Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility

Hsiu Mei Hsieh-Li[1], David P. Witte[2], Michael Weinstein[1], William Branford[1], Hung Li[1], Kersten Small[1] and S. Steven Potter[1],*

Divisions of [1]Basic Science Research and [2]Pathology, Children's Hospital Research Foundation, and Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45229-3039, USA

*Author for correspondence

SUMMARY

Hoxa 11 is a murine *Abdominal-B*-type homeobox gene. The structure of this gene is presented, including genomic and cDNA sequence. The cDNA includes the complete open reading frame and based on primer extension results is near full length. Surprisingly, the antisense strand of *Hoxa 11* was found to be transcribed. Moreover, these antisense transcripts were processed and polyadenylated. The developmental expression patterns for both sense and antisense transcripts were examined using serial section and whole-mount in situ hybridizations. *Hoxa 11* transcription patterns were defined in the limbs, kidney and stromal cells surrounding the Mullerian and Wolffian ducts. Of particular interest, in the developing limbs, the sense and antisense transcripts showed complementary expression patterns, with antisense RNAs increasing in abundance in regions where sense RNAs were diminishing in abundance. Furthermore, targeted mutation of *Hoxa 11* is shown to result in both male and female sterility. The female mutants produce normal ova, which develop properly post-fertilization when transferred to wild-type surrogate mothers. The *Hoxa 11* homozygous mutants are shown to provide a defective uterine environment. The mutant males exhibited a malformation of the vas deferens that resembles a partial homeotic transformation to an epididymis. In addition, the mutant testes fail to descend properly into the scrotum and, likely as a result, spermatogenesis is perturbed.

Key words: homeobox gene, antisense transcripts, limb development, spermatogenesis, uterus, mouse

INTRODUCTION

The 38 mammalian Hox genes are organized in four clusters (McGinnis and Krumlauf, 1992). These *Antennapedia*-like homeobox genes encode transcription factors with a helix-turn-helix motif. The homeobox genes are remarkably well-conserved during evolution, with mammalian and *Drosophila* orthologous homeodomains often showing greater than 95% amino acid sequence identity. These genes are also surprisingly conserved in function; transgenic *Drosophila* that misexpress mammalian Hox genes display homeotic transformations that closely reproduce the results of misexpression of endogenous *Drosophila* homeobox genes (Malicki et al., 1990). This suggests that significant portions of the genetic circuitry of development are conserved between mammals and flies. Furthermore, both mammalian and *Drosophila* homeobox genes show the property of colinearity, with genes located more 3′ within clusters having developmentally earlier and more anterior expression domains.

The *Abdominal-B* (*Abd-B*) type Hox genes are located at the 5′ ends of the mammalian Hox clusters, and are of particular interest for several reasons. These genes have undergone a considerable expansion in number in mammals, with but a single *Abd-B* gene in *Drosophila*, and a total of 15 *Abd-B*-type genes in mammals. These genes show overlapping domains of expression in developing limbs, suggesting that they may play an important role in specifying regional identity (Dolle et al., 1989; Yokouchi et al., 1991). Targeted mutations in three of these genes, *Hoxa 11*, *Hoxd 11* and *Hoxd 13*, have in each case resulted in both axial and appendicular skeletal defects (Small and Potter, 1993; Davis and Capecchi, 1994; Dolle et al., 1993). The *Hoxa 11* mutants showed posteriorization of the thirteenth thoracic vertebra to form a first lumbar vertebra and anteriorization of the sacral region by one segment. In addition, the ulna and radius of the forelimb, and tibia and fibula of the hindlimb, were dramatically malformed, with extensively altered contour shapes. Moreover, there were apparent gene dosage effects, with hemizygotes showing phenotypes of intermediate severity.

In this report, we extend our studies of the *Hoxa 11* gene. The structure of *Hoxa 11* is presented, including the complete cDNA sequence and 9 kb of genomic DNA sequence. Surprisingly, during the course of this work, it was found that the antisense strand of the *Hoxa 11* gene gives rise to elaborately spliced mRNAs. We also provide a more detailed presentation of the *Hoxa 11* expression pattern during development. *Hoxa*

11 expression in the early limb has previously been studied by serial section (Haack and Gruss, 1993) and whole-mount (Small and Potter, 1993) in situ hybridizations. Here we use both techniques to examine expression of both *Hoxa 11* sense and antisense transcripts in the limbs, including later time points, and in the body of the embryo. The sense and antisense transcripts show a complementarity in the developing limbs, with antisense RNAs appearing in regions where sense RNAs are diminishing, that suggests a possible antisense regulatory role. Furthermore, as the *Hoxa 11* homozygous mutants matured, we observed that both males and females were sterile. In this report, we describe the nature of these infertility phenotypes.

MATERIALS AND METHODS

Cloning of cDNAs

A mouse embryo E10.5 cDNA library (Novagen) was screened with a labeled 300 bp fragment (probe A shown in Fig. 3) which was created by PCR from the region immediately 3′ of homeobox. The most 5′ end 250 bp fragment (nucleotides 4830 to 5080 shown in Fig. 1) of the *Hoxa 11* cDNA from this screen was then used to screen an E12 limb library (kindly provided by M. Solursh) to extend the 5′ end of cDNA. One of the antisense transcripts, 3C, was obtained from this screen; another antisense transcript, 23A, was obtained by using probe B (Fig. 3) to rescreen the limb library; other antisense transcripts, 59 and 40, were obtained from screening an E12.5 embryo library (Novagen) by using the same probe. Plaque hybridization was performed as previously described (Kern et al., 1992).

DNA sequencing

Genomic DNA was sequenced using the USB sequenase system. Both strands were determined and for difficult regions both inosine and deaza nucleotides were used. *Hoxa 11* cDNA and antisense transcripts were sequenced using the ds cycle-sequencing system (GIBCO BRL) according to supplier's protocols.

Primer extension

The procedure for primer extension was modified from Sambrook et al. (1989). 5′-end labeled primer (5×10^5 cts/minute) was hybridized with 10 µg total RNA at 95°C for 10 minutes and cooled to 70°C. The primer-RNA hybrid templates were chilled in ice for 1 minute and then precipitated by ethanol. Reverse transcription was performed at 42°C for 30 minutes by incubating the resuspended templates with 10 units/µl of Superscript II reverse transcriptase (GIBCO BRL). RNA was removed by digestion with 1 unit/µl of RNase H at 55°C for 5 minutes followed by precipitation with ethanol. The products were electrophoresed on a 6% polyacrylamide-8 M urea gel.

Northern blot analysis

Poly(A)$^+$ RNA was isolated from E10.5, 11.5, 12.5 and 14.5 mouse embryos or E11.5 limbs by using either oligo(dT) cellulose (Celano et al., 1993) or the Qiagen d(T) kit according to supplier's protocols. Probes A and B (Fig. 3) were random primer labeled and used to detect *Hoxa 11* and antisense transcripts separately. Probe A is a PCR product of 300 bp immediately 3′ of the *Hoxa 11* homeobox and probe B is the first 650 bp fragment of antisense transcript 3C. Both fragments were subcloned into pBS (KS, Stratagene) to make both sense and antisense riboprobes.

Serial section and whole-mount in situ hybridizations

The procedures for serial section and whole-mount in situ hybridizations were performed as previously described (Li et al., 1994; Small and Potter, 1993). Both strands of probes A and B (Fig. 3) were used to detect *Hoxa 11* sense and antisense transcript expression in different stage mouse embryos or adult tissues.

Zygote transfer

3.5- to 4.5-week-old *Hoxa 11* mutant female mice were superovulated and E0.5 zygotes were transferred into the oviducts of pseudopregnant wild-type mice as described (Hogan et al., 1986).

RESULTS

Hoxa 11 cDNA and genomic structure

The *Hoxa 11* gene was originally cloned using a degenerate oligonucleotide screen of mouse genomic DNA (Singh et al., 1991). To define the coding potential of this gene, a genomic DNA fragment of 300 bp located immediately 3′ of the homeobox was used to screen an embryonic day 10.5 cDNA library. The most 5′ fragment of the cDNAs isolated was used to screen an E12 limb-specific cDNA library to extend the 5′ end. The longest *Hoxa 11* cDNA isolated was completely sequenced on both strands. To determine exon-intron organization and define potential regulatory regions, 9 kb of genomic DNA including the *Hoxa 11* gene was also sequenced on both strands. The *Hoxa 11* cDNA sequence is shown in Fig. 1 in bold letters within the genomic sequence. The homeobox is underlined and the predicted amino acid sequence is shown below.

The *Hoxa 11* cDNA includes the entire open reading frame, with an in frame termination codon just 33 bp 5′ of the translation initiation ATG. The *Hoxa 11* amino acid sequence shows an interesting series of mono-amino acid runs prior to the homeodomain. Most dramatic is a tract of 14 amino acids that includes 11 alanines (A). Such polyalanine tracts have been previously observed in several genes of developmental interest, including the *even-skipped* (Poole et al., 1985), *engrailed* (Macdonald et al., 1986), *cut* (Blochinger et al., 1988), *runt* (Kania et al., 1990) and *ovo* (Mevel-Ninio et al., 1991) genes in *Drosophila*. It is thought that this domain may form an alpha helix structure involved in transcription repression (Han and Manley, 1993a,b).
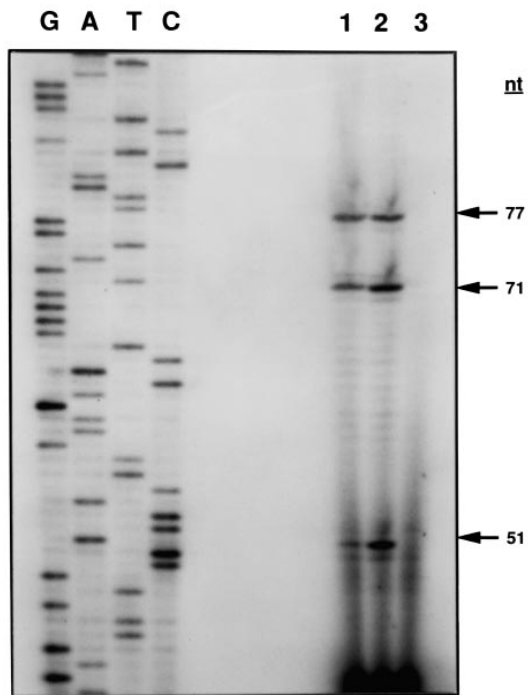
Primer extension analysis was used to define the 5′ transcription start site. The results, shown in Fig. 2, suggest start sites at nucleotides 4641, 4647 and 4667 of Fig. 1.

Antisense transcripts

A restriction segment probe corresponding to the 5′ end of the *Hoxa 11* cDNA was used to rescreen cDNA libraries in order to obtain a full-length cDNA. Surprisingly, this screen resulted in several fold more positive clones than had been previously found using a probe corresponding to the 3′ end of the *Hoxa 11* cDNA. Moreover, when these cDNAs were characterized, it was found that most represented transcripts from the *Hoxa 11* antisense strand of DNA. For cDNAs from the poly(dT)-primed E12.5 cDNA library, the positions of the poly(A) tails served to orient them. In addition, sequence analysis of all four cDNAs shown in Fig. 3, and comparison with genomic DNA, allowed a definition of intron-exon boundaries. The positions of the splice donor and acceptor sites served to confirm that these cDNA transcripts were from the *Hoxa 11* antisense strand. All splice junctions were conventional, with good similarity to established consensus sequences.

```
CTGGTGAAGACCTCTTTCTTTAAACCTGTGTAGTAGTTTCCTGCCAGGGA ACACATCAGCTCTTTCTGAGTCAAGGAGAGACATCAGAAAAGGCAGGAGA 100
AAGCCTAAGAAAGGATAGATGGGAAAAGTCAAGATGGGAGACAAATGGGG CAGGAGGAACAGCCTGGGAACAAAACAGGCTTTGAGGAGGATCTGGAGAG 200
TGGTTGGGATATTCACGCTATGCTGATCCTTCCAGGAACTGCACTGGGTAT TTCCAATGGATGCCCAGGAGTTAGAGGAGTTAGAGATAGCTGGCCTCCAA 300
CCACCCCCTTTGACAGATGGTGGTTGATGGAAGGAGTTGATGCTAGGAAG CCCAGAGTCAAGAAGTGCTTATGGTTGCCCATGTGACTCTTTCTAGTCTC 400
CCTTGAAGTGGCATAATGCTAAGGAGCCTAGCAACTGATAGCGGTCCATC TCAAGGCTCTTGAATGGCTGTGTTGGAACCTGGCCTGGATCAGACAATCT 500
GGCCACAAGGCCTTCCATTGCCTTTCTACAAATACACGATTCCTCAAAAG CTATAATCACTCTAGGCAGCTTTCCCCGAGGCTCCCAACTACTGCTCAT 600
ACATCTCCGACTGCAAGGAAAATCAAATTCAGCAGAGGCACCTGAATACA CTTTATTCTGGTATTCTATCTCAGAAGGAGAGTGCAGCCAGCGCAAGGTG 700
CCTGAATCTCTCTCCCTCCCTCCCTTCTTGGTACTTTTAGCCGGCGGGA GGAAAACAGCAGCCACCCCAGCCGCTGTGGACACTAGCATATACCCTTTA 800
GGACCAAATCAAAAGAAACCTCCGAAAAGGGTTGGTTTGTTCTGCCGATC AAATCTCCTTTCAGCCCGACCGAGGCTTAAATCGTTGGCCACCTCCAAAT 900
GCTGCGCCGCTGGGGGTGTGGGGGGCGAGGTCTCGGCAAGGCCCTAAGGGGG TCTGCCTAGACGTTTGAAGTTCGGGAAAACGAGGGGGGAGGTTGAACCGGCTC 1000
CCTCCCTTCTTTACAGACACACACGCACCGTCAGGCCGCGGCCACTTCCT GCCCTTTGTTTTTCTTCCTCCAAACATTTTTGGCCTCTTGCATTAGCCGCC 1100
TTAGGGCTCCGGGCGACGATTAAAAGAAATAAACGTTGAAGAGCGTTTTC TGGAGACAGCCCCGTGCTGTGGAGTTCTTGAGAAAAAGCAGCTCCTGGGA 1200
AACCAATGTCCAGGGCCAATGTCATAGCCTTGCCGGAAGCCGCCTCTTCT AGAAATGGAACTCGGACTTGGCCTTTTGGTGTGCTCTTATCCCTAGTTGG 1300
GTCTTCCCCAGATCCTGGTGGGCTGAAATCAACCTTGGGCCTCCTGGAAA GCCAAAGAACCAGTTAAGCTCAAAGCAAAGTGAAAACCAAGGACCCTCAG 1400
AGACCTCCCTTTCTTCCTCCAGATTTGGTCCGAACCTCAGGGCTGCTCGC ACCTTGTACCCTGATGCATGACTGAAGAACATTCACACTACCAGGGCCATC 1500
GCAGGCAAAGCGGCTACTATTGAGCAGCCTTAGGGATCTGAGCATCCATT TCTTCTCTCTGCTTAAGAGAACTAAACGCAATCGGCATCTCCAGGCCATC 1600
CGCCTCCCCCCCACCCCCATCAAAAGCCTTTCCCTCATCCATCTTCATCG CAGCGCCCCACACGCTCCCTTCGAGGATGCCCCTCTGGGCACCCTCTT 1700
CTGTAGAGCCTGTCGAATTGCTCTGTCAGCCTCTGTGGGGCTCCTTCTTC TTCCGCAGTCTTTGCTCGGGAATTTGTCTTGTCTCCTGCTGACACCCATG 1800
GGATACAGACCTGCCGGCTCCTATTCAATTAGCCTCTCCCCGAGAACCTAG GAAGCACTGGCTGTGCAGAACACTGCACAGTGCCGCACCGACCTTTCTGT 1900
GGGTACCCCTTCTCCCTGGCATCCCGGGAAGTCGAGTTAATTATGAGAGA AGCTGCCATGTAGTCAGGGGAGGCGGCCTGGCGTCCCCCGCGGCTGAGCC 2000
ACCTCGGGCCCCACTGCTCAGGCACAAGCTGAAGTTGACTGCCAAAACTG CTATAAAACCACAGGTCTGTCTAGAACTGTTTTCGGTTCTCCTAGACGCC 2100
CCCGGGTTCAATTATGGTCTTAAAAGCCCTAATGCAAGTGCAAGAATTTG TTTGCCTTCACCCAGGGGCGGGGCGAGCCCTGCAATGGTGGCGCCTTTGT 2200
TATAGCAAAGTCACTGAAACCCTTCATGCATTCCTTTTGTTCGAGGGAAG GGGTATTGAGGACTGGGTGGGCTGCTGAAGAAAGTTTGAGGTAGCAAAC 2300
CAACCCAGAGAGCTGTTCTTGTAAGGCCCTTTGCAAGCCCACCTCCCTTT TCTATCCGGATCTCCAGAAAGACAACCGTGAAGAGACAGATATGCTTGCC 2400
TCCTTGGTAACTGATAGAAATAGAAGTTTTTGCTGCAGGACCATCCTGTC GTTGTGATCCTCAGCTTTAGAAGACCTGAAGCTACCCTTTCCCAACACCC 2500
TCTCTCCCTCCTGAATCTCATGCAAACGCCGTGGCAGTCATCTATGAAGC CACAGTGAACAAATATGTCAAGCTGAAGATGCACAAATAATTTCAAGTTC 2600
AGCTCTCAGGGATTTCAAGGGGCATGAAAGTTCCTGCAGCAAATCCTGAA TAAAGAGTTCATTAAACCAATGTGTCCGAGTAGCTTGATTTGTGATAAGC 2700
AGCCCTTGTAAAAGAGAGGGGGTTTGGAGCCACATAAACATTAATGCTAA TAAATTAGTTTACTCCACTACACAGTAATTACTTGATATTATTGATATTT 2800
ACAGTGGGCATTCATATGCAATGGTGGGTCATTATTTGGAGAAAATACAT AACGAGAGTAGTTCTTTTCCAGAGCAATACATGATTAAATTTGTTATTGA 2900
GTCAGTTACAGTCAGCTCAGCAATAAATAAATAAATCGATGTTGACACTT AAATACCAAAGATCTTAGCGTTTATACTCTAAATCTCCCCAAGATATGTA 3000
AATACTTTAACTATTTCCTGAATGAGGAAAGGGAAAAAGGTTTTCATTT TTCACTTTTGACTGTTGTGACACCCTAGCACCTTCTAAGAAGGGAGTTGT 3100
GGTGAGACCCCTGGTCTGGAGCCCCATCTCTGCTCAATTTGCCTAGCTGT CCTGCGCCCTCTTTTTTCTCTTTCATTTTCCCCTAAGTACATCTTTGGTC 3200
TTGCAGAATGGCAGTCGTTATTTTATCTTCTTAATCCCCATTTCCTCCTC GGGGAGAAAGCTAAATGCAACTTAGCAGCCAAGGCCTCTTTAGAATTATCT 3300
ATCTGTACAGCAAGCCTTGCCCTTTTGAGGCATCTCTGTGGCCCAAGGTG CAGGCAGGCAGGAGGAGCATCTCTCTTTTCTTTGGCATTGTGCTTCCTG 3400
TGGGCTAAAGAACAAGAATGATGGGAGCTGTCTCAATTTCTGCCTTGAAT ATCCTTTGCTTTCACTCTTAGGTTTTGCAAGAAAGAAGAGACTCTTTTCT 3500
TAGCCAGGACTCCTTGGGCCAGTACTCCCTCTTTGCCCACCTCCTTTTCC CTCACAGCCTCTGAAGCCCTGCAAGCAGCCTGGTGCCCCCCCGACCCCCA 3600
CCCCAAGGTGCAGCATACCCAGGAGAAAGGCTCCTCAGCGCCTCTAAAACA GCCCAAGGTTCGCCTCAATTGCATGGTTTCCCGATTCGTCAGCTTCAGCA 3700
GACCAGGCAGCTTGCGGACCGGTGAGCAGAAGTGCCTGCTGGTCTGAAGGA GTCATCCGAACCCCTACAGGCAGCGAGGAGAAGGGGTTCCTTCAAGGAC 3800
AGGGTGTGGGGTAAGGAAGTAAGCAGAAAGATACAGGGAAGGTGGAACCC GGGTAAGACAAGGCCCTTCCGGACAGCGCGGATTATTGACAAACCGCGG 3900
AGGGAAGCCGCTTTGCCTGTTGGCGGTTTGGGGACGGAAGGCACTAAAGC GCTTCGGAAGTGACCGTGAATGAGAGCGTGTAATCAAGTCACCGTGCAAA 4000
TCGGACAAGCCACGAGGCCGGCACATCCGTGCCGGCGAAGCCTGGCCCCA AATGGGGTCCAACCAGGGTCTGAGGCTGGGGGGCTTCCAGAGCAAACCT 4100
GTGGGAGGGGGGCTGCGAAGAAGGTGCTGAACGTAAGCTTCGGATAACAG ACTCCGGGACCTGAAGCTATACAATTCCGGGCTTTCACCTGGCCACCCCC 4200
AGCCCAGAGAGTATCCCTCACCCACCGGGGAGTGGGGGGAGGCGTCAGGA GTTGCGCTCTCTTGCCGCCCTCAGAGCACTCGTGGACCAGGACCATGTAG 4300
CTGAGCAAAGGAGAGCTGCCCGGGGGCGCACCCCAGCCTTTCTGCGCGCG GGGAGGCCCCCCAGCCAACATGAGTTACACCGGCGCGGATTACGTGCTTTCGG 4400
TGAGAACACCGAGTGACGATCTGTTGCTTCCCCTGAGGTGGCTACAAAGA AAGGAAGCCGAGGAGGGAGGGGAGGGGGAAAAAAGGAAAGGGAGGGGGGTT 4500
AAAAAAAAAAAACCGGGACTAGGCTCGCGGCTTGTCAATTTCAACAT CGGGTCACATGACCAGCACCTCCCTGCTAAGGATGGGGGATAGATTTCCAC 4600
GTCAGCTTACGTCTCCAAATTTCTACTTCACGGATCCGCTTCAAAGA**GGC** **AGCTGCAGTGGAGAATCATGT**TAA**GCTCGGCTACTGCGGAGAGCCCAAGG** 4700
```

```
**TAGCCCAATGATGGATTTTGATGAGCGTGGTCCCTGCTCCTCTAACATGT** **ATTTGCCAAGTTGTACTTACTACGTCTCGGGTCCAGATTTCTCCAGCCTC** 4800
          M   M   D   F   D   E   R   G   P   C   S   S   N   M       Y   L   P   S   C   T   Y   Y   V   S   G   P   D   F   S   S   L
**CCTTCTTTTTTGCCCCAGACCCCGTCTTCGCGCCCAATGACATACTCCTA** **CTCCTCCAACCTGCCCCAGGTCCAACCCGTGCGCGAAGTGACCTTCAGAG** 4900
      P   S   F   L   P   Q   T   P   S   S   R   P   M   T   Y   S   Y       S   S   N   L   P   Q   V   Q   P   V   R   E   V   T   F   R
**AGTACGCCATTGAGCCCGCCACTAAATGGCACCCCCGCGGCAATCTGGCC** **CACTGCTACTCCGCGGAGGAGCTCGTGCACAGAGACTGTCTGCAGGCGCC** 5000
      E   Y   A   I   E   P   A   T   K   W   H   P   R   G   N   L   A       H   C   Y   S   A   E   E   L   V   H   R   D   C   L   Q   A   P
**CAGCGCGGCCGGCGTGCCTGGCGACGTGCTGGCCAAGAGCTCGGCCAACG** **TCTACCACCACCCCCACCCCCGCCGTCTCGTCCAATTTCTATAGCACGGTG** 5100
      S   A   A   G   V   P   G   D   V   L   A   K   S   S   A   N       V   Y   H   H   P   T   P   A   V   S   S   N   F   Y   S   T   V
**GGCAGGAACCGGCGTTCTGCCACAGGCTTTCGACCAGTTTTTTCGAGACGGC** **TTACGGCACCCCGGAAAACCTGCTTCCTCCGACTACCCCGGGGACAAGA** 5200
      G   R   N   G   V   L   P   Q   A   F   D   Q   F   F   E   T   A       Y   G   T   P   E   N   L   A   S   S   D   Y   P   G   D   K
**ACGCGCGAGAAGGGGCCCAAGCAGCAGCTCGACGCCTCCGCTGCGGCGGTG** **GCGGCGGCGGCCACGGGCGGCCGGCAACTTCAAGTTCGGACGGCGGCGG** 5300
      N   A   E   K   G   P   Q   A   A   A   T   S   A   A   V       A   A   A   A   T   G   A   P   A   T   S   S   S   D   G   G
**TGGCGGCGGCTGTCAGGAGGCGGCGGCGGAGGAGAAGGAGCGGACGGCGGC** **GACCCGAGAGCAGCAGCAGCCCCGAGTCGTCTTCCGGCCACACTGAGGAC** 5400
      G   G   G   C   Q   E   A   A   A   E   E   K   E   R   R       R   P   E   S   S   S   P   E   S   S   S   G   H   T   E   D
**AAGGCCGGTGGCTCCG**GTGCGTATAAAGGCAGCGCCCAGCGCTTTATGAA AGGGGAGTTGGGAATCTAGCGCAGCACCGCTGTTAAATTTTTATATGCTG 5500
      K   A   G   G   S
TTTTTATAAGCATATAAGGTTTATGAAGGGCCTTTAGGTTTCCCCCAACAA AACTTTGTTATAAAAGGAGGGGATCACGTGGCGAGTGCTGAAATTGGCCGT 5600
GAAATACCCACCGGCCAAAGCATGCCTGCAGAAAAAAAAAAAAAAACACGA CTCCCCTTTTTCCTTTTTTCCCTTTCTCAGCTCCCAAACTCAACAGCACGA 5700
TCCCCGCTAGAGACCCAGCTCCAGGCTGGGGAGCCGGAGGTGGGCGCCTGT AAATTCCCCCTACAAAGGGGCTTCAGGAGCCTTTGATAGCGGGGTAATCCGG 5800
AGATTTTTATAAAAGCCATGTGTTGCCGGCGGACTCCAGTCCCGGCTGGG ATTCAAAGGCATCAGGTTTAATGATGGCGCTAAGAGCGCATTTAACAG 5900
ATAAAGTAGCCCACTTAGCTCAAGCAATATATTAAAAGAAACAAATTAAC TTAAATTTGGCCTTTTTGTCCAAAGGAAGCCATTTTCCTCTGGTCCGGTT 6000
GCAATTACATGCTGCCTTCTGTTATTCACAATTGTTTTTAAATCAGTCAA AAGTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 6100
TGTGTGTTTCTTAAAGGTGAGGGTGAACAAATGGATCCTGCCTGAGACTT AAGGTAGCCTCTCTTTCCTCTACCCTTCCTAAACTAGCATCCCTAACCCTG 6200
CTGAGGGCCCTCTGGCAGTATGCTGACGGTGTGTACTGAGTGAAGGCCTG GGTAGGGGTTTCCTCTGCCACCTCCCACCTCGAGAGAGCTGGGAGGCTGG 6300
CGGAAGAGGCAAGAACTGGTGATCAGAAGTCAGGCTGACAAAGACCCCTC AGCCGCCCCAGATGTTAAGGAGGGCTAAATCATGGCGGCAAGCAGGAAAG 6400
CCTGGGGCCAGGATGCATGCCTTTACTGCCTCGGGAACCAGGCGGCCCAGG TGCTGAAGGTGAAATGTCCTCGGGCGGCTGTTTGTTTAGTCTGGCGGGGC 6500
AAGGAGCAGCCCGCCAACCCCCACTTGCCTCCGTTCAAAGCAGTTCAACCA GCGCAAACCTCGCTCGTTGCTGTGAGGAGCCTTCCTTTCAGCTCAAGGGGC 6600
AGCCTGGAGTTCCTTTCCCTGGGGGGCCTATCTAGTGAGCTGGCTTTTATCTG AAGCCGGCTCCAGACTGCCAACTTGTCCCTGACCAGCAATAGAAGCCTGAA 6700
GGCAGGGGCCTGGAAAGACTTTGAAAGGGCCTGGATTGGGGCCTCCAGACC CTGTGGCCAGACCTTCTCCAATGGGGGTCCTGGGAGAGGGGTTGCACCCC 6800
TATCGCCCTCTCACTCTTTGCCTCTCTCCTTCCTTAG**TGGCCAACGCAC** **CCGCAAAAGCGCTGCCCTTATACCCAAATACCAGATCCGAGAGCTGGAGC** 6900
                             G   G   Q   R   T       R   K   K   R   C   P   Y   T   K   Y   Q   I   R   E   L   E
**GAGAGTTCTTCTTCAGCGTCTACATCAACAAAGAGAACGGCGTCTGCAACTG** **TCCCGCATGCTCAACCTCACCGACCGTCAAGTCAAAATCTGGTTTCAAGAA** 7000
      R   E   F   F   F   S   V   Y   I   N   K   E   N   R   K   L   Q   L       S   R   M   L   N   L   T   D   R   Q   V   K   I   W   F   Q   N
**CAGAAGAATGAAGGAAAAAAAGATTAACAGA**GACCGGTTACAGTACTACT CAGCTAATCCACTTCTCTAAGGCTCCAGCCTACTGGAATTGGGAGGGGGG 7100
      R   R   M   K   E   K   K   I   N   R   D   R   L   Q   Y   Y       S   A   N   P   L   L   *
```

```
CTTCATACATGTGAAATAATATGCAGATTTTGCCCTTGACAAGGTCAAGC CACATTGCGACTTGGAAAGGGGGTGATATAGGAGAGGAGAGGTTTC 7200
CTTATCAGAGATTCGTCCAATAGGAAGGGGGCCTGAAACTCTCTTGGTTC AGATCTCCGTGGTTAAGCTTCCTATTAGTAATTGAATGCTAAGGGTTGTG 7300
AATCAACAGAATGAATGGTTTGGGACCCCTGGTGGTTCACTCTTAGAGCC TGCAAAAGCTTTGAGCTGGGTGTGGTCACTGGAGATTGGGCCCCCTGCCA 7400
AACCCTAAGGAGGCCACCCGCTCAAGACCCTCTGCACAGAGCCTGCCCAT CCCCAGGTCTCCCAAGTCAAGAAACTGTGTGCCCCAAGCCTAGTTCAGCT 7500
TGGGGATAGGAGCAGGAGGAAAGGGGATGACTCCTAGAGATTCAGGCTGG GGTTGCTTTCCAAAAGTAATGTGAAAGACAGTTAACCAAAAGGTGGCTGT 7600
GAAAAGGGAGAGAGAGTGGGAGGGGGGTGTCCTGGAGGAAGGAGAAACCT ATCCAGATGTCATTAATATCATCCCACCACTGATCTGCACCCAAACCTGA 7700
GCTCCTGTCAGCCACCTTCTGCAGCCTACCTGGGTTTTTCTACCTCCCAAT GGTGGTGATGGGTCTTCCATCTCAGAGCTGACTATTCCTAGGGTAGAAGA 7800
CATTGAAAACTCCAGAGGCTGTGCTGGTTCCTGGCCCCTCAATCTCTACA GGAAATGCCTGCGGAGGGCTACAGTTTGGCAAACTCTTCACCCACATTAAG 7900
GAAACTTGGATTTGGTGTCTGTCAGGCAGCAGAGCAGCAGCAGCTGGCCAG TGGGTCCTTGGCCTTTGGCTGTGAGGGGCTGAAGTTCCCCCCCCTCCCA 8000
AATCCCTCATTCCTACCCTGCACTCTGTGCTCATCAACTCAGAACTGAGC TTGGAAGCTTCTGGCAATCTTCTAAGAGCCGGAGAGTGATTTGGAGCTAT 8100
TTTAAAAGATAAATATTATATTATATATATATATATATTTCCCCGAAGGAAC CAAACGAATTTAAGAAGCAATGTAGAGGGGGGGGGTGGGGATAATAAA 8200
AATATTTAAAGGCTCTATCTGTTTACAGTGTTCAAAGTTTAAACTCGCTC ACTGACTAAAAATATTTGAAGTGTATGCTTCATACAGGGATGGTGTTCAAAAG 8300
**ACTGTA**AAATAAA**GAAACCATAACCTA**CTGTGTTTCAATACTCCTTTCCCT ATCCCCATTAGCTGACTTATTTGAGGGTGCTGGGGAGAGTGGTCTGGGGA 8400
TGGTCTTTTAATCTCTTACAACACTGTGACATTTGGAAAGGGCTGGCATA CTGAGAGCCACGATCTACTCAAGACCTCGAGAATGTCCTCTGGGAGCCAGGC 8500
CTCTCCAAATGAATCTCTTCTGCTCCCTCCCTGCCCTTAAGAAATCCCTGGTG TTCTGTCTGTAGTGAGCGCCTTTGGCCGGCCCTTACTCAACATACATTTC 8600
CTATACATTTTAACACAGAGAGAGAGAGCAATGGGGCATACCAAGAATAAAAA GGGTAGAAATTTGAGTATCGAGTCCTGGTGTGACATTTCTCTCTCCCTGT 8700
TCCGCTTTTTCCTGCCAAATGTCTCCCCCTCAGCTAGCATCCTTTTAGCAG TAAAAGGGGTGAGTTACCACGCCCCCGAAAGGTGATGGAATCAGGATTGA 8800
GCTGCAATTGAAGCTAAGCATAGGAAGCTCGAGGCTATACGTCCTTGTGG AAGGCAGCATATTAGGGATGGGATATGTCAGGGCCTATGGTGTGAGACCT 8900
CTCTCCAACTGACCTTAGCTCCGGAGAAGGCCTTGTGGGTAACTGCAATA GCCAATGGCCTAACCCAGGCTAATCTTGCTCACCTCAGGACTGAATAAAA 9000
ACCCTCTGGGTGTACCTTGTGAGTCCTTTCAGCATGAG
```
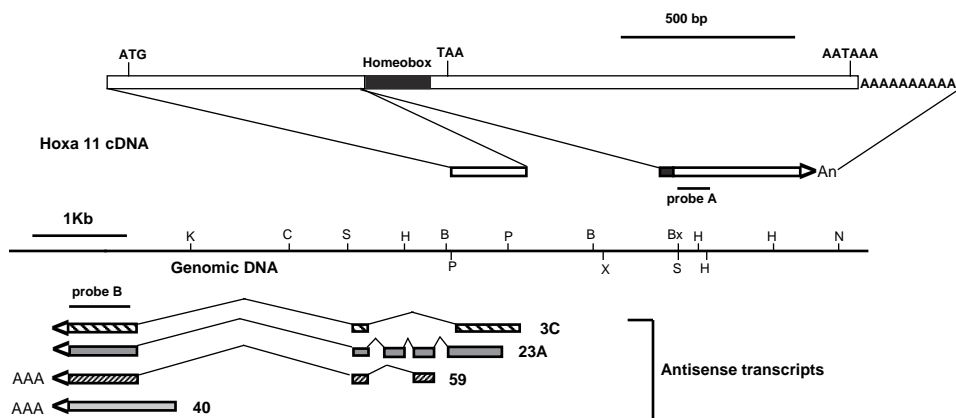
Fig. 2. Determination of the transcription initiation site of *Hoxa 11* by primer extension. A radiolabeled oligonucleotide corresponding to position 51-70 of the *Hoxa 11* cDNA (Fig. 1) was used as a primer. Three primer-extended products of 51, 71 and 77 bases were detected with total RNA from E11.5 mouse embryo hindlimbs (lane 1 and 2) but not from brain (lane 3). Marker lanes G, A, T and C indicate a sequencing ladder.

As shown in Fig. 3, the processing of the antisense transcripts is complex. The sequences of the last two exons of the 3C cDNA are shared with the final exons of the 23A and 59 cDNAs. Nevertheless, the preceding exons of each of these cDNAs are distinct, as is the exon arrangement of the 40 cDNA.

Antisense cDNA sequences are shown in Fig. 4. It is interesting to note that the 3C cDNA, in particular, shows significant overlap with the open reading frame of the *Hoxa 11* cDNA. The longest open reading frames of these antisense cDNAs are marked by the encoded amino acid sequences shown. For the 3C cDNA a potential amino acid sequence of 140 amino acids is encoded by the 5′ end. If this cDNA is

Fig. 4. The nucleotide sequence and ORFs of antisense cDNAs 3C, 23A, 59 and 40. The predicted encoded amino acids are shown below and the CTG codon of 23A is underlined.

incomplete at the 5′ end, as appears likely, then the actual open reading frame could be significantly longer. To date, attempts to extend these antisense cDNAs have instead resulted in the characterization of additional processed forms of the transcripts.

It is interesting to note that the potential 3C protein shown in Fig. 4 is actually encoded on the DNA strand opposite of the *Hoxa 11* coding strand. That is, the same region of DNA specifies *Hoxa 11* amino acid sequence on one strand and potentially encodes a distinct protein on the other strand. This putative 3C protein includes a proline-rich motif, with 50% of residues from position 5 to 35 being proline (P). In other proteins, such proline-rich motifs have been associated with transcription activation and repression domains (Mitchell and Tjian, 1989; Han and Manley, 1993a,b). The amino region of the protein is also particularly basic, with 33 of 140 amino acids being arginine (R), suggesting again possible interaction with DNA. It is also interesting that in a different frame 3C carries a 966 bp open reading frame, from bases 63 to 1029. The placement of the first ATG codon, however, at base 666, permits the encoding of only a 121 amino acid protein. Nevertheless, for some genes, CUG has been reported to serve efficiently as a translation initiator codon (Hann et al., 1994; Florkiewicz and Sommer, 1989; Prats et al., 1989). The CTG at base 210 of the 3C cDNA might therefore initiate a protein of 270 amino acids encoded by this large open reading frame.

The 23A cDNA has two significant open reading frames, with one encoding 165 amino acids at the 5′ end, and the other encoding 123 amino acids near the 3′ end. Splicing differences bring these two open reading frames together in the 3C cDNA to contribute to the 966 bp open reading frame observed there.

The abundance of statistically significant open reading frames, including one of almost 1 kb on the 3C cDNA, coupled with the richness of splicing alternatives, suggests but does not prove that some *Hoxa 11* antisense transcripts encode protein. Further characterization of additional cDNAs and immunocytochemistry will be required to establish this point.

Northern blot analysis

Northern blot analysis of embryonic mRNAs was conducted to better define the developmental time course of *Hoxa 11*

Fig. 3. The gene structure of the *Hoxa 11* genomic locus. *Hoxa 11* sense cDNA is shown above the genomic fragment and the four antisense cDNAs identified are shown below. Both 3C and 23A cDNAs were isolated from E12 limb library; cDNAs 59 and 40 were from an E12.5 library. Probes A and B were used to detect *Hoxa 11* and antisense transcript expression in northern blot and in situ hybridization (both serial section and whole mount). B, *Bss*H2; Bx, *Bst*X1; C, *Cla*1; H, *Hin*d3; K, *Kpn*1; N, *Nhe*1; P, *Pvu*2; S, *Sca*1; X, *Xho*1.

3C.

```
CGTCGCTCCTTCTCCTCCGCCGCCGCCTCCTGACAGCCGCCGCCACCGCC GCCGTCCGAACTTGAAGTTGCCGGCGCGCCCGTGGCCGCCGCCGCCACCG 100
      V  A  P  S  P  P  P  P  P  P  D  S  R  R  H  R    R  R  P  N  L  K  L  P  A  R  P  W  P  P  P  P  P
                                                             S  C  R  R  A  R  G  R  R  R  H  R
CCGCAGCCGGAGGTCGCAGCTGCTGCTTGGGGCCCCTTCTCGGCGTTCTTG TCCCCGGGGTAGTCGGAGGAAGCGAGGTTTTCCGGGGTGCCGTAAGCCGT 200
      P  Q  R  R  S  Q  L  L  L  G  A  P  S  R  R  S  C    P  R  G  S  R  R  K  R  G  F  P  G  C  R  K  P
         R  S  G  G  R  S  C  C  L  G  P  L  L  G  V  L    V  P  G  V  V  G  G  S  E  V  F  R  G  A  V  S  R
CTCGAAAAACTGGTCGAAAGCCTGTGGCAGAACGCCGTTCCTGCCCACCG TGCTATAGAAATTGGACGAGACGGCGGGGGTGGGGTGGTGGTAGACGTTG 300
      S  R  K  T  G  R  K  P  V  A  E  R  S  C  P  P    C  Y  R  N  W  T  R  R  R  G  W  G  G  R  R  W
      L  E  K  L  V  E  S  L  W  Q  N  A  V  P  A  H  R    A  I  E  I  G  R  D  G  G  G  V  V  V  D  V
GCCGAGCTCTTGGCCAGCACGTCGCCAGGCACGCCGGCCGCGCTGGGCGC CTGCAGACAGTCTCTGTGCACGAGCTCCTCCGCGGAGTAGCAGTGGGCCA 400
      P  S  S  W  P  A  R  R  Q  A  R  R  P  R  W  A    P  A  D  S  L  C  A  R  A  P  P  R  S  S  S  G  P
      G  R  A  L  G  Q  H  V  A  R  H  A  G  R  A  G  R    L  Q  T  V  S  V  H  E  L  L  R  G  V  A  V  G  Q
GATTGCCGCGGGGGTGCCATTTAGTGGCGGGCTCAATGGCGTACTCTCTG AAGGTCACTTCGCGCACGGGTTGGACCTGGGGCAGGTTGGAGGAGTAGGA 500
      D  C  R  G  G  A  I  *
         I  A  A  G  V  P  F  S  G  G  L  N  G  V  L  S    E  G  H  F  A  H  G  L  D  L  G  Q  V  G  G  V  G
GTATGTCATTGGGCGCGAAGACGGGGTCTGGGGCAAAAAAGAAGGGAGGC TGGAGAAATCTGGACCCGAGACGTAGGGTTCGGAGTGACTCCTCAGAGCC 600
      V  C  H  W  A  R  R  R  G  L  G  Q  K  R  R  E  A    G  E  I  W  T  R  D  V  G  F  G  V  T  P  Q  S
AGAGGCACTTCTGCTCACCGGTCCGCAAGCTGCCTGGTCTGCTGAAGCTG ACGAATCGGGAAACCATGCAATTGAGGCGAACCTTGGGCTGTTTTAGAGG 700
      Q  R  H  F  C  S  P  V  R  K  L  P  G  L  L  K  L    T  N  R  E  T  M  Q  L  R  R  T  L  G  C  F  R  G
CGCTGAGGAGCCTTCTCCTGGGAGGCCCAAGGTTGATTTCAGCCCACCAG GATCTGGGGAAGACCCAACTAGGGATAAGAGCACACCAAAAGGCCAAGTC 800
      A  E  E  P  S  P  G  R  P  K  V  D  F  S  P  P    G  S  G  E  D  P  T  R  D  K  S  T  P  K  G  Q  V
CGAGTTCCATTTCTAGAAGAGGCGGCTTCCGGCAAGGCTATGACATTGGC CCTGGACATTGGTTTCCCAGGAGCTGCTTTTTTCTCAAGAACTCCACAGCA 900
      R  V  P  F  L  E  E  A  A  S  G  K  A  M  T  L  A    L  D  I  G  F  P  G  A  A  F  S  Q  E  L  H  S
CGGGGCTGTCTCCAGAAAACGCTCTTCAACGTTTATTTCTTTTAATCGTC GCCCGGAGCCCTAAGGCGGCTAATGCAAGAGGCCAAAAATGTTTGGAGGA 1000
      T  G  L  S  P  E  N  A  L  Q  R  L  F  L  L  I  V    A  R  S  P  K  A  A  N  A  R  G  Q  K  C  L  E  E
AGAAAAACAAAGGCAGGAAGTGGCCGCGGCCTGACGGTGCGTGTGTGTCT GTAAAGAAGGGAGGGAGCCGGTTCAACCTCCCCTCGTTTTCCCGAACTTC 1100
      E  K  Q  R  Q  E  V  A  A  A  *
AAGGTCTAGGCAGACCCCCTTAGGGCCTTGCCGAGGCTCGCCCCCACACC CCCAGCGGCGCAGCATTTGGAGGTGGCCAACGATTTAAGCCTCGGTCGGG
CTGAAAGGAGATTTGATCGGCAGAACAAACCAACCCTTTTCGGAGGTTTC TTTTTGATTTGGTCCTAAAGGGTATATGCTAGTGTCCACAGCGGCTGGGGT 1300
GGCTGCTGTTTTCCTCCCGCCGGGCTAAAAGTACCAAGAAGGGAGGGAGG GAGAGAGATTCAGGCACCTTGCGCTGGCTGCACTCTCCTTCTGAGATAGA 1400
ATACCAGAATAAAGTGTATTCAGGTGCCTCTGCT
```

23A.

```
GTCGGAGGAAGCGAGGTTTTCCGGGGTGCCGTAAGCCGTCTCGAAAAACT GGTCGAAAGCCTGTGGCAGAACGCCGTTCCTGCCCACCGTGCTATAGAAA 100
  V  G  G  S  E  V  F  R  G  A  V  S  R  L  E  K  L    V  E  S  L  W  Q  N  A  V  P  A  H  R  A  I  E
TTGGACGAGACGGCGGGGGTGGGGTGGTGGTAGACGTTGGCCGAGCTCTT GGCCAGCACGTCGCCAGGCACGCCGGCCGCGCTGGGCGCCTGCAGACAGT 200
  I  G  R  D  G  G  G  V  V  V  D  V  G  R  A  L    G  Q  H  V  A  R  H  A  G  R  A  G  R  L  Q  T  V
CTCTGTGCACGAGCTCCTCCGCGGAGTAGCAGTGGGCCAGATTGCCGCGG GGGTGCCATTTAGTGGCGGGCTCAATGGCGTACTCTCTGAAGGTCACTTC 300
  S  V  H  E  L  L  R  G  V  A  V  G  Q  I  A  A    G  V  P  F  S  G  G  L  N  G  V  L  S  E  G  H  F
GCGCACGGGTTGGACCTGGGGCAGGTTGGAGGAGTAGGAGTATGTCATTG GGCGCGAAGACGGGGTCTGGGGCAAAAAAGAAGGGAGGCTGGAGAAATCT 400
  A  H  G  L  D  L  G  Q  V  G  G  V  G  V  C  H  W    A  R  R  R  G  L  G  Q  K  R  R  E  A  G  E  I
GGACCCGAGACGTAGTAAGTACAACTTGGCAAATACATGTTAGGAGGAGCA GGGACCACGCTCATCAAAATCCATCATTGGGCTACCTTGGGCTCTCCGCA 500
  W  T  R  D  V  V  S  T  T  W  Q  I  H  V  R  G  A    G  T  T  L  I  K  I  H  H  W  A  T  L  G  S  P  Q
GTAGCCGAGCTTAACATGATTCTCCACTGCAGCTGCCTCTTTGAAGCGGA TCCGTGAAGTAGAAATTTGGAGACCCACCTCAGGGGAAGCAACAGATCGT 600
  *
CACTCGGTGTTCTCACCGAAAGCACGTAATCGCCGGTGTAACTCATGTTG GCTGGGGGGCCTCCCCGCGCGCAGAAAGGCTGGGGTGCGCCCCCGGGCAG 700
CTCTCCTTTGCTCAGCTACATGGTCCTGGTCCACGAGTGCTCTGAGGGCG GCAAGAGAGCGCAACTCCTGACGCCTCCCCCCACTCCCCGCCCCCCTCCC 800
ACAGGTTTGCTCTGGGAAGCCCCCCAGCCTCAGACCCTGGCTGGACCCCA TTTGGGCCCAGGCTTCGCCGGCACGGATGTGCCGGCCTCGTGGCTTGTCC 900
GATTTGCACGGTGACTTGATTACACGCTCTCATTCACGGTCACTTCCGAA GCGCTTTAGTGCCTTCCGTCCCCAAACCGCCAACAGGCAAAGCGGCTTCC 1000
CTCCGCGGGGTTCGGAGTGACTCCTCAGAGCCAGAGGCACTTCTGCTCAC CGGTCCGCAAGCTGCCTGGTCTGCTGAAGCTGACGAATCGGGAAACCATG 1100
                                                                                                   M
CAATTGAGGCGAACCTTGGGCTGTTTTAGAGGCGCTGAGGAGCCTTCTCC TGGGAGGCCCAAGGTTGATTTCAGCCCACCAGGATCTGGGGAAGACCCAA 1200
  Q  L  R  R  T  L  G  C  F  R  G  A  E  E  P  S  P    G  R  P  K  V  D  F  S  P  P  G  S  G  E  D  P
CTAGGGATAAGAGCACACCAAAAGGCCAAGTCCGAGTTCCATTTCTAGAA GAGGCGGCTTCCGGCAAGGCTATGACATTGGCCCTGGACATTGGTTTCCC 1300
  T  R  D  K  S  T  P  K  G  Q  V  R  V  P  F  L  E    E  A  A  S  G  K  A  M  T  L  A  L  D  I  G  F  P
AGGAGCTGCTTTTTCTCAAGAACTCCACAGCACGGGCTGTCTCCAGAAA ACGCTCTTCAACGTTTATTTCTTTTAATCGTCGCCCGGAGCCCTAAGGCC 1400
  G  A  A  F  S  Q  E  L  H  S  T  G  L  S  P  E    N  A  L  Q  R  L  F  L  L  I  V  A  R  S  P  K  A
GCTAATGCAAGAGGCCAAAAATGTTTGGAGGAAGAAAAACAAAGGCAGGA AGTGGCCGCGGCCTGACGGTGCGTGTGTGTCTGTAAAGAAGGGAGGGAGC 1500
  A  N  A  R  G  Q  K  C  L  E  E  E  K  Q  R  E    V  A  A  A  *
CGGTTCAACCTCCCCTCGTTTTCCCGAACTTCAAGGTCTAGGCAGACCCC CTTAGGGCCTTGCCGAGGCTCGCCCCCACACCCCCAGCGGCGCAGCATTT 1600
GGAGGTGGCCAACGATTTAAGCCTCGGTCGGGCTGAAAGGAGATTTGATC GGCAGAACAAACCAACCCTTTTCGGAGGTTTCTTTTGATTTGGTCCTAAA 1700
GGGTATATGCTAGTGTCCACAGCGGCTGGGGTGGCTGCTGTTTTCCTCCC GCCGGGCTAAAAGTACCAAGAAGGGAGGGAGGGAGAGAGATTCAGGCACC 1800
TTGCGCTGGCTGCACTCTCCTTCTGAGATAGAATACCAG
```

59.

```
CTCAGGGGAAGCAACAGATCGTCACTCGGTGTTCTCACCGAAAGCACGTA ATCGCCGGTGTAACTCATGTTGGCTGGGGGGCCTCCCCGCGCGCAGAAAG 100
GCTGGGGTGCGCCCCCGGGCAGCTCTCCTTTGCTCAGCTACATGGTCCTG GTCCACGAGTGCTCTGAGGGCGGCAAGAGAGCGCAACTCCTGACGCCTCC 200
CCCCACTCCCCGGGGTTCGGAGTGACTCCTCAGAGCCAGAGGCACTTCTG CTCACCGGTCCGCAAGCTGCCTGGTCTGCTGAAGCTGACGAATCGGGAAA 300
CCATGCAATTGAGGCGAACCTTGGGCTGTTTTAGAGGCGCTGAGGAGCCT TCTCCTGGGAGGCCCAAGGTTGATTTCAGCCCACCAGGATCTGGGGAAGA 400
    M  Q  L  R  R  T  L  G  C  F  R  G  A  E  E  P    S  P  G  R  P  K  V  D  F  S  P  P  G  S  G  E  D
CCCAACTAGGGATAAGAGCACACCAAAAGGCCAAGTCCGAGTTCCATTTC TAGAAGAGGCGGCTTCCGGCAAGGCTATGACATTGGCCCTGGACATTGGT 500
    P  T  R  D  K  S  T  P  K  G  Q  V  R  V  P  F    L  E  E  A  A  S  G  K  A  M  T  L  A  L  D  I  G
TTCCCAGGAGCTGCTTTTTTCTCAAGAACTCCACAGCACGGGGCTGTCTCC AGAAAACGCTCTTCAACGTTTATTTCTTTTAATCGTCGCCCGGAGCCCTA 600
    F  P  G  A  A  F  S  Q  E  L  H  S  T  G  L  S  P    E  N  A  L  Q  R  L  F  L  L  I  V  A  R  S  P
AGGCGGCTAATGCAAGAGGCCAAAAATGTTTGGAGGAAGAAAAACAAAGG CAGGAAGTGGCCGCGGCCTGACGGTGCGTGTGTGTCTGTAAAGAAGGGAG 700
    K  A  A  N  A  R  G  Q  K  C  L  E  E  E  K  Q  R    Q  E  V  A  A  A  *
GGAGCCGGTTCAACCTCCCCTCGTTTTCCCGAACTTCAAGGTCTAGGCAG ACCCCCTTAGGGCCTTGCCGAGGCTCGCCCCCACACCCCCAGCGGCGCAG 800
CATTTGGAGGTGGCCAACGATTTAAGCCTCGGTCGGGCTGAAAGGAGATT TGATCGGCAGAACAAACCAACCCTTTTCGGAGGTTTCTTTTGATTTGGTC 900
CTAAAGGGTATATGCTAGTGTCCACAGCGGCTGGGGTGGCTGCTGTTTTC CTCCCGCCGGGCTAAAAGTACCAAGAAGGGAGGGAGGGAGAGAGATTCAG 1000
GCACCTTGCGCTGGCTGCACTCTCCTTCTGAGATAGAATACCAGAATAAA GTGTATTCAGGTGCCTCAAAAAAAAA
```

40.

```
ACAGAGGCTGACAGAGCAATTCGACAGGCTCTACAGAAGAGGGGTGCCCA GAGGGGCATCCTCGAAGGGAGCGTGTGGGGCGCTGCGATGAAGATGGAT 100
GAGGGAAAGGCTTTTGATGGGGGTGGGGGGGAGGCGGATGGCCTGGAGAT GCCGATTGCGTTTAGTTCTCTTAAGCAGAGAGAAGAAATGGATGCTCAGA 200
TCCCTAAGGCTGCTCAATAGTAGCCGCTTTGCCTGCTTGGCCCTGGTAGT GTGAATGTTCTTCAGTCATGCATCAGGGTACAAGGTGCGAGCAGCCCTGA 300
                                                        M  F  F  S  H  A  S  G  Y  K  V  R  A  A  L
GGGTTCGGACCAAATCTGGAGGAAGAAAGGGAGGTCTCTGAGGGTCCTTGG TTTTCACTTTGCTTTGAGCTTAACTGGTTCTTTGGCTTTCCAGGAGGCCC 400
    R  F  G  P  N  L  E  E  E  R  E  V  S  E  G  P  W    F  S  L  C  F  E  L  N  W  F  F  G  F  P  G  G  P
AAGGTTGATTTCAGCCCACCAGGATCTGGGGAAGACCCAACTAGGGATAA GAGCACACCAAAAGGCCAAGTCCGAGTTCCATTTCTAGAAGAGGCGGCTT 500
    R  L  I  S  A  H  Q  D  L  G  K  T  Q  L  G  I    R  A  H  Q  K  A  K  S  E  F  H  F  *
CCGGCAAGGCTATGACATTGGCCCTGGACATTGGTTTCCCAGGAGCTGCT TTTTCTCAAGAACTCCACAGCACGGGCTGTCTCCAGAAAACGCTCTTCA 600
ACGTTTATTTCTTTTAATCGTCGCCCGGAGCCCTAAGGCGGCTAATGCAA GAGGCCAAAAATGTTTGGAGGAAGAAAAACAAAGGCAGGAAGTGGCCGCG 700
GCCTGACGGTGCGTGTGTGTCTGTAAAGAAGGGAGGGAGGGAGCCGGTTCAACC TCCCCTCGTTTTCCCGAACTTCAAGGTCTAGGCAGACCCCCTTAGGGCCT 800
TGCCGAGGCTCGCCCCCACACCCCCAGCGGCGCAGCATTTGGAGGTGGCC AACGATTTAAGCCTCGGTCGGGCTGAAAGGAGATTTGATCGGCAGAACAA 900
ACCAACCCTTTTCGGAGGTTTCTTTTGATTTGGTCCTAAAGGGTATATGC TAGTGTCCACAGCGGCTGGGGTGGCTGCTGTTTTCCTCCCGCCGGGCTAA 1000
AAGTACCAAGAAGGGAGGGAGGGAGAGAGATTCAGGCACCTTGCGCTGGC TGCACTCTCCTTCTGAGATAGAATACCAGAATAAAGTGTATTCAGGTGCC 1100
TCTGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

expression and the size of sense and antisense transcripts produced. Poly(A)$^+$ RNA was isolated from embryos at 10.5, 11.5, 12.5 and 14.5 days of gestation and used to produce the northern blot panel shown in both Fig. 5A and B. For Fig. 5A, a 300 bp (probe A, Fig. 3) fragment from immediately 3′ of the homeobox was used. None of the antisense cDNAs isolated to date have sequence overlap with this region. Furthermore, when sense controls of this segment were used as probes for whole-mount and serial section in situ hybridizations, no signal was detected. This probe therefore appears specific for the *Hoxa 11* sense transcripts. In Fig. 5A, three distinct bands of hybridization are detected, corresponding to mRNAs of 1.9, 2.5 and 4.4 kb. The most abundant mRNA, of 2.5 kb, approximately matches the predicted polyadenylated size of the cDNA transcript described in this report. The other two fainter bands are not likely the result of non-specific sticking to residual 18S and 28S ribosomal RNA, as a flanking lane (not shown) was loaded with total RNA, which contained much more 18S and 28S rRNA, yet did not show these bands of hybridization. It is interesting to note that, in a separate northern blot, using mRNA purified from developing limbs, only a single intense band corresponding to the *Hoxa 11* 2.5 kb transcript was observed (data not shown). These observations suggest that the developing limbs specifically express the 2.5 kb transcript, with additional transcripts expressed in the developing body. Consistent with this, the original *Hoxa 11* cDNA was isolated from a total embryo E10.5 library, but all subsequent work to complete the sense cDNA used a limb-

specific cDNA library, and no evidence of *Hoxa 11* sense transcripts other than the one described in this report was found.

The embryonic northern panel in Fig. 5A was allowed to decay and rehybridized with the antisense probe designated B in Fig. 3. The resulting hybridization pattern is shown in Fig. 5B. Several bands, indicating transcripts of 1.4, 1.9, 2.9 and 3.2 kb, were observed. Consistent with the several fold greater number of antisense cDNAs found in cDNA library screens, the hybridization signal observed in Fig. 5B was significantly stronger than in Fig. 5A. This suggests that antisense transcripts are more abundant than standard *Hoxa 11* sense transcripts.

In situ hybridizations

The *Hoxa 11* antisense transcripts may function as regulators of the *Hoxa 11* transcripts. This would require an overlap in the *Hoxa 11* sense and antisense expression patterns. To better define the *Hoxa 11* expression domains and to establish the expression pattern of the antisense transcripts, we have conducted extensive whole-mount and serial section in situ hybridizations.

The *Hoxa 11* mutant mice show severely altered shapes for the tibia and fibula bones of the hindlimbs, and ulna and radius bones in the forelimb (Small and Potter, 1993). At E14, sections through the developing forelimbs and hindlimbs showed abundant expression of *Hoxa 11* mRNA. As shown in Fig. 6A, cross sections show localization of the signal to layers of immature mesenchymal cells that surround the cartilaginous forms of the developing long bones as well as the layers that separate the immature muscle bundles.

Expression of *Hoxa 11* was also observed in several organs in the embryo, particularly the urogenital tract. Fig. 6B shows expression in the metanephric kidney of an E14 embryo. The expression is restricted to the subcapsular zones of nephrogenesis in the blastemal cell population. No expression was present in the primitive tubules or glomerular structures. In the external genitalia of the male embryos, expression localized to the glans penis (Fig. 6C). Fig. 6D shows expression of *Hoxa 11* in the squamous epithelial layers of the skin and in the foregut mucosa of the stomach. Although no expression was observed throughout development in the male or female gonads, there was signal associated with the developing genital duct system (Fig. 6E,F). The immature stromal cells surrounding the paramesonephric duct (Mullerian duct) showed significant *Hoxa 11* transcription. No hybridization signal, however, localized to the epithelial cells lining these ducts. For the expression patterns in the body, single-stranded riboprobes that were sense (probe A of Fig. 3) and antisense (probe B of Fig. 3) specific were used, with indistinguishable results, indicating overlapping expression domains (data not shown). It remains possible that subtle differences will be revealed by future double label experiments.

A distinct complementary pattern of expression was seen for sense and antisense transcripts in the developing limbs using whole-mount in situ hybridization. We have previously shown that, at E9.5, *Hoxa 11* sense transcripts are abundant throughout the distal region of the limb buds (Small and Potter, 1993). As shown in Fig. 7, at E10.5, the *Hoxa 11* antisense transcripts are found at the distal tips of limb buds and simultaneously the sense transcripts disappear from this region. The embryos in Fig. 7A,B show forelimbs and hindlimbs, respectively, at
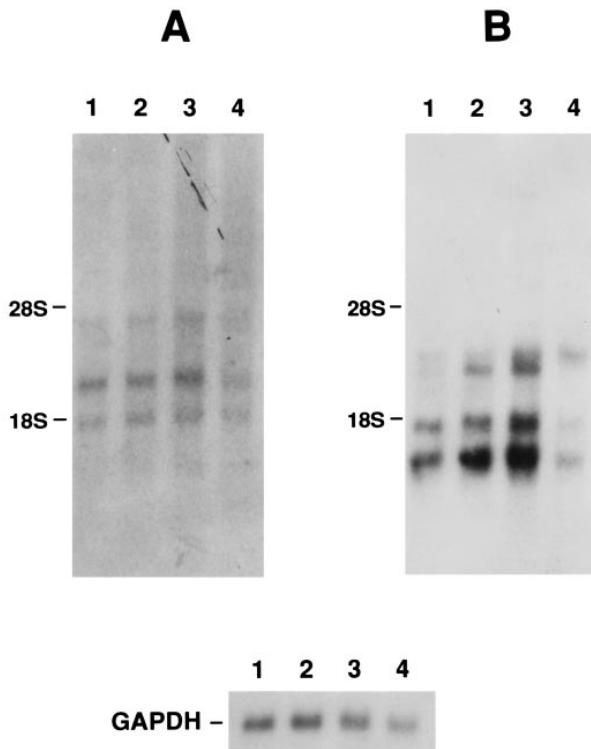

Fig. 5. Northern blot hybridization analysis of the temporal expression of (A) *Hoxa 11* sense and (B) antisense transcripts in mouse embryos. 10 μg of poly(A)$^+$ RNA from E10.5, 11.5, 12.5 and 14.5 embryos were loaded on lanes 1 to 4. Probes used are shown in Fig. 3.
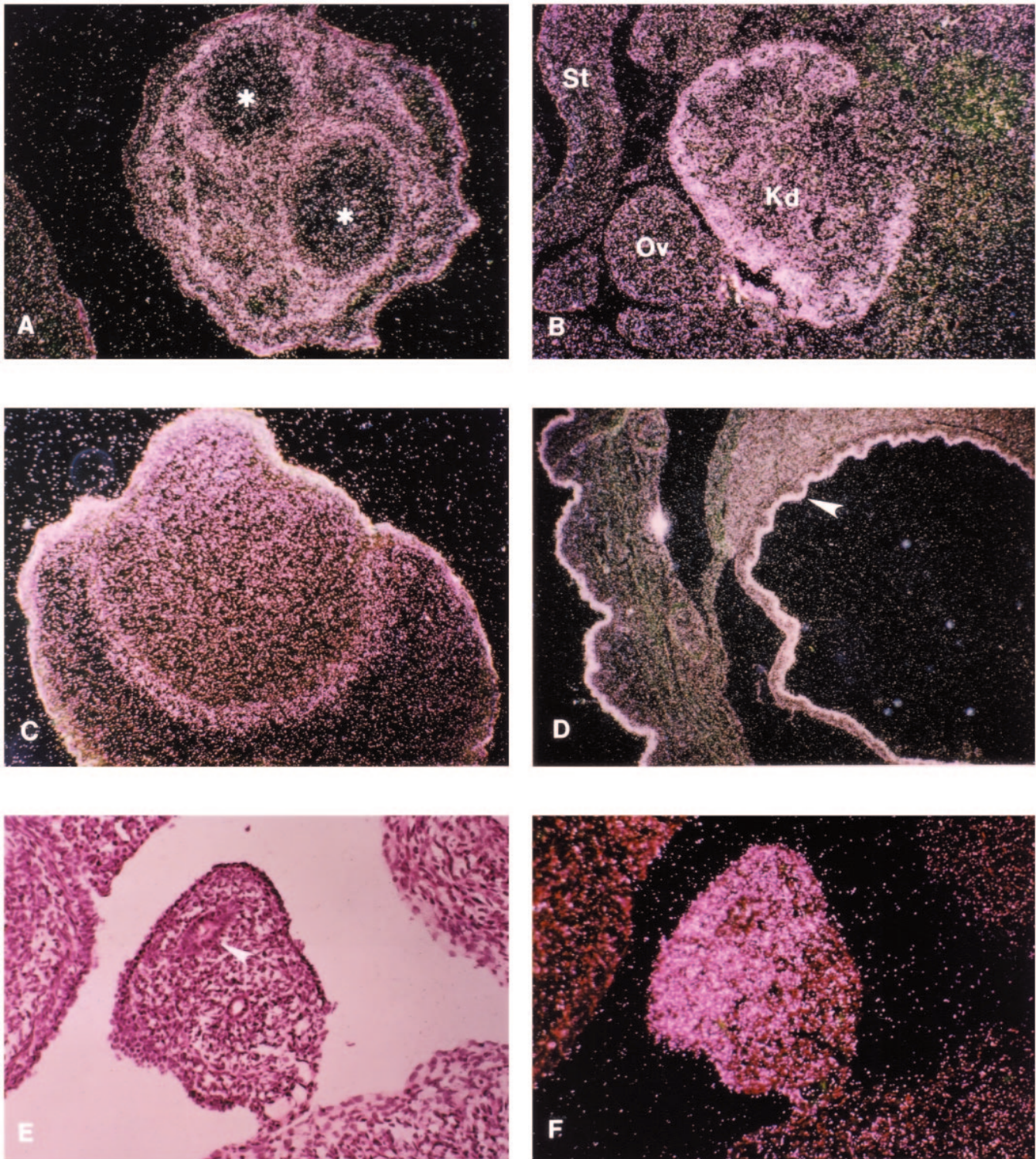
Fig. 6. Embryonic expression pattern of *Hoxa 11* by in situ hybridization. (A) Darkfield illumination of a section through the forelimb of an E14 mouse embryo hybridized with antisense probe for *Hoxa 11*. The white grains indicate localization of the *Hoxa 11* mRNA, which is restricted to bands of mesenchymal cells that surround the developing radius and ulna (*) as well as in bands of cells that separate developing muscle bundles. (B) Section from an E14 mouse embryo. The positive signal indicating *Hoxa 11* mRNA is localized to the nephrogenic zone in the developing kidney (kd). No signal is present in the ovary (Ov) or stomach (St). (C) Developing external genitalia of an E16 male embryo. The *Hoxa 11* mRNA is localized to the glans penis. (D) Section through the stomach and abdominal wall of an E16 mouse. There is strong hybridization in the squamous epithelium lining the forestomach (arrowhead) and in the skin covering the abdominal wall. (E,F) Bright-field and dark-field illuminations respectively of sections through the mesonephric and paramesonephric ducts in an E14 mouse. The stromal cells surrounding the ducts show a high level expression of *Hoxa 11*. No signal is present in the epithelium lining the mesonephric duct (arrowhead) or the adjacent paramesonephric duct. (A-D, 120× magnification; E,F, 240× magnification.)

E10.5. This trend, with sense transcripts disappearing as antisense transcripts appear, continues at E11.5, as shown in Fig. 7C. The antisense transcripts are first found in restricted domains at the most distal section of the limb, but by E11.5 antisense expression is much more widespread. The sense transcripts, in turn, are found at more proximally restricted regions as the antisense expression domain expands. This overall
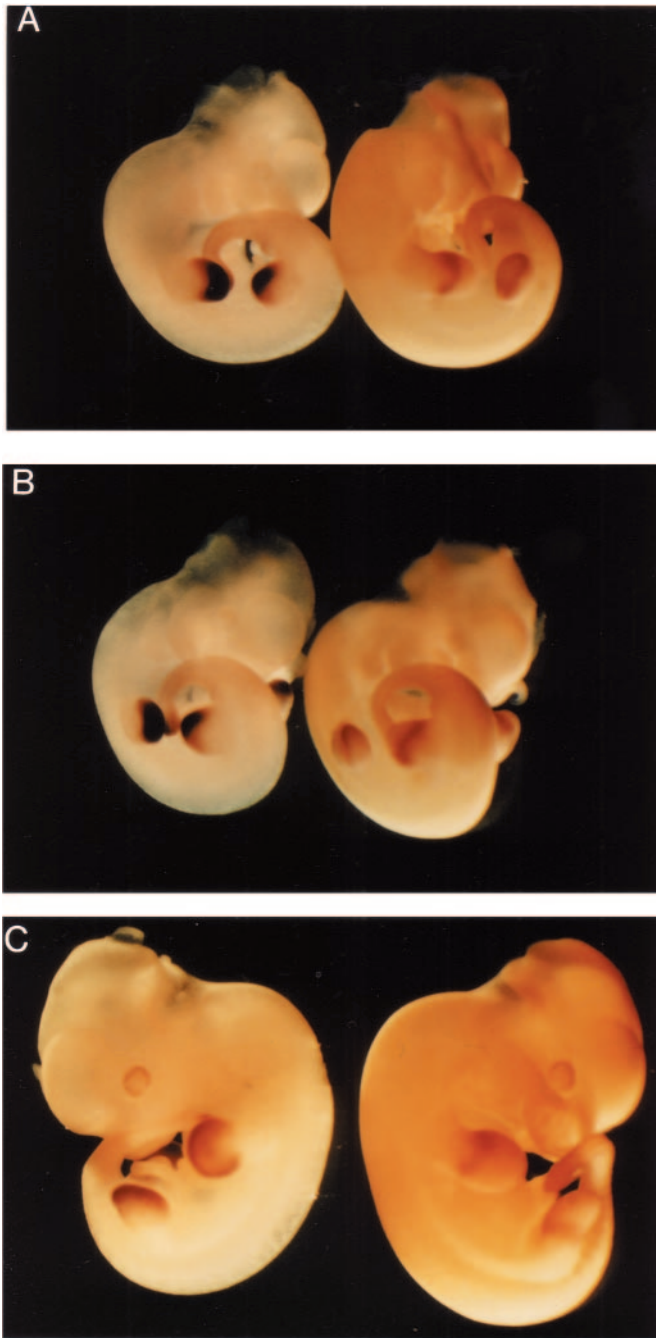
Fig. 7. Complementary expression patterns of *Hoxa 11* sense and antisense transcripts in the embryonic limb. Whole-mount in situ hybridizations show distinct complementary expression patterns of *Hoxa 11* antisense transcripts (embryo on left in A-C) and *Hoxa 11* sense transcripts (embryo on right in A-C) in E10.5 forelimbs (A) and hindlimbs (B) and in E11.5 limbs (C). Probes used are shown in Fig. 3.

pattern is consistent with an antisense role in regulating *Hoxa 11* sense transcript stability, as previously shown for the bFGF gene in *Xenopus* (Kimmelman and Kirschner, 1989).

Hoxa 11 mutants are sterile

Multiple matings set up between mature *Hoxa 11* homozygous mutant males and females failed to produce any progeny. To determine if the mutant males, or females, or both were sterile, matings between mutant and wild-type mice were performed. The mutant males copulated with wild-type females, producing vaginal plugs at a normal frequency, but no pups were born. Mutant females likewise mated normally with wild-type males, but produced no progeny. *Hoxa 11* homozygous mutant males and females therefore both appeared sterile.

The female sterility appeared to represent a maternal effect. From previous breeding of heterozygous *Hoxa 11* mutants, it was clear that zygotes with no functional copy of the *Hoxa 11* gene survived to birth and beyond at normal frequency. It was likely, therefore, that the *Hoxa 11* homozygous females were either incapable of producing ova, or that these ova were defective, or that the mothers were incapable of properly nurturing embryos to birth. Histologic examination of the ovaries of *Hoxa 11* mutants revealed no abnormalities (data not shown). Furthermore, on the day of observing the vaginal plug following mating with a wild-type male, normal numbers of zygotes were retrieved from the oviducts of mutant females. Similarly, at 1.5 and 2.5 days post coitus, development of embryos within the oviducts appeared normal. Morphologically normal blastocysts were also observed at E3.5.

Expression of *Hoxa 11* during development was not detected in the gonads, but was found in the cells surrounding the Wolffian and Mullerian ducts. It is therefore perhaps not surprising that, in the adult female reproductive system, *Hoxa 11* expression was restricted to the stromal cells of the uterus (Fig. 8A). These observations suggested that the reproductive defect in these females resided in the uterus. To test this hypothesis, we conducted embryo transfers, removing zygotes from *Hoxa 11* homozygous mutant females and placing them in the oviducts of surrogate mothers that had copulated with vasectomized males. Six such embryo transfers resulted in the birth of normal litters, strongly supporting the hypothesis. Further suggesting that the *Hoxa 11* female infertility is the result of a uterine defect is the observation that the heterozygous females

Fig. 8. *Hoxa 11* is expressed in the genital tract of the adult mouse. (A) The localization of *Hoxa 11* mRNA in a section through the uterus of a wild-type mouse. The signal is localized to the stromal cells of the endometrial lining but no signal is present in the glandular epithelium (arrowhead). (B) The localization of *Hoxa 11* mRNA in the vas deferens. The signal is present in the stromal cells beneath the mucosal epithelium which is negative. (C,D) Sections of wild-type and mutant vas deferens, respectively. (E) Stained paraffin section of testis from a wild-type adult mouse. Numerous sperm are shown along the luminal lining of the seminiferous tubules (arrowheads). (F) Section from the *Hoxa 11* mutant mouse. Although spermatogonia are present, there is no maturation in this population of germ cells. Note the extensive apoptosis in the germ cell population. (G) Section through the epididymis of a mature wild-type mouse. The tubules contain numerous mature sperm cells. (H) The epididymis from a *Hoxa 11* mutant. No mature sperm cells are present in the tubules. (A-E and G, 120× magnification; F,H, 240× magnification.)
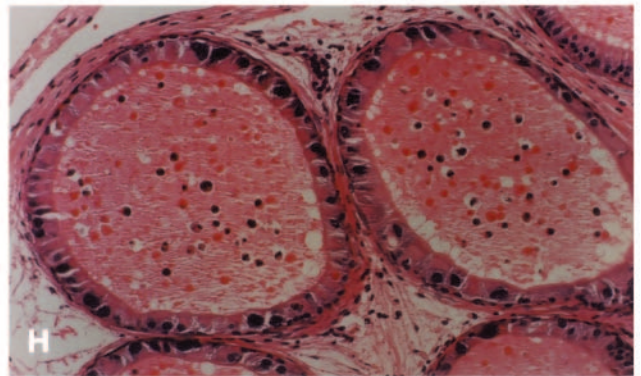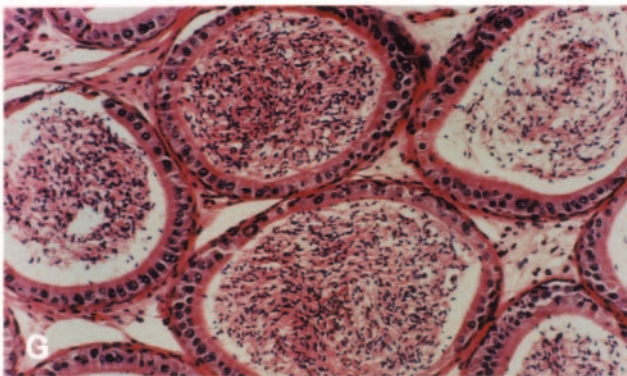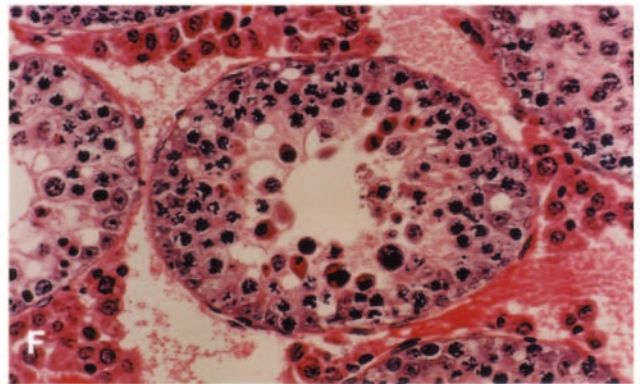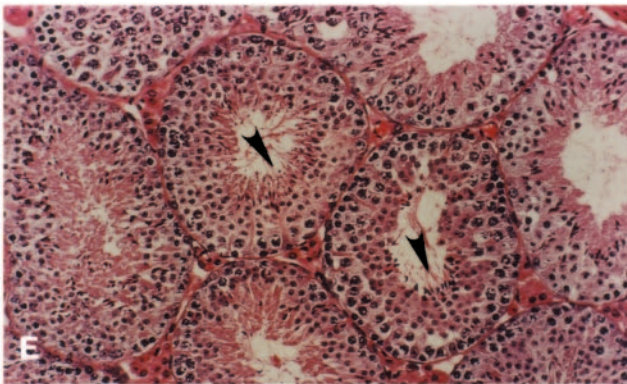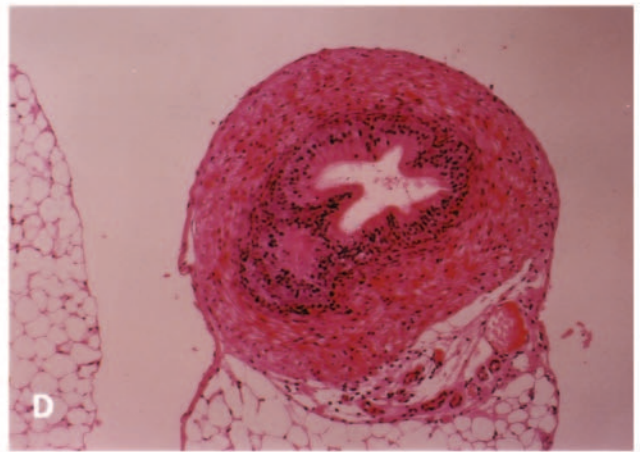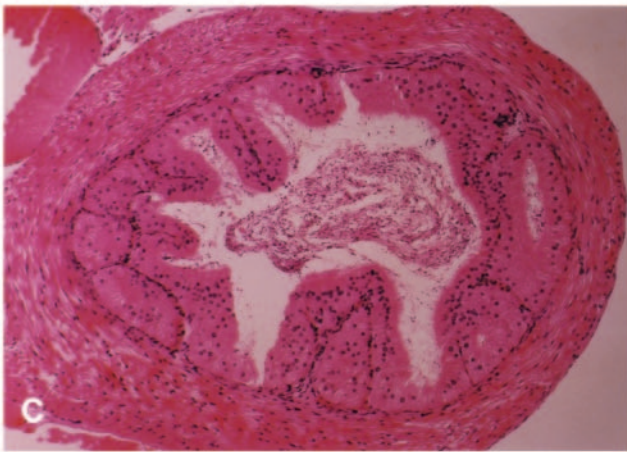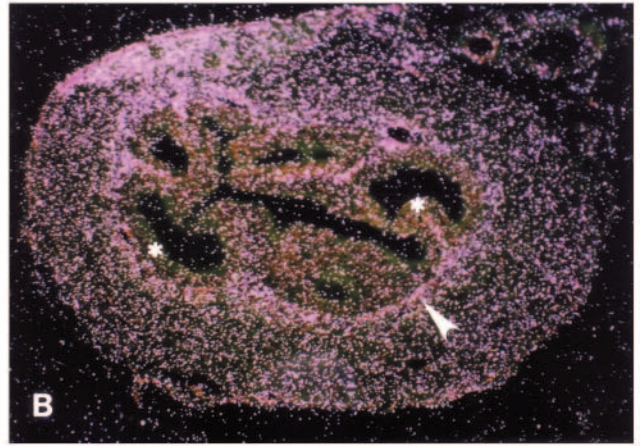
have severely reduced litter sizes. The uteri of pregnant heterozygous females were found to have only 165 living embryos out of a total of 480 implantation sites when examined at days 10.5 and 11.5 of gestation.

This indicates that the heterozygote uteri are only marginally capable of supporting embryonic development to term. Such a gene dosage effect is consistent with the observed intermediate malformations of bones and homeotic axial segment transformations in *Hoxa 11* heterozygotes. The accumulated data strongly suggest that in the *Hoxa 11* homozygous mutants there is defective uterine support of embryonic development. The precise nature of this defect is being further investigated.

The *Hoxa 11* homozygous mutant males copulate normally, and produce vaginal plugs, but no pups are born, even when mated with wild-type females. Similar to the females, no *Hoxa 11* expression was observed during gonad development. During development, the male reproductive tract showed *Hoxa 11* expression restricted to the cells surrounding the Wolffian ducts and, in the adults transcripts were localized to the stromal cells of the vas deferens (Fig. 8B). No expression was detected in the cells of the epithelial lining. This expression pattern is also consistent with malformations observed in the vas deferens of the mutant males. Histologic sections of the vas deferens showed the lumens were patent, but the luminal diameter was consistently smaller in the mutants (Fig. 8C,D). The epithelial lining of the mutant vas deferens also showed abnormalities. The epithelial lining was more simple in appearance as a result of less complex mucosal folding. The mutant epithelial lining cells were more narrow and columnar in appearance. The nuclei were also more basal in location in contrast to the normal vas deferens where the nuclei tended to be located more in the central or apical regions of the cells. The epithelial lining of the mutant animals also had more of a stratified organization compared to the wild type, where
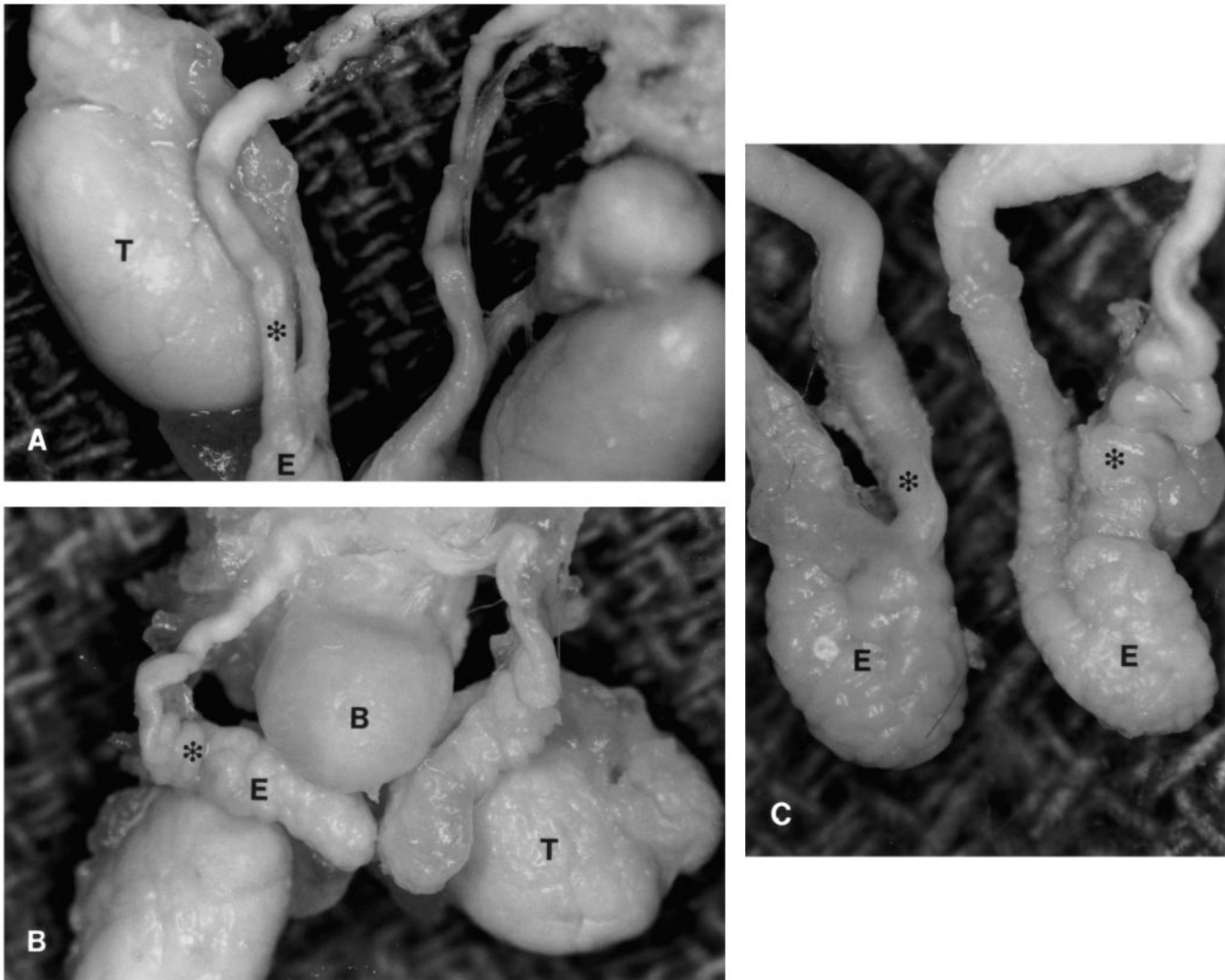
Fig. 9. *Hoxa 11* mutant mice have malformations in the genital tract. (A) The dissected reproductive tract from a wild-type mouse. (T, testis; E, cauda epididymis; *, vas deferens near the junction with the epididymis). (B) The dissected reproductive tract from a *Hoxa 11* mutant. The testes (T) are slightly smaller compared to the wild type. The vas deferens (*) is hypoplastic and markedly coiled. This is most notable at the junction between the vas deferens and the cauda epididymis (E). (B, bladder) (C) Side-by-side dissections of normal mouse epididymis (E) and vas deferens (*) on the left compared to the *Hoxa 11* mutant mouse on the right. Note the extensive folding in the proximal vas deferens in the mutant mouse making it difficult to distinguish where epididymis (E) ends and vas deferens (*) begins.

generally only a single layer of epithelial cells lined the vas deferens.

Prior to sexual maturity, the histology of the mutant testis was indistinguishable from wild type (data not shown). However, sections of sexually mature mutant testes showed evidence of altered spermatogenesis. This aspect of the phenotype showed variable expressivity, with Fig. 8E-H representing a severe example of failed spermatogenesis. The seminiferous tubules of the mutant testis (Fig. 8F) were lined by Sertoli cells and contained numerous spermatogonia but no mature sperm were present. Instead, many of the germ cells appear to be undergoing cell death by apoptosis. Similarly, the epididymis in the wild-type mouse contains numerous mature sperm but, in the mutant, it only contains cellular debris and fluid (Fig. 8G,H). In occasional mutant testes, some mature sperm cell populations were present, again consistent with the observed variable expressivity.

Gross level malformations were also observed in the reproductive tracts of the *Hoxa 11* homozygous mutant males (Fig. 9). The vas deferens in the mutant mice were smaller in diameter than normal, and generally had a highly coiled configuration, particularly at the proximal end near the junction of the vas deferens and the cauda epididymis. Indeed, the mutant vas deferens now resembles the epididymis in appearance. Furthermore, the testes of the mutant mice were smaller than normal in size and consistently showed incomplete descent into the scrotal sac. The degree of maldescent was variable, but in some animals the testes remained completely within the abdominal cavity and were located near the inferior pole of the kidneys.

DISCUSSION

In this report, the *Hoxa 11* gene structure is shown to be fairly typical among Hox genes, with a single intron and two exons spanning a total of only 3679 bp. The encoded homeodomain, which we had previously described (Singh et al., 1991), is also unremarkable for an *Abd-B* type Hox gene. There is an absence of the YPWM motif often found prior to the homeodomains of many Hox genes. In this same region, there is an interesting series of homopolymeric amino acid tracts, with polyalanine, polyglutamine, polyserine and polyarginine runs.

The characterization of the *Hoxa 11* cDNA led to the finding of abundant *Hoxa 11* antisense transcripts. In screening cDNA libraries, several fold more cDNAs corresponding to antisense, rather than sense *Hoxa 11* transcripts, were found. The sequences of four of these antisense cDNAs are presented in this report. These cDNAs, although likely incomplete, present an interesting picture of RNA processing alternatives, significant sequence overlap with the 5′ end of *Hoxa 11*, and coding potential.

It is interesting to consider the possible functions of these antisense transcripts. At one extreme, these transcripts may be 'spurious' and without function. A second possibility is that the antisense transcripts function in the traditional sense by encoding protein(s). The *Hoxa 11* antisense transcripts might also serve to regulate expression of the *Hoxa 11* gene. The occurrence of natural antisense regulator RNAs has now been shown in several organisms. In prokaryotes antisense transcripts appear to have numerous functions, including regula-

tion of mobile DNA element transposition, control of plasmid replication and incompatibility, regulation of bacteriophage genes and, in some cases, control of *E. coli* chromosomal genes (Simons, 1988). In *C. elegans*, the *lin-4* gene is believed to transcribe two small non-coding RNAs of 22 and 61 nucleotides that base pair with the 3′ UTR of the *lin 14* gene and block translation (Lee et al., 1993). In higher organisms, several genes with antisense transcripts have now been described. Two *c-erb* related genes overlap on the opposite strands of DNA (Miyajima et al., 1989; Lazar et al., 1989). For the dihydrofolate reductase gene short, non-coding, non-polyadenylated antisense transcripts have been detected (Farnham et al., 1985). Swiatek and Gridley (1993) have used *Hoxb 3* sense strand probes to detect the presence of antisense transcripts. For the insulin-like growth factor-2 (Rivkin et al., 1993; Baccarini et al., 1993), *c-myc* (Celano et al., 1992; Kindy et al., 1987) and *N-myc* (Krystal et al., 1990) genes, antisense transcripts with potential regulatory function have been described. Perhaps most interesting, however, is the *bFGF* gene, with a spliced, polyadenylated antisense mRNA that encodes an evolutionarily conserved $24{\times}10^3$ M_r protein of unknown function (Volk et al., 1989). Moreover, the antisense mRNA appears to direct the covalent modification of the *bFGF* transcripts, resulting in their rapid degradation (Kimmelman and Kirschner, 1989). The expression pattern of the *Hoxa 11* antisense RNA is consistent with a regulatory function. The limbs are of particular interest, with the appearance of the antisense transcripts coincident with the disappearance of the normal *Hoxa 11* transcripts in the distal regions of the limb bud.

It is interesting to note that the targeted disruption of the *Hoxa 11* gene described by Small and Potter (1993) deleted the entire coding sequence of the *Hoxa 11* cDNA presented here. This targeting therefore also disrupted antisense transcription units, at least as represented by the 3C and 23A cDNAs described in this report. The resulting mutant phenotype may therefore in part be due to the absence of some or all antisense transcripts.

The expression pattern of *Hoxa 11* in the developing kidney is quite suggestive. Although no histologic defects are observed in the *Hoxa 11* homozygous mutants, this may be the result of genetic redundancy, as ten other clustered Hox genes have now been reported expressed in the developing kidney (Bard et al., 1994). The metanephric kidney begins through inductive interactions between the ureteric bud and the metanephric mesenchyme. In the developing kidney, stem cells are located at the periphery, with differentiating aggregates immediately subperipheral and maturing nephrons more central. The expression of *Hoxa 11*, at the outer regions of the kidney, suggests function in early stages of nephron formation. Further experiments involving two or more targeted Hox genes will be required to test this proposition.

Hoxa 11 is required for both male and female fertility. *Hoxa 11* homozygote females produce normal eggs, which develop properly postfertilization when transplanted to a surrogate mother. The *Hoxa 11* mutant females provide a defective uterine environment, however, with even hemizygotes suffering extremely high embryo resorption rates. The molecular basis of this uterine defect is being further investigated. The *Hoxa 11* mutant phenotype is quite similar to that observed in mice with a targeted mutation in the leukemia

inhibitory factor (LIF) gene (Stewart et al., 1992). In both cases, oogenesis is normal, but embryos fail to thrive in the mutant uterus. This suggests the possibility of their involvement in a common genetic pathway.

In contrast, the *Hoxa 11* homozygous mutant males suffer a severe perturbation of gametogenesis. Although the required cellular components of spermatogenesis, including spermatogonia, Leydig cells and Sertoli cells are present and appear morphologically normal, the process is aberrant with spermatocytes undergoing cell death rather than forming spermatozoa. It is important to note that this phenotype does not show 100% penetrance, as we have observed one homozygote male capable of producing progeny. The molecular cause of the altered spermatogenesis remains uncertain, but gross dissection indicated that the mutant testes failed to descend from the abdomen. Mammalian testes generally require the reduced scrotum temperature to allow completion of spermatogenesis. Therefore the defective spermatogenesis is likely the result of the maldescent of the testes. Both androgen and Mullerian inhibitory substance have been implicated in the hormonal control of testicular descent (Hutson and Donahue, 1986). Moreover, it is relevant to note that several studies have correlated undescended testes in humans with epididymal abnormalities (Minenberg and Schlossberg, 1983; Heath et al., 1984; Marshall and Shermeta, 1979). It has been speculated that since the epididymis and vas deferens physically precede the testes, they may play a role in testicular descent.

The observed alteration of the mutant vas deferens is rather unique. At the gross level, the mutant vas is thinner and more coiled, in some cases even severely kinked, in appearance. The mutant vas deferens structurally resembles the epididymis. For this phenotype, there was variable expressivity, with the vas sometimes bearing a striking similarity to the epididymis, while in other cases the more caudal regions of the vas appeared normal. This apparent partial homeotic transformation of the vas deferens into epididymis was also suggested by histological examination, with the cellular architecture, including the morphology and location of nuclei of the epithelial lining, now resembling that seen in the epididymis. This is the first reported homeotic transformation of a non-skeletal element caused by a Hox gene mutation in mice.

REFERENCES

Baccarini, P., Fiorentino, M., D'Errico, A., Mancini, A. M. and Grigioni, W. F. (1993). Detection of anti-sense transcripts of the insulin-like growth factor-2 gene in Wilson's tumor. *Am. J. Pathol.* **143**, 1535-1542.

Bard, J. B. L., McConnell, J. E. and Davies, J. A. (1994). Towards a genetic basis of kidney development. *Mech. Dev.* **48**, 3-11.

Blochinger, K., Bodmer, R., Jack, J., Jan, L. Y. and Jan, Y. N. (1988). Primary structure and expression of a product from *cut*, a locus involved in specifying sensory organ identity in *Drosophila*. *Nature* **333**, 629-635.

Celano, P., Berchtold, C. M., Kizer, D. L., Weeratna, A., Nelkin, B. D., Baylin, S. B. and Casero, R. A. (1992). Characterization of an endogenous RNA transcript with homology to the antisense strand of the human *c-myc* gene. *J. Biol. Chem.* **267**, 15092-15096.

Celano, P., Vertino, P. M. and Casero, J., R.A. (1993). Isolation of

polyadenylated RNA from cultured cells and intact tissues. *BioTechniques* **15**, 27-28.

Davis, A. P. and Capecchi, M. R. (1994). Axial homeosis and appendicular skeleton defects in mice with a targeted disruption of *hoxd-11*. *Development* **120**, 2187-2199.

Dolle, P., Izpisúa-Belmonte, J.-C., Falkenstein, H., Renucci, A. and Duboule, D. (1989). Coordinate expression of the murine *Hox-5* complex homeobox-containing genes during limb pattern formation. *Nature* **342**, 767-772.

Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P. and Duboule, D. (1993). Disruption of the *Hox d-13* gene induces localized heterochrony leading to mice with neotenic limbs. *Cell* **75**, 431-441.

Farnham, P. J., Abrams, J. M. and Shinke, R. T. (1985). Opposite-strand RNAs from the 5′ flanking region of the mouse dihydrofolate reductase gene. *Proc. Natl. Acad. Sci. USA* **82**, 3978-3982.

Florkiewicz, R. Z. and Sommer, A. (1989). Human basic fibroblast growth factor gene encodes four polypeptides: Three initiate translation from non-AUG codons. *Proc. Natl. Acad. Sci. USA* **86**, 3978-3981.

Haack, H. and Gruss, P. (1993). The establishment of murine *Hox-1* expression domains during patterning of the limb. *Dev. Biol.* **157**, 410-422.

Han, K. and Manley, J. (1993a). Transcriptional repression by the *Drosophila* Even-skipped protein: definition of a minimal repression domain. *Genes Dev.* **7**, 491-503.

Han, K. and Manley, J. L. (1993b). Functional domains of the *Drosophila* Engrailed protein. *EMBO J.* **12**, 2723-2733.

Hann, S. R., Dixit, M., Sears, R. C. and Sealy, L. (1994). The alternatively initiated c-Myc proteins differentially regulate transcription through a noncanonical DNA binding site. *Genes Dev.* **8**, 2441-2452.

Heath, A. L., Man, D. W. K. and Eckstein, H. B. (1984). Epididymal abnormalities associated with maldescent of the testis. *J. Plastic Surg.* **19**, 47-49.

Hogan, B., Costantini, F. A. and Lacy, E. (1986). *Manipulating the Mouse Embryo*. New York: Cold Spring Harbor Laboratory.

Hutson, J. M. and Donahoe, P. K. (1986). The hormonal control of testicular descent. *Endocrine Rev.* **7**, 270-283.

Kania, M. A., Bonner, A. S., Duffy, J. P. and Gergen, J. P. (1990). The *Drosophila* segmentation gene *runt* encodes a novel nuclear regulatory protein that is also expressed in the developing nervous system. *Genes Dev.* **4**, 1701-1713.

Kern, M. J., Witte, D. P., Valerius, M. T., Aronow, B. J. and Potter, S. S. (1992). A novel murine homeobox gene isolated by a tissue specific PCR cloning strategy. *Nucl. Acids Res.* **20**, 5189-5195.

Kimmelman, D. and Kirschner, M. W. (1989). An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus* oocytes. *Cell* **59**, 687-696.

Kindy, M. S., McCormack, J. E., Buckler, A. J., Levine, R. A. and Sonensheim, G. E. (1987). Independent regulation of transcription of the two strands of the *c-myc* gene. *Mol. Cell. Biol.* **7**, 2857-2862.

Krystal, G. W., Armstrong, B. C. and Battey, J. F. (1990). *N-myc* mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol. Cell. Biol.* **10**, 4180-4191.

Lazar, M. A., Hodin, R. A., Darling, D. S. and Chin, W. W. (1989). A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat *c-erb A* and transcriptional unit. *Mol. Cell. Biol.* **9**, 1128-1136.

Lee, R., Feinbaum, R. L. and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854.

Li, H., Witte, D. P., Branford, B. W., Aronow, B. J., Weinstein, M., Kaur, S., Wert, S., Singh, G., Schreiner, C. M., Whitsett, J. A., Scott, W. J. and Potter, S. S. (1994). *Gsh-4* encodes a LIM-type homeodomain, is expressed in the developing central nervous system and is required for early postnatal survival. *EMBO J.* **13**, 2876-2885.

Macdonald, P. M., Ingham, P. and Struhl, G. (1986). Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell* **47**, 721-734.

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.

Marshall, F. F. and Shermeta, D. W. (1979). Epididymal abnormalities associated with undescended testis. *J. Urology* **121**, 341-343.

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

Mevel-Ninio, M., Terracol, R. and Kafatos, F. C. (1991). The *ovo* gene of *Drosophila* encodes a zinc finger protein required for female germ line development. *EMBO J.* **10**, 2259-2266.

Minenberg, D. T. and Schlossberg, S. (1983). The role of the epididymis in testicular descent. *J. Urology* **129**, 1207-1208.

Mitchell, P. J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

Miyajima, N., Horinchi, R., Shibuya, Y., Fukushige, S., Matsubara, K., Toyoshima, K. and Yamamoto, T. (1989). Two *erb A* homologs encoding proteins with different T_3 binding capacities are transcribed from opposite DNA strands of the same genetic locus. *Cell* **57**, 31-39.

Poole, S. J., Kauvar, L. M., Drees, B. and Kornberg, T. (1985). The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. *Cell* **40**, 37-43.

Prats, H., Kaghad, M., Prat, A. C., Klagsburn, M., Lélias, J. M., Liauzin, P., Chalon, P., Tauber, J. P., Amalric, F., Smith, J. A. and Caput, D. (1989). High molecular mass forms of basic fibroblast growth factor are initiated by alternative GUC codons. *Proc. Natl. Acad. Sci. USA* **86**, 1836-1840.

Rivkin, M., Rosen, K. M. and Villa-Komaroff, L. (1993). Identification of an antisense transcript from the IGF-II locus in mouse. *Mol. Rep. Dev.* **35**, 394-397.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning, A Laboratory Manual.* 2nd Edition. pp. 771-783. Cold Spring Harbor: Cold Spring Harbor Press.

Simons, R. W. (1988). Naturally occurring antisense RNA control-a brief review. *Gene* **72**, 35-43.

Singh, G., Kaur, S., Stock, J. L., Jenkins, N. A., Gilbert, D. J., Copeland, N. G. and Potter, S. S. (1991). Identification of 10 murine homeobox genes. *Proc. Natl. Acad. Sci. USA* **88**, 10706-10710.

Small, K. M. and Potter, S. S. (1993). Homeotic transformation and limb defects in *Hox A11* mutant mice. *Genes Dev.* **7**, 2318-2328.

Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., Kontgen, F. and Abbondazo, S. J. (1992). Blastocyst implantation depends on maternal expression of leukemia inhibitory factor. *Nature* **359**, 76-79.

Swiatek, P.J. and Gridley, T. (1993). Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene Krox20. *Genes Dev.* **7**, 2071-2084.

Volk, R., Koster, M., Poting, A., Hartmann, L. and Knochel, W. (1989). An antisense transcript from the *Xenopus laevis bFGF* gene coding for an evolutionarily conserved 24 kd protein. *EMBO J.* **8**, 2983-2988.

Yokouchi, Y., Sasaki, H. and Kuroiwa, A. (1991). Homeobox gene expression correlated with the process of limb cartilage development. *Nature* **353**, 443-445.