and 13½ of gestation and in extraembryonic tissues at day 9½. Adult expression is detectable in kidney and testis but not in liver, spleen and brain. One major transcript is detectable on Northern blots of kidney and day-13½ embryo RNA. In kidney, this transcript is 2·7 kb whereas in embryos the major transcript is smaller at 1·9 kb, a much fainter band being visible at 2·7 kb. Localized expression of *Hox-6.1* is observed in the spinal cord and prevertebral column of day-12½ embryos, and in the posterior mesoderm and ectoderm of day-8¼ embryos. An anterior boundary of expression is located just behind the hindbrain whereas the boundary in the mesoderm is located at the level of the 7th prevertebra.

† The locus has been designated according to the nomenclature system accepted by the International Committee for Standardized Genetic Nomenclature and agreed with T. Roderick, Jackson Laboratory.

Key words: homeobox, Hox-6 complex, chromosome 14, *Mus musculus*, DNA.

Introduction

Many of the genes that appear to play key roles in the control of embryonic development in *Drosophila*, such as segmentation and homeotic genes, have been found to share a conserved DNA sequence of 180 nucleotides termed the homeobox (Gehring 1985; McGinnis *et al.* 1984).

Sequence analysis demonstrated that this sequence is translated into protein and that the predicted homeobox protein sequences show greater homology than the corresponding DNA sequence, implying conservation of a functional domain. This protein domain is believed to be involved in DNA binding and shows discernible homology to several characterized DNA-binding proteins (Shepherd *et al.* 1984). Two further characteristics of homeobox-containing genes are that the genes occur in clusters in the Antennapedia and Bithorax loci and that homeobox gene expression is observed in embryos (McGinnis *et al.* 1984). Homeobox sequences have been demonstrated to exist in a wide variety of nondipteran genomes including sea urchin, *Xenopus*, mouse and man (Dolecki *et al.* 1986; Harvey *et al.* 1986; Hart *et al.*

1985; Boncinelli *et al.* 1985). In the mouse, at least 15 homeobox sequences have been identified. These sequences are all potentially protein coding and once again the predicted protein sequences show greater conservation than the nucleotide sequences. Further work has shown that mammalian homeobox genes are expressed during embryogenesis and frequently occur in clusters (Hart *et al.* 1985). At least two loci containing five or more homeoboxes have been described. Hox-1 on chromosome 6 and Hox-2 on chromosome 11 (Hart *et al.* 1985; Duboule *et al.* 1986). Thus, the mouse homeobox genes show many parallels with their *Drosophila* counterparts and must be considered as possible developmental control genes. Here we describe the identification of another mouse homeobox-gene locus, Hox-6, on chromosome 14. One gene of this locus, *Hox-6.1*, has been studied in detail and appears to be a homologue of the *Xenopus Xeb1* homeobox gene and the human *c8* gene (Carrasco *et al.* 1984; Simeone *et al.* 1987).

Materials and methods

cDNA library production and screening

Adult F_1 male mouse kidney total RNA was extracted by precipitation in guanidine thiocyanate (Chirgwin *et al.* 1979). Poly(A)$^+$ RNA was extracted by oligo(dT)-cellulose affinity column chromatography (Aviv & Leder, 1972). 10 μg of poly(A)$^+$ RNA were used to construct a cDNA library λgt10 (Watson & Jackson, 1985). The library consisting of 5×10^5 recombinant phage was amplified and screened with a nick-translated 220 bp probe derived from the murine *Hox-1.5* gene containing 145 bp of homeobox sequence (McGinnis *et al.* 1984). One of the positive clones isolated after three successive screenings was mapped and restriction fragments subcloned. Hybridizations were carried out under low stringency conditions, 1 M-NaCl, 1 % SDS, 10 % dextran sulphate, 10 μg ml^{-1} salmon sperm DNA at 60 °C for 12 h.

Restriction mapping, subcloning and sequencing

1–2 μg of phage DNA from a single plaque was digested with restriction endonucleases, separated on a 0·8 % Tris–acetate agarose gels and transferred to Gene Screen plus membranes (NEN). Hybridizations with the *Hox-1.5* fragment were carried out using the same conditions as for library screening. Probes were nick translated using ^{32}P-labelled dCTP (Amersham). The whole 540 bp *Eco*RI gel-purified insert of the clone was ligated into M13mp8. Single-stranded templates were prepared and sequenced by the dideoxy method (Sanger *et al.* 1977). Reverse primer sequencing was used on a double-stranded template to confirm the sequence (Hong, 1981). The 490 bp *Eco*RI/*Bgl*II fragment of the insert was subcloned into the *Eco*RI/*Bam*HI sites of pGEM 4 (Promega Biotech) for production of riboprobes.

RNase protection assays

RNase protection was carried out essentially as described by Zinn *et al.* (1983). Total RNAs were extracted as described from adult F_1 male mouse tissues. RNA was extracted from embryos and extraembryonic tissues obtained from natural matings between male F_1 and female CFLP mice. Extraembryonic tissue consisted of allantois and amnion dissected from the embryo and decidua and contained no visible trophoblast. Midday on the day of the copulation plug was designated day $\frac{1}{2}$ of pregnancy, making the assumption that mating had taken place at midnight. 50 μg samples of RNA were dissolved in 30 μl of 80 % formamide, 40 mM-Pipes buffer pH 6·4, 400 mM-NaCl and 1 mM-EDTA. Radioactive RNA probes complementary to Hox-6.1 mRNA were prepared after digestion of pGEM4 plasmid DNA with *Hin*dIII followed by *in vitro* transcription using SP6 RNA polymerase with ^{32}P-UTP (Amersham; 50 μCi, 0·2 μg per 20 μl reaction). The specific activity obtained was about $1{\cdot}0\times10^8$ disints min^{-1} μg^{-1}. 5×10^5 cts min^{-1} of riboprobe was added, heated at 85 °C for 10 min and incubated overnight at 45 °C. 350 μl of 10 mM-Tris–HCl (pH 7·5), 5 mM-EDTA, 300 mM-NaCl containing 40 μg ml^{-1} RNase A and 100 i.u. ml^{-1} RNase T1 (Sigma), previously boiled for 5 min and allowed to cool were added and incubated at 30 °C for 30 min. 2·5 μl of proteinase K (20 mg ml^{-1}) and 20 μl 10 % SDS were added and incubated at 37 °C for 15 min. Following phenol/chloroform extraction, the RNAs were ethanol precipitated at −70 °C with 5 μg tRNA as carrier. The precipitates were washed in ethanol, dried and resuspended in 4 μl 50 % formamide, 10 mM-EDTA, 0·5 % xylene cyanol and bromophenol blue. After heating at 65 °C for 5 min, 2 μl of each sample was loaded onto a 6 % acrylamide/urea sequencing gel and run at 30 V cm^{-1}. The gel was fixed in 10 % methanol/10 % acetic acid, and dried onto 3MM paper before autoradiography.

Hybridization to mouse genomic DNA

5 μg of mouse DNA was digested with *Eco*RI, *Hin*dIII or *Bam*HI and electrophoresed on a 0·6 % agarose gel. DNA was transferred to Gene Screen plus membranes by the method of Southern (1975). Nick-translated probes derived from *Eco*RI/*Bgl*II fragment of the insert were hybridized under conditions of high stringency (1 M-NaCl, 1 % SDS, 10 % poly(ethylene glycol) at 65 °C and washed under similar stringent conditions of 2×SSC (0·3 M-sodium chloride/0·03 M-sodium citrate), 1 % SDS at 65 °C.

Northern blot analysis

50 μg of total RNA from adult kidney, liver and day-13$\frac{1}{2}$ embryos were electrophoresed at 250 V for 2 h in 1×MOPS buffer. The gels were blotted onto Hybond N membranes according to the Amersham protocol. Blots were prehybridized in 5×SSPE, 50 % formamide, 5×Denhardt's solution and 0·5 % SDS containing denatured salmon sperm DNA at 500 μg ml^{-1} at 42 °C for 2 h. ^{32}P-labelled anti-sense riboprobe was added and hybridized overnight at 42 °C. Blots were washed in 0·1×SSPE, 0·1 % SDS at 65 °C for 30 min before incubation with RNase A at 10 μg ml^{-1} in 2× SSPE at 37 °C for 10 min. The RNase A treatment was required to remove all nonspecifically bound probe which

was monitored by the control sample of liver RNA which had been shown by RNase protection assays (Fig. 4) not to contain transcripts of Hox-6.1.

A combination of pBr322 digests and rehybridization of blots with rRNA probes was used as molecular weight markers.

Isolation of Hox 6.1 genomic clone

A mouse genomic library in Charon 35 was screened under conditions of high stringency with a probe derived from a nonhomeobox-containing fragment of a *Hox-6.1* cDNA clone. Five positive plaques were identified and one of these was mapped by restriction digests of the phage, Southern blotting and probing with the *Hox-6.1* fragment. The blots were washed and reprobed with the *Hox-1.5*-derived fragment to identify other homeobox sequences.

In situ *chromosome hybridization*

Mitotic preparations were obtained from third subculture mouse embryonic fibroblasts in a growth phase and the chromosomes banded for recognition after *in situ* hybridization by substitution with BrdU using an adaptation of a method described for human cells (Zabel *et al.* 1983). Slides were probed with the entire 540 bp insert labelled with tritium using the method described by Lyon *et al.* (1986), dipped in Ilford Nuclear Emulsion L4 and exposed for between 15 and 28 days before developing in D19 (Kodak). After developing, replication bands were revealed by treating the slides with Hoechst 33258 at a concentration of 10 $\mu g\,ml^{-1}$ in 2×SSC for 30 min, placing in 2×SSC for 1 h at a distance of 15 cm away from a long-wave u.v. light and finally staining in 5–10 % Giemsa (Merck) in pH 6·8 buffer.

The slides were scored under a ×100 oil immersion lens by counting the total number of grains overlying all the chromosomes in a metaphase spread and relating grain positions with respect to the chromosomes and G-bands as identified from the standard nomenclature for the mouse karyotype (Nesbitt & Francke, 1973).

In situ *hybridization to embryo sections*

^{35}S-labelled probes for use in *in situ* hybridization experiments were prepared from the 490 bp *Eco*RI/*Bgl*II fragment of *Hox-6.1* subcloned in pGEM4. The probe (antisense probe) used for hybridization with Hox-6.1 mRNA was synthesized in an SP6 RNA polymerase reaction after linearization of the plasmid with *Hin*dIII. In some experiments, a shorter anti-sense probe of about 200 bp not including the homeobox region was prepared after digestion of plasmid with *Bst*NI. The control probe (sense probe), of opposite sense, was synthesized in a T7 RNA polymerase reaction after linearization of plasmid with *Eco*RI. Methods used for the production and alkaline hydrolysis of radioactive probes, for the preparation of embryo sections, and for *in situ* hybridization were all as previously described (Gaunt *et al.* 1986; Gaunt, 1987).

Interpretation of embryo sections in autoradiograms was made with reference to Holland & Hogan (1988). These authors distinguished between 'somites' of the early embryo (such as seen in our 8¼-day embryo sections) and the 'prevertebrae' of later stages (such as seen in our 12½-day sections). This distinction is important since prevertebra 1 does not develop from somite 1, but instead may develop from somites 5 and 6 (Holland & Hogan, 1988).

Results

Molecular cloning and sequence analysis of Hox 6.1

A cDNA library from adult mouse kidney poly(A)$^+$ RNA was constructed in λgt10 and screened with a probe derived from the *Hox-1.5* (Mo-10) gene containing 145 bp of homeobox sequence (McGinnis *et al.* 1984). One of the clones isolated was studied in detail and found to contain an insert of 540 bp with a *Bgl*II site, in the homeobox (Fig. 1). The nucleotide sequence of the insert from the 5′ end to the first inframe termination codon was determined and is shown in Fig. 2.

Nucleotide sequence homology of the *Hox-6.1* homeobox with other mouse homeoboxes is greatest at 86% with *Hox-1.2* (Colberg-Poley *et al.* 1985). However, homology at the amino acid level is greatest with a *Xenopus* homeobox *Xeb1* (previously called *AC1*) and a human homeobox *c8*, being complete at 100% within the 60 amino acids of the homeobox (Carrasco *et al.* 1984; Muller *et al.* 1984; Boncinelli *et al.* 1985). The amino acid sequences of Hox-6.1 and c8 are identical over the regions compared, namely from 10 amino acids upsteam of the homeobox to the termination codon (Simeone *et al.* 1987). Homology with Xeb1 extends both immediately upstream and downstream of the homeobox (Fig. 2). Published sequence of Xeb1 is available for seven amino acids upstream and six amino acids downstream of the homeobox. Six of the seven upstream amino acids are conserved with the nonconserved being alanine in Hox-6.1 (GCG) at position −2 which is serine (TCG) in Xeb1. Five of the six downstream amino acids are also conserved, the exception being threonine (ACG) at position +63 in Hox-6.1 and serine (TCG) in Xeb1 which is a conservative change. No significant homology in these regions is evident with other homeobox genes. Further sequence of Xeb1 has not yet been published but from a personal communication (Eddy De Robertis), it is clear that the downstream amino acid

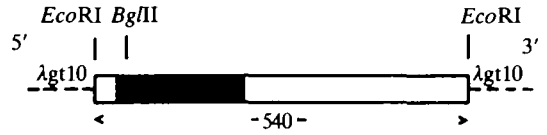

Fig. 1. Restriction map of adult kidney homeobox cDNA clone of *Hox6.1*. Mouse DNA is shown as open boxes with the homeobox region filled in. All *Eco*RI sites were inserted during cloning with synthetic linkers. The whole *Eco*RI insert was subcloned into M13 mp8 for sequencing and the *Eco*RI/*Bgl*II fragment into pGEM4 for production of riboprobes.

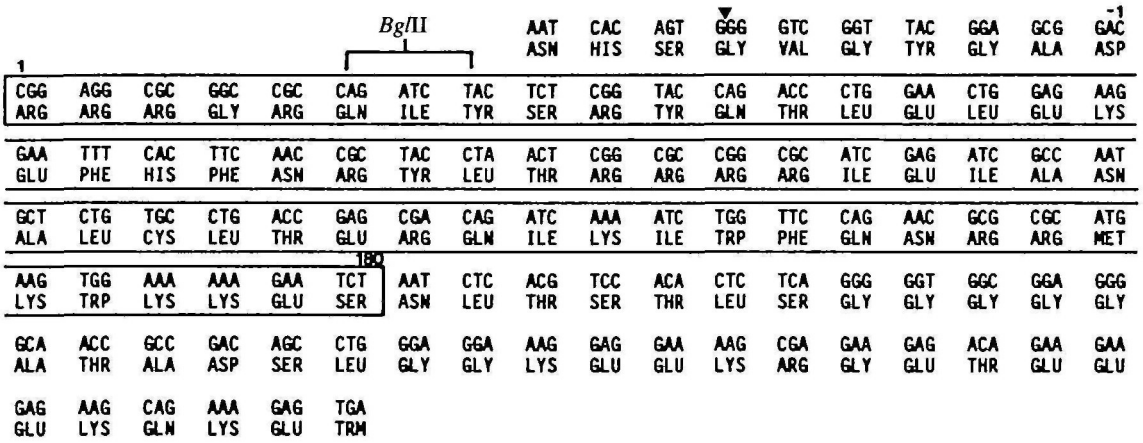

Fig. 2. Nucleotide sequence and conceptual amino acid sequence of the *Hox6.1*, cDNA clone. Boxed region is the homeobox sequence. Nucleotides and amino acids are numbered with respect to the first base of the homeobox. The position of the splice site is arrowed, which is identical to that in *c8*.

sequences of Xeb1 and Hox-6.1 are highly conserved but not identical.

Hox-6.1 has been sequenced as far as the first inframe termination codon which is found 35 amino acids downstream of the homeobox and is followed by at least 220 bp of 3′-untranslated sequence. Long 3′-untranslated sequences are a common feature of many *Drosophila* homeobox genes and have also been found in at least one other mouse gene, *Hox-2.1* (Krumlauf *et al.* 1987).

Several other clones isolated from the kidney cDNA library were found to be derived from transcripts of Hox-6.1 but none contained any more coding sequence than that shown in Fig. 2. Surprisingly, the majority of Hox-6.1 cDNA clones examined from the kidney library were found to be derived from unspliced RNA transcripts of *Hox-6.1* having an intact 3′ splice acceptor consensus sequence at position −22 (Fig. 2) (unpublished) which is the same position as the splice site in *c8* (Simeone *et al.* 1987). The clone sequenced in Fig. 2 is derived from a spliced transcript of *Hox-6.1* with the 5′ limit of the clone ending just 5′ of the position of the 3′ splice site. The possible significance of these unspliced RNA transcripts will be discussed in a later publication.

When a Southern blot of mouse DNA was probed with the insert under conditions of high stringency, a single band was visible with each enzyme used, showing there to be only one copy of the gene (Fig. 3).

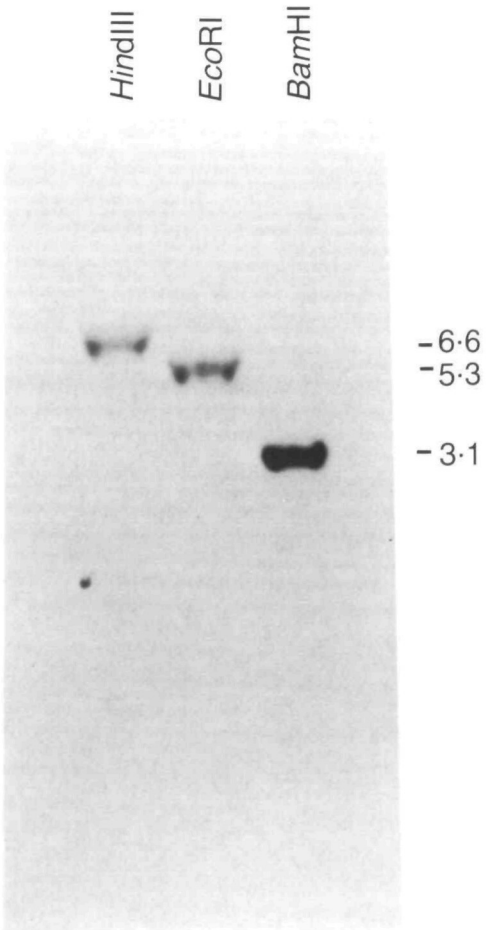

Fig. 3. Genomic Southern blot of restriction digested mouse DNA probed with the 540 bp insert of the cDNA clone, under high-stringency conditions.

Expression of Hox-6.1

(a) *RNase protection assays*

RNase protection assays were used to detect transcripts of the *Hox-6.1* gene during embryonic development and in adult tissues. The 490 bp *Eco*RI/*Bgl*II fragment from the insert was subcloned into pGEM4 and ^{32}P-labelled anti-sense riboprobes used for the protection assays. Protected fragments were detected in RNA extracted from embryos on days 9½ and 13½ of gestation and in RNA from day-9½ extraembryonic

tissue, adult kidney and testis (Fig. 4). Very faint protected fragments were detected in day-7½ embryo and extraembryonic RNA after prolonged exposure of the gels but these were not visible when photographed and were therefore not included in Fig. 4. No protected fragments were detected with RNA from adult liver, brain or spleen (Fig. 4). Control experiments were also conducted with the sense riboprobes and no protected fragments were observed (not shown).

Although protected fragments were occasionally visible at 490 bp, corresponding to the size of the intact riboprobe, these were faint even after prolonged exposure of the gel. The major protected fragment was consistently 265 bp (Fig. 4).

(b) *Northern blot analysis*

Northern blots of total RNA from adult kidney, day-13½ embryos and, as a control, adult liver were probed with the same ^{32}P-labelled anti-sense riboprobe used for the RNase protection assays. After stringent washings and RNase treatment (see Methods), bands were visible in the tracks from kidney and embryo RNA but not in the liver track. One major transcript was visible from kidney and embryo RNA but these were of different sizes. In kidney, the single major transcript detectable was 2·7 kb. With day-13½ embryo RNA, the major transcript was 1·9 kb with a faint band also visible at 2·7 kb (Fig. 5).

(c) In situ *hybridization*

The localization of *Hox-6.1* transcripts in developing embryos was examined by *in situ* hybridization. At 12½ days (Fig. 6), intense labelling was seen in the spinal cord, but no labelling above background was seen in any part of the brain. The boundary between anterior (unlabelled) and posterior (labelled) parts of the nervous system was sharply defined and was located within the spinal cord at a position just posterior to the hindbrain (Fig. 6A). No reduction in intensity of labelling was detected between anterior parts of the spinal cord (Fig. 6A) and more posterior parts

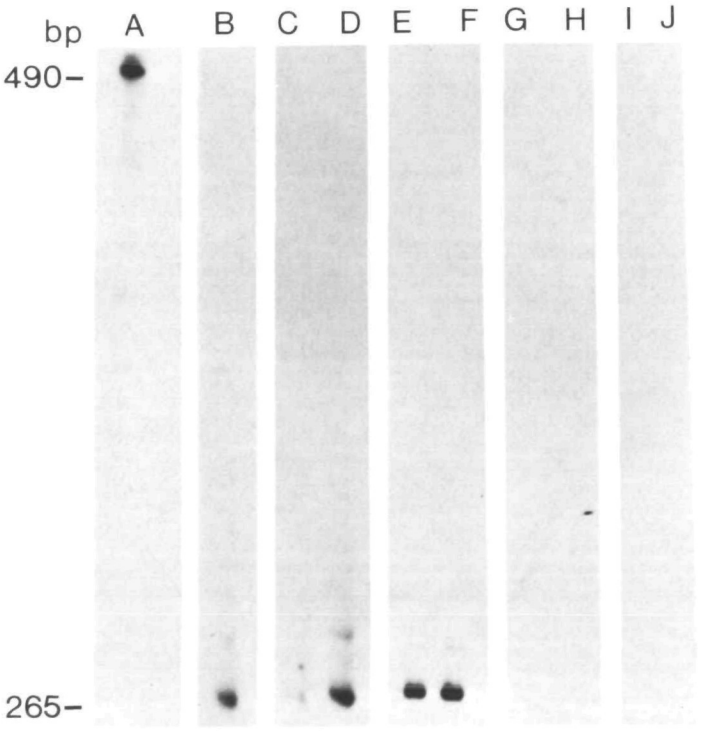

Fig. 4. RNase protection assay of total RNA from adult tissues and embryos with the 490 bp *Eco*RI/*Bgl*II fragment of *Hox6.1*. Protected fragments are visible with adult kidney and testis day-9½, -13½ embryos and day-9½ extraembryonic tissues. The size of the major protected fragment is approximately 265 bp as determined by markers prepared from a sequencing reaction. Faint bands were also visible on the original autoradiographs at 490 bp corresponding to the size of the intact transcript. A, Intact riboprobe; B, testis; C, day-9½ extraembryonic tissue; D, day-9½ embryo; E, day-13½ embryo; F, kidney; G, liver; H, spleen; I, brain; J, RNase-treated riboprobe. Lanes B–D and E–I are from different experiments.

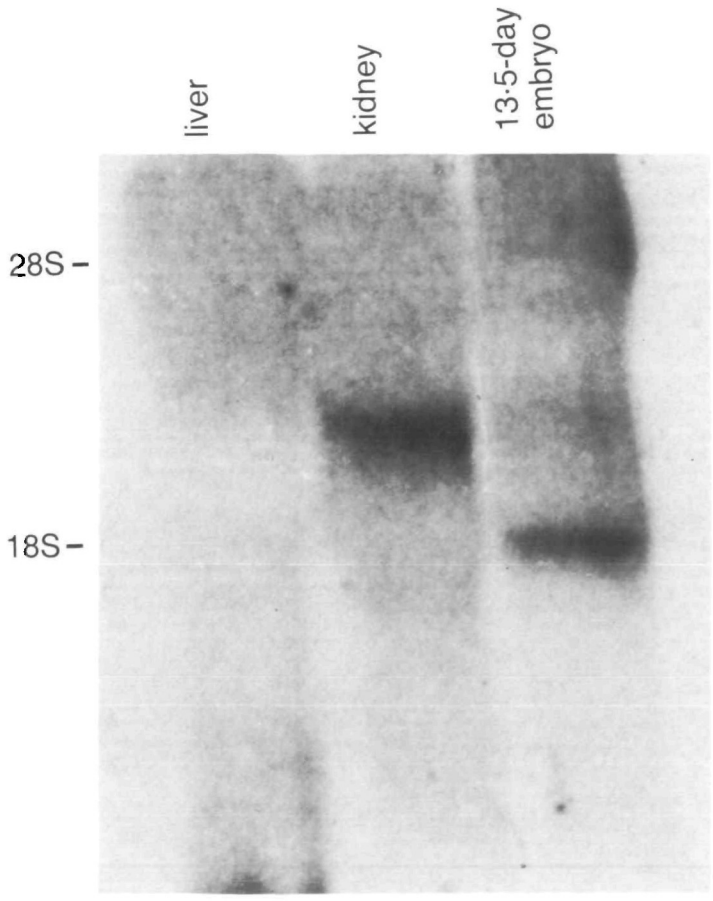

Fig. 5. Northern blot of adult liver, kidney and day-13½ embryo RNA probed with the 490 bp *Eco*RI/*Bgl*II fragment of *Hox6.1*. Molecular weight markers are derived from rehybridization of the blot with rRNA probes which were also confirmed with pBr markers run on the gel (not shown).

(Fig. 6D). Spatial restriction of *Hox-6.1* expression was also evident within the column of prevertebrae (Fig. 6D), but here the boundary between labelled and unlabelled regions was more posteriorly located than in the nervous system. Prevertebrae 1–6 were not labelled above background. Prevertebra 7 was weakly labelled, and labelling was progressively stronger in prevertebrae 8 and 9. Prevertebrae 9–16 were intensely labelled. Labelling was progressively weaker over prevertebrae 17–20 but persisted at a

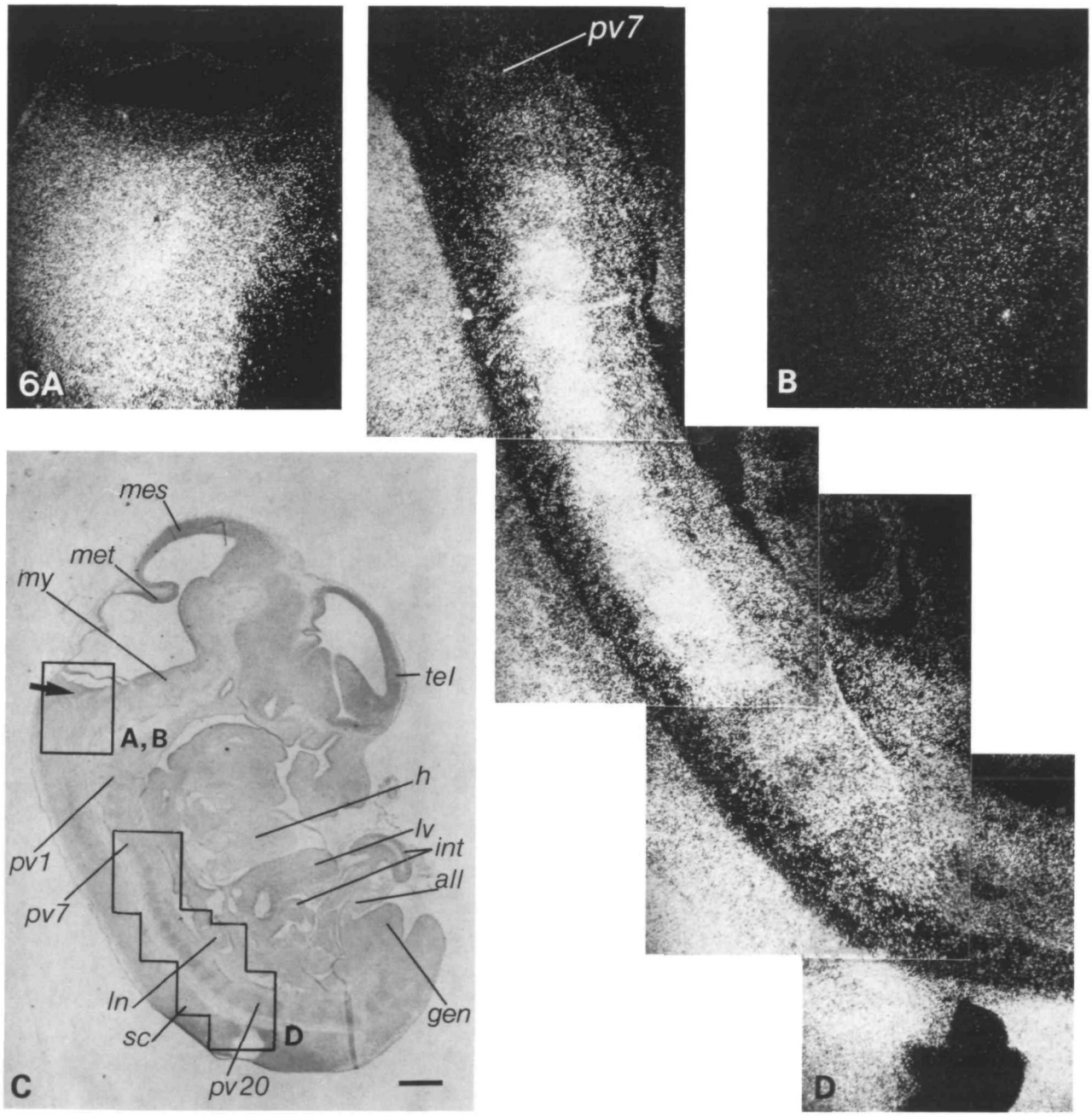

Fig. 6. *In situ* hybridization to probe for *Hox6.1* transcripts in a 12½-day embryo section.

Sections are parasagittal and areas shown under dark-field illumination (A,B,D) are outlined on the corresponding bright-field view (C). Anti-sense probe was used in A,D; sense (control) probe was used in B. *tel*, telencephalon; *mes*, mesencephalon; *met*, metencephalon (cerebellum); *my*, melencephalon (medulla oblongata); *sc*, spinal cord; *pv1*, *pv7*, *pv20*, prevertebrae 1, 7 and 20; *ln*, lung; *h*, heart; *lv*, liver; *int*, intestine; *all*, allantoic stem; *gen*, genital eminence. Bar, 0·5 mm.

The boundary between anterior, unlabelled and posterior, labelled nervous tissue (arrowed in C) was located just posterior to the junction of the brain and spinal cord. No specific labelling or boundary was seen in an adjacent section of the same 12½-day embryo hybridized to the control probe (B). Only prevertebrae posterior to pv7 were labelled.

low, apparently uniform level over all more posterior prevertebrae. In addition, labelling was also seen in several other mesodermal derivatives in the posterior part of the body. Most obvious was labelling in the metanephric kidney, the gonad and the mesodermal components (but not the endodermally derived lining epithelia) of the lung (Fig. 6D), the stomach and some regions of the intestine. No specific labelling was detected over the heart or liver (not shown). *In situ* hybridization experiments using the sense (control) probe produced only background labelling of tissues with no evidence of boundaries in either the nervous system (Fig. 6B) or the prevertebral column (not shown). *In situ* hybridization using an anti-sense probe that did not include the homeobox (see Materials and methods) produced an identical pattern of labelling to that described for the full-length probe (not shown). This finding, together with the results of the Northern blot analysis, suggests strongly that the Hox-6.1 probe used for *in situ* hybridization did not cross hybridize with the transcripts of other homeobox genes.

At 8¼ days gestation, the embryo was at the 6-somite stage (see Materials and methods for distinction between somites and prevertebrae). *In situ* hybridization at this stage (Fig. 7) showed that *Hox-6.1* transcripts were restricted to the ectoderm and mesoderm layers in the posterior part of the embryo. Labelling did not include any of the somites already formed but was seen in the presomitic mesoderm posterior to somite 6. Labelling in the mesoderm layer extended posteriorly into the allantois, but other extraembryonic tissues including decidual tissue and extraembryonic membranes were not labelled above background (Fig. 7). Experiments using sense (control) probe on the 8¼-day embryo showed no specific labelling of tissues (not shown).

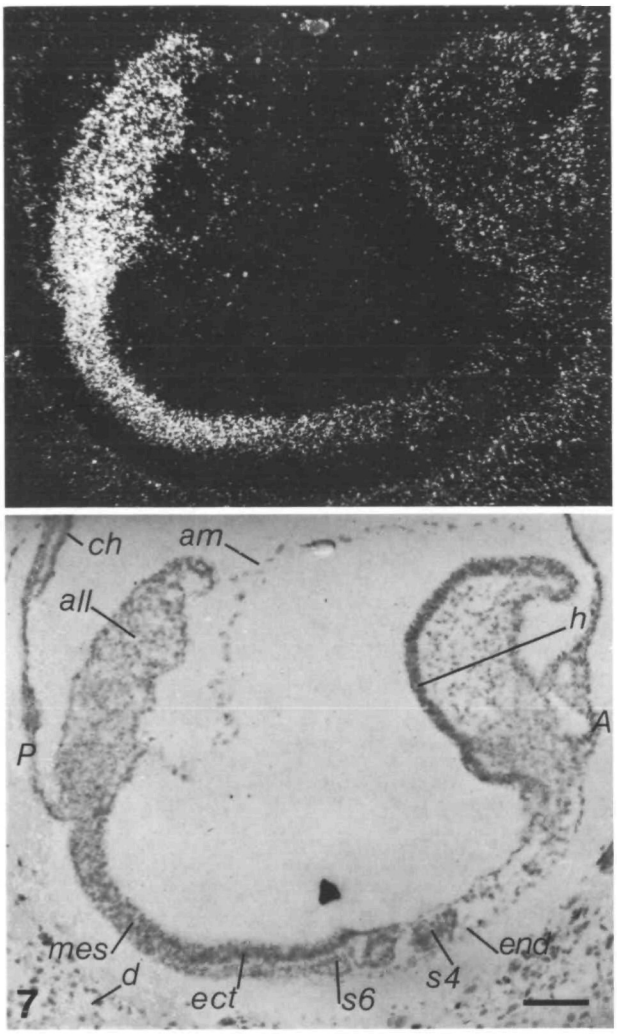

Fig. 7. Expression of *Hox 6.1* in the 6-somite embryo. Upper panel, dark-field illumination; lower panel, phase contrast. Section is parasagittal. *A*, anterior; *P*, posterior; *ect*, ectoderm; *mes*, mesoderm; *end*, endoderm; *hf*, head fold; *S4*, *S6*, somites 4 and 6; *all*, allantois; *am*, amnion; *ch*, chorion; *d*, decidual tissue. Bar, 0·1 mm. *Hox 6.1* transcripts were restricted to the ectoderm and mesoderm tissue in the posterior region of the embryo. Labelling within mesoderm began posterior to the newly formed somite 6 and extended caudally into the allantois. Other extraembryonic tissues were not labelled above background.

Chromosomal location of Hox-6.1

The chromosomal location of *Hox-6.1* was determined by *in situ* hybridization with ^{3}H-labelled probe from the 540 bp insert. A total of 133 grains were scored in a sample of 100 mitotic cells. If the grains were randomly distributed and not a positive indication of a successful hybridization, they would be expected to be distributed at random according to unit length of chromosome. Therefore, estimates were made for the number of grains expected to overlie each homologous pair by reference to the total number of grains scored in relation to the chromosome lengths given by Nesbitt & Franke (1973). Apart from chromosome 14, expected numbers were observed to overlie the other chromosomes. Chromosome 14 represents 4·31 % of the haploid genome and would be expected to carry approximately 6 grains (4·31 % of 133 = 5·7); however, 23 grains were observed, a fourfold increase of observed over expected. Further, of the 23 grains, 20 overlay the distal segment of the chromosome with a peak distribution over the middle of band 14E2 (Fig. 8). Clearly, this hybridization indicates that *Hox-6.1* maps to this region of chromosome 14.

Isolation of a genomic clone of Hox 6.1

In order to begin to map the Hox 6 locus and determine whether other homeobox-containing genes

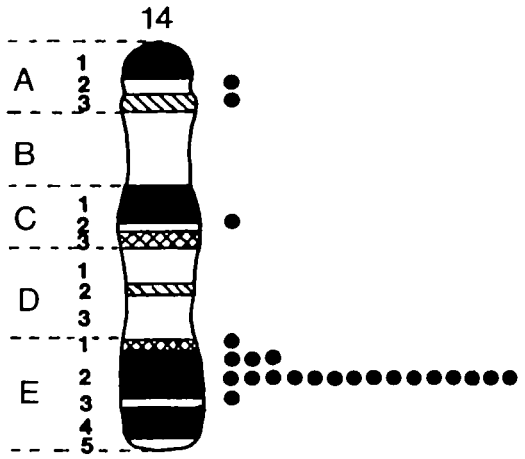

Fig. 8. Grain distribution along chromosome 14 after *in situ* hybridization with a *Hox6.1* probe.

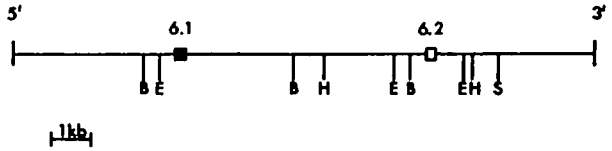

Fig. 9. Map of a genomic clone of Hox6. The putative homeobox sequence *Hox6.2* has been detected by Southern blotting and has not yet been confirmed by sequencing. E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; S, *Sal*I.

might be located close to *Hox-6.1*, a genomic clone containing *Hox-6.1* was isolated from a library in Charon 35. The probe used was derived from a region of a cDNA clone of *Hox-6.1* that contained no homeobox sequence and was shown by low-stringency probing of Southern blots not to cross hybridize with any other mouse sequences. The genomic library was probed under high-stringency conditions and five positive plaques were obtained. One of these five phage has been mapped, with the position of *Hox-6.1* being detected by hybridization of digested fragments to the probe used to screen the library. The same blots were reprobed with the probe derived from *Hox-1.5* which identified a possible second homeobox sequence within the clone approximately 7 kb downstream of *Hox-6.1* and this putative homeobox-containing gene has been named *Hox-6.2* (Fig. 9).

Discussion

Homeobox-containing genes have been isolated from a variety of vertebrate species including mammals. The conservation of the homeobox region of these genes and between their *Drosophila* counterparts is striking, suggesting an important role for these genes in development. It also appears that regions outside the homeobox are conserved between different genes from different vertebrate species (Boncinelli *et al.* 1985). Furthermore, there are now at least three instances of possible homologous homeobox genes where large portions of the amino acid sequences (if not all), are highly conserved in genes isolated from widely different vertebrate species. The mouse *Hox-2.1* gene (Jackson *et al.* 1985) has homologues in the human (*Hu1*) (Hauser *et al.* 1985) and in *Xenopus* (*Xhox1B*) (Harvey *et al.* 1986). Similarly there is high amino acid homology between the mouse *Hox-1.4* (Duboule *et al.* 1986) *Xenopus Xhox1A* (Harvey *et al.* 1986) and human *c13* (Boncinelli *et al.* 1985) genes. Here, we report the isolation of a new mouse homeobox gene which appears to be homologous to the *Xenopus Xeb1* and human *c8* genes (Carrasco *et al.* 1984; Boncinelli *et al.* 1985). If these highly conserved genes have an important function in development it might be expected that homologous genes would serve similar functions in their respective species.

The *in situ* hybridization results now obtained for *Hox-6.1* can be compared with results already published for several other mouse homeobox genes (Awgulewitsch *et al.* 1986; Gaunt *et al.* 1986; Gaunt, 1987; Utset *et al.* 1987; Krumlauf *et al.* 1987). In 12½-day mouse embryos, all of the homeobox genes studied show transcripts in the posterior, but not anterior, regions of the nervous system. Individual genes may differ, however, in the position of the boundary between labelled and unlabelled parts. Thus, the boundary for *Hox-6.1*, located just behind the hindbrain, is clearly posterior to the boundary for *Hox 1.5* (Gaunt *et al.* 1986; Gaunt, 1987), but anterior to that of *Hox-3.1* (Awgulewitsch *et al.* 1986; Utset *et al.* 1987). The *Hox-6.1* transcript boundary within somitic mesoderm derivatives is located posterior to the boundary in nervous tissue and is at the position of the 7th prevertebra. In contrast, all prevertebrae are labelled with *Hox 1.5* (Gaunt *et al.* 1986; Gaunt, 1987). Like *Hox-6.1*, *Hox-3.1* (Utset *et al.* 1987) shows a transcription boundary within somitic mesoderm derivatives posterior to that in the nervous system. The location of *Hox-6.1* transcripts now described within the developing lung seems identical to that previously noted for *Hox-2.1* (Krumlauf *et al.* 1987).

The pattern of *Hox-6.1* labelling observed in the 8¼-day embryo shows similarities to that already described for *Hox-1.5* (Gaunt *et al.* 1986; Gaunt, 1987). Thus, labelling is restricted to posterior ectoderm and mesoderm tissues, and extends posteriorly into the allantois. *Hox-6.1* transcripts at this stage do not, however, extend as far forward in the embryo as *Hox-1.5* transcripts, which are detected in anterior somites (Gaunt, 1987). The observations for *Hox-6.1* are consistent with a view, as suggested for *Hox 1.5* (Gaunt, 1987), that this gene serves as one of a series

of homeobox genes whose expression provides positional cues during the determination of tissues along the body axis.

Since *in situ* hybridization data are already published for *Xebl* (Carrasco & Malacinski, 1987), the *Xenopus* homologue of mouse *Hox-6.1*, it is now possible for the first time to compare the spatial distribution of homologous homeobox genes during the development of two different vertebrate species. The transcription boundaries within the central nervous system are similar in both species, being located at the junction of the spinal cord and the brain. Also similar is the concentration of transcripts in the posteriormost parts of earlier-stage embryos (gastrula- and neurula-stage *Xenopus* embryos, and $8\frac{1}{4}$-day late-primitive-streak-stage mouse embryos). The striking pattern of *Hox-6.1* transcription seen in the somitic mesoderm derivatives of the mouse does not seem to be repeated in *Xenopus*. The somitic mesoderm of *Xenopus* showed no labelling by the *Xebl* probe, although labelling was noted in the lateral mesoderm around the blastopore of gastrula stages (Carrasco & Malacinski, 1987). It is possible that homeobox gene transcription in the mesoderm of *Xenopus* is more transient than in mouse so that it does not persist at stages of somite formation.

The full coding sequence of *Hox-6.1* has not yet been determined since, in the cDNA clones so far studied, there appears to be a preponderance of clones derived from unspliced transcripts of *Hox-6.1* where the 3′ splice acceptor consensus is intact (unpublished data). In view of the complete amino acid homology between Hox-6.1 and c8 in the regions compared (Fig. 2) it seems probable that the upstream region of Hox-6.1 beyond that currently sequenced will be identical to c8 (Simeone *et al.* 1987). We are currently sequencing a genomic clone of *Hox-6.1* to confirm this.

The position of the 3′ splice acceptor consensus sequence of *Hox-6.1* is identical to that of the *c8* gene (Simeone *et al.* 1987). The clones isolated that do contain spliced transcripts of *Hox-6.1* terminate just upstream of the position of the 3′ splice site (Fig. 2). It seems probable from the Northern blot data that this preponderance of unspliced *Hox-6.1* transcripts may be a feature of the adult kidney and may not occur to the same extent in embryos. The predominant message size in embryos is 1·9 kb which is consistent with the size of other homeobox gene transcripts (e.g. Gaunt *et al.* 1986; Krumlauf *et al.* 1987). The major transcript size in the kidney is significantly larger at 2·7 kb and this may reflect transcripts containing unspliced intron sequences. A faint band at 2·7 kb in the embryo RNA is visible which may suggest the occurrence of a small proportion of similar unspliced transcripts in embryos.

The RNase protection data show that *Hox-6.1* is expressed in adult kidney and testis but not in liver, spleen or brain. The transcripts are also detected in day-$9\frac{1}{2}$ embryonic and extraembryonic tissues and day-$13\frac{1}{2}$ embryos. The size of the protected fragments seen is around 265 bp which is 225 bp shorter than the size of the intact probe used. We do not, as yet, have an explanation for this but a fragment of 265 bp from the 5′ end would end about 30 bp past the termination codon and it is possible that this may reflect processing of 3′ untranslated sequences.

Transcripts were detected in extraembryonic tissues (allantois and amnion) from day-$9\frac{1}{2}$ embryos. Previously examined extraembryonic tissues, placenta (Jackson *et al.* 1985) and yolk sac plus placenta (Colberg-Poley *et al.* 1985) showed no evidence of homeobox gene transcription. These findings for *Hox-6.1* are readily explained by the *in situ* hybridization data, which show transcripts in the allantois but not in other extraembryonic tissues of the $8\frac{1}{4}$-day embryo. Similar findings have been obtained for *Hox-1.5* transcripts (Gaunt, 1987). We have examined the expression of *Hox-6.1* in day-$7\frac{1}{2}$ embryos and extraembryonic tissues and found barely detectable levels of expression using RNase protection (data not shown).

The Hox-6 locus has been located on chromosome 14 in the E2 region. This location does not appear to contain any known obvious developmental mutations. In common with the Hox-1 and Hox-2 loci (Hart *et al.* 1985; Duboule *et al.* 1986), Hox-6 appears to contain a cluster of homeobox-containing genes. A single genomic clone of 15 kb contains *Hox-6.1* and another putative homeobox (*Hox-6.2*) detected by Southern blotting, within 7 kb. We are currently examining other genomic clones to determine the extent of this cluster.

Between amino acids 68 and 72 of Hox 6.1 are five consecutive glycine residues. Such a region may be devoid of secondary structure and could form a short flexible hinge-type region in the protein linking two spatially distinct domains. A similar putative hinge-type region has been identified in the *Ubx* homeobox gene in *Drosophila* but no such regions have been identified downstream in the homeobox genes so far described (Beachy *et al.* 1985). We have recently isolated a pig homeobox gene, however, which has a glycine-rich region between amino acids 70–75 downstream of the homeobox.

We wish to thank Kenneth Krauter for the gift of the mouse genomic library, Eddy De Robertis for the unpublished cDNA sequence of *Xebl*, Andres Carrasco for detailed information on the localization of *Xebl* transcripts and Don Powell for invaluable help and assistance.

References

AWGULEWITSCH, A., UTSET, M. F., HART, C. P., MCGINNIS, W. & RUDDLE, F. H. (1986). Spatial restriction in expression of a mouse homoeobox locus within the central nervous system. *Nature, Lond.* **320**, 328–335.

AVIV, H. & LEDER, P. (1972). Purification of biological active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412.

BEACHY, P. A., HELFAND, S. L. & HOGNESS, D. S. (1985). Segmental distribution of bithorax complex proteins during *Drosophila* development. *Nature, Lond.* **313**, 545–551.

BONCINELLI, E., SIMEONE, A., LA VOLPE, A., FAIELLA, A., FIDANZA, V., ACAMPORA, D. & SCOTTO, L. (1985). Human cDNA clones containing homeobox sequences. *Cold Spring Harb. Symp. quant. Biol.* **50**, 301–306.

CARRASCO, A. E. & MALACINSKI, G. M. (1987). Localization of *Xenopus* homeobox gene transcripts during embryogenesis and in the adult nervous system. *Devl Biol.* **121**, 69–81.

CARRASCO, A. E., MCGINNIS, W., GEHRING, W. J. & DE ROBERTIS, E. M. (1984). Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. *Cell* **37**, 409–414.

CHIRGWIN, J. M., PRYZBYLA, A. E., MACDONALD, R. J. & RUTTER, W. J. (1979). Isolation of biological active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299.

COLBERG-POLEY, A. M., VOSS, S. D., CHOWDHURY, K., STEWART, C. L., WAGNER, E. F. & GRUSS, P. (1985). Clustered homeoboxes are differentially expressed during murine development. *Cell* **43**, 39–45.

DOLECKI, G. J., WANNAKRAIROJ, S., LUM, R., WANG, G., RILEY, H. D., CARLOS, R., WANG, A. & HUMPHREYS, T. (1986). Stage-specific expression of a homeobox-containing gene in the non-segmented sea urchin embryo. *EMBO J.* **5**, 925–930.

DUBOULE, D., BARON, A., MAHL, P. & GALLIOT, B. (1986). A new homeobox is present in overlapping cosmid clones which define the mouse Hox-1 locus. *EMBO J.* **5**, 1973–1980.

GAUNT, S. J. (1987). Homeobox gene *Hox-1.5* expression in mouse embryos: earliest detection by *in situ* hybridization is during gastrulation. *Development* **101**, 51–61.

GAUNT, S. J., MILLER, J. R., POWELL, D. J. & DUBOULE, D. (1986). Homeobox gene expression in mouse embryos varies with position by the primitive streak stage. *Nature, Lond.* **324**, 662–664.

GEHRING, W. J. (1985). The homeobox: A key to the understanding of development. *Cell* **40**, 3–5.

HART, C. P., AWGULEWITSCH, A., FAINSOD, A., MCGINNIS, W. & RUDDLE, F. H. (1985). Homeobox gene complex on mouse chromosome 11: Molecular cloning, expression in embryogenesis and homology to a human homeobox locus. *Cell* **43**, 9–18.

HARVEY, R. P., TABIN, C. J. & MELTON, D. A. (1986). Embryonic expression and nuclear localisation of *Xenopus* homeobox (Xhox) gene products. *EMBO J.* 5, 1237–1244.

HAUSER, C. A., JOYNER, A. L., KLEIN, R. D., LEARNED, T. K., MARTIN, G. R. & TIJAN, R. (1985). Expression of homologous homeobox containing genes in differentiated human teratocarcinoma cells and mouse embryos. *Cell* **43**, 19–28.

HOLLAND, P. W. H. & HOGAN, B. L. M. (1988). Spatially restricted patterns of expression of the homeobox-containing gene *Hox 2.1* during mouse embryogenesis. *Development* **102**, 159–174

HONG, G. F. (1981). A method for sequencing single-stranded cloned DNA in both directions. *Bioscience Reports* **1**, 243–252.

JACKSON, I. J., SCHOFIELD, P. & HOGAN, B. (1985). A mouse homoeobox gene is expressed during embryogenesis and in adult kidney. *Nature, Lond.* **317**, 745–748.

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

LYON, M. F., ZENTHON, J., EVANS, E. P., BURTENSHAW, M. D., WAREHAM, K. A. & WILLIAMS, E. D. (1986). Lack of inactivation of mouse X-linked gene physically separated from the inactivation centre. *J. Embryol. exp. Morph.* **97**, 75–85.

MCGINNIS, W., HART, C. P., GEHRING, W. J. & RUDDLE, F. H. (1984). 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., KURIOWA, A. & GEHRING, W. J. (1984). A conserved DNA sequence found in homeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature, Lond.* **308**, 428–433.

MULLER, M. M., CARRASCO, A. E. & DE ROBERTIS, E. M. (1984). A homeobox containing gene expressed during oogenesis in *Xenopus*. *Cell* **39**, 157–162.

NESBITT, M. N. & FRANCKE, U. (1973). A system of nomenclature for band patterns of mouse chromosomes. *Chromosoma* **41**, 145–158.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467.

SHEPHERD, J. C. W., MCGINNIS, W., CARRASCO, A. E., DE ROBERTIS, E. M. & GEHRING, W. J. (1984). Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. *Nature, Lond.* **310**, 70–71.

SIMEONE, A., MAVILIO, F., ACAMPORA, D., GIAMPAOLA, A., FAIELLA, A., ZAPPAVIGNA, V., D'ESPOSITO, A. D., PANNESE, M., RUSSO, G., BONCINELLI, E. & PESCHLE, C. (1987). Two human homeobox genes c1 and c8: structure analysis and expression in embryonic development. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4914–4918.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. molec. Biol.* **98**, 503–517.

UTSET, M. F., AWGULEWITSCH, A., RUDDLE, R. H. & MCGINNIS, W. (1987). Region-specific expression of two mouse homoeobox genes. *Science* **235**, 1379–1382.

WATSON, C. J. & JACKSON, J. F. (1985). An alternative procedure for the synthesis of double-stranded cDNA for cloning in phage and plasmid vectors. *DNA Cloning* **1**, 79–88. (ed. D. M. Glover). IRL Press.

ZABEL, B. U., NAYLOR, S. L., SAKAGUCHI, A. Y., BELL, G. I. & SHOWS, T. B. (1983). High-resolution chromosomal localisation of human genes for amylase, proopiomelanocortin, somatostatin, and a DNA fragment (D351) by in situ hybridisation. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6932–6936.

ZINN, K., DI MAIO, D. & MANIATIS, T. (1983). Identification of two distinct regulatory regions adjacent to the human β-interferon gene. *Cell* **34**, 865–879.

(*Accepted 22 October 1987*)