Development 119, 419-431 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants

Thierry J. Lints, Linda M. Parsons, Lynne Hartley, Ian Lyons and Richard P. Harvey*

The Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Victoria 3050, Australia

*Author for correspondence

SUMMARY

We have isolated two murine homeobox genes, *Nkx-2.5* and *Nkx-2.6*, that are new members of a sp sub-family of homeobox genes related to *Drosophila NK2, NK3* and *NK4/msh-2*. In this paper, we focus on the *Nkx-2.5* gene and its expression pattern during post-implantation development. *Nkx-2.5* transcripts are first detected at early headfold stages in myocardiogenic progenitor cells. Expression preceeds the onset of myogenic differentiation, and continues in cardiomyocytes of embryonic, foetal and adult hearts. Transcripts are also detected in future pharyngeal endoderm, the tissue believed to produce the heart inducer. Expression in endoderm is only found laterally, where it is in direct apposition to promyocardium, suggesting an interaction between the two tissues. After foregut closure, *Nkx-2.5* expression in endoderm is limited to the pharyngeal floor, dorsal to the developing heart tube. The thyroid primordium, a derivative of the pharyngeal floor, continues to express *Nkx-2.5* after transcript levels diminish in the rest of the pharynx. *Nkx-2.5* transcripts are also detected in lingual muscle, spleen and stomach.

The expression data implicate *Nkx-2.5* in commitment to and/or differentiation of the myocardial lineage. The data further demonstrate that cardiogenic progenitors can be distinguished at a molecular level by late gastrulation. *Nkx-2.5* expression will therefore be a valuable marker in the analysis of mesoderm development and an early entry point for dissection of the molecular basis of myogenesis in the heart.

Key words: heart, myogenesis, myocardium, homeobox gene, heart inducer, endoderm, mesoderm formation

INTRODUCTION

The homeobox gene superfamily encodes transcription regulatory proteins that act at critical points in development and ontogeny (Scott et al., 1989; Shashikant et al., 1991; McGinnis and Krumlauf, 1992). The hallmark of these proteins is a structurally conserved DNA-binding motif termed the homeodomain. Biochemical and genetic data have shown that homeodomain proteins are sequence-specific DNA-binding factors (Hayashi and Scott, 1990).

The most characterized homeobox genes are those belonging to the *HOM-C* complex of *Drosophila* and the *Hox* clusters of mammals (McGinnis and Krumlauf, 1992). The presence of similar clusters in lower metazoa (Kenyon and Wang, 1991; Schummer et al., 1992) indicates that these genes have been maintained throughout evolution as a highly conserved and presumably functionally homologous genetic unit. *HOM-C* and *Hox* clusters show conserved gene structure, gene order and relative timing of gene activation, as well as similarities in embryonic expression patterns (McGinnis and Krumlauf, 1992). The suggested functional conservation of the *HOM-C* and *Hox* genes and associated regulatory regions is highlighted by their equivalence in interspecific gene transfer experiments (Malicki et al., 1990, 1992; McGinnis et al., 1990; Awgulewitsch and Jacobs, 1992). Expression of the *HOM-C/Hox* genes appears to specify positional identity within cellular fields, both metameric, such as the *Drosophila* cuticle, and non-metameric, such as the vertebrate limb (Hunt and Krumlauf, 1991; Tabin, 1992; McGinnis and Krumlauf, 1992). Mutation, or ectopic expression of *HOM-C/Hox* genes in flies and mice can lead to homeotic changes within metameres (Kessel et al., 1990; Pollock et al., 1992; McGinnis and Krumlauf, 1992; Jegalian and De Robertis, 1992; Le Mouellic et al., 1992).

In addition to *HOM-C/Hox* genes, numerous other homeobox genes have been identified (Shashikant et al., 1991) and these are typically dispersed in the genome. Some of them also appear to be involved in specifying positional information (Ruiz i Altaba and Melton, 1990), while others may play a role in lineage maturation and cell type discrimination (Ingraham et al., 1988; Guazzi et al., 1990; Bodmer et al., 1990; Tronche and Yaniv, 1992), morphogenesis (Dressler and Douglass, 1992) and/or growth control (Kamps et al., 1990; Hatano et al., 1991; Song et al., 1992; Lin et al., 1992).

The participation of homeobox genes in lineage commitment is not well defined. Commitment has been most extensively studied in the skeletal muscle lineage (Weintraub et al., 1991), where transcription factors of the helix-loop-helix

(HLH) superfamily (*myoD*, *myogenin*, *myf-5* and *MRF-4*) play an important and dominant role in commitment and differentiation. Although HLH myogenic genes have the remarkable property of inducing myogenesis when transfected into a broad range of cell types, their action is nonetheless context-dependent (Weintraub et al., 1989; Dekel et al., 1992) and the absence of their expression in limb myoblasts suggests that they act late in progression to the committed state (Sassoon, 1993). Furthermore, multiple routes to the myogenic phenotype might be possible, since HLH myogenic genes are not expressed in mammalian cardiac muscle (but see Litvin et al., 1993) and are not absolutely required for muscle formation in *C. elegans* (Chen et al., 1992).

The structural and functional conservation of the *Drosophila HOM-C*/*Hox* genes throughout evolution suggests that non-*HOM-C* homeobox genes involved in ancient developmental processes will be similarly conserved. The aim of this study was to screen for murine homologues of the *Drosophila* homeobox gene *NK4* (also called *msh-2*), expressed during *Drosophila* development in mesodermal progenitor cells and later in the heart and gut muscle lineages (Kim and Nirenberg, 1989; Bodmer et al., 1990). *NK4/msh-2* mutants have no detectable heart or gut muscle progenitors, suggesting a role for the gene in muscle commitment. Using polymerase chain reaction (PCR) and hybridization screening, we have isolated three murine homeobox genes, *Nkx-2.3*, *Nkx-2.5* and *Nkx-2.6*, closely related to *NK4/msh-2* as well as to *DrosophilaNK2* and *NK3* (Kim and Nirenberg, 1989). *Nkx-2.5* and *Nkx-2.6* are novel genes, whereas *Nkx-2.3* is one of four mammalian genes already isolated (*Nkx-2.1*, *Nkx-2.2*, *Nkx-2.3*, *Nkx-2.4*) that are members of the same homeobox gene subfamily (Price et al., 1992). The thyroid transcription factor-1 gene (*TTF-1*, now called *Nkx-2.1*) was the first of these genes cloned and codes for a transcription regulator of the thyroid-specific genes, thyroglobulin and thyroid peroxidase (Guazzi et al., 1990; Mizuno et al., 1991). In addition to thyroid, *Nkx-2.1* is expressed in lung and forebrain (Lazzaro et al., 1991). *Nkx-2.2* is also expressed in the forebrain where *Nkx-2.1* and *Nkx-2.2* may cooperate in the delimitation of forebrain structures (Price et al., 1992). The expression patterns of *Nkx-2.3* and *Nkx-2.4* have not been reported.

In this paper, we present a detailed study of the structure and expression pattern of the *Nkx-2.5* gene. Like *NK4/msh-2*, *Nkx-2.5* is expressed in early heart progenitor cells and may function in commitment to the myocardial lineage. In support of this, *Nkx-2.5* is also expressed in primitive pharyngeal endoderm, the tissue thought to secrete the heart inducer (Jacobson and Sater, 1988).

MATERIALS AND METHODS

Recombinant DNA techniques

All recombinant DNA techniques were performed essentially as described (Sambrook et al., 1989). PCR was carried out using primers corresponding to the sense strand of the 5′ region (5′CGAAAGCCTCGCGTGCTCTT3′) and the antisense strand of the third helix (5′TAGCGCCGATTCTGGAACCA3′) of the *NK4/msh-2* homeobox. PCR reactions were performed with *Thermus aquaticus* (*Taq*) polymerase (Cetus) at varying concentrations of $MgCl_2$ (1.5-4.0 mM) under the following conditions: 40 cycles of 90 seconds duration at 94°C; 90 seconds annealing at 55°C and 180 seconds extension at 72°C. The adult mouse heart cDNA library (Stratagene) was screened at moderate stringency (42% formamide, 5× SSC, 42°C) with the *Nkx-2.5* PCR-derived homeobox fragment. Washing was at low stringency (2× SSC, 0.1% SDS, 42°C). To obtain *Nkx-2* genomic clones, a BALB/c mouse genomic library (constructed by Lynn Corcoran, The Walter and Eliza Hall Institute, Melbourne, Australia) was screened at moderate stringency (5× SSC, 50% formamide, 37°C) with the insert from the cNkx-2.5b cDNA clone. Filters were washed at low stringency (4× SSC, 0.1% SDS, RT).

RNase protection

RNase protections were performed as described (Krieg and Melton, 1988). All hybrids were digested in 8 μg/ml RNase A and 0.4 μg/ml RNase T1 (Sigma) at 37°C for 30 minutes. The *Nkx-2.5* probe was synthesized with T7 polymerase from a PCR-derived homeobox fragment (see text) after cloning into pBluescript SK+ and linearization with *Sal*I. The rat myosin heavy chain probe was transcribed with T7 RNA polymerase from the *Hin*dIII-linearized plasmid, MHC β150. The GAPDH probe was synthesized with T7 RNA polymerase from the *Sau*3A-linearized plasmid pGAPDH.

Cell lines and culture

All stromal cell lines were grown in DME, 10% newborn calf serum and 50 μM 2-mercaptoethanol. Bone marrow lines were as described (Li and Johnson, 1985; Pietrangeli et al., 1988; Nishikawa et al., 1988). Lymph node lines were the gift of A. Strasser (The Walter and Eliza Hall Institute, Melbourne, Australia) and were derived from lymph nodes of mice, which carried *bcl-2* and *myc* transgenes (Strasser et al., 1990), by infection with a retrovirus expressing SV40 large T antigen. The thymic stromal line B6T-ea was as described (Miyazawa et al., 1980). The SCL-5, SCL-13 and SCL-19 lines were derived by Gary Waanders (Ludwig Institute for Cancer Research, Lausanne, Switzerland) from the thymuses of 6- to 8-week-old AKR mice by passaging cells in bulk culture for 3-4 weeks, followed by two rounds of limit dilution cloning (unpublished data). The C3H $10T_{\frac{1}{2}}$ myogenic cell line was derived by 5-azacytidine treatment (Lassar et al., 1986). The C2C12 cell line was grown and differentiated as for the $10T_{\frac{1}{2}}$ myoblast line (Lassar et al., 1986). The D3 ES cell line was grown on STO fibroblasts in modified DME (Robertson, 1987) containing 15% FCS and 10^3 u/ml leukaemia inhibitory factor (LIF). For differentiation experiments, cells were weaned from STO feeders at high density in the presence of LIF. For differentiation in aggregates, cultures were plated into bacteriological dishes in growth media without LIF. For adherent differentiation, treatment was identical except that cultures were plated into gelatin-coated tissue culture dishes.

In situ hybridization

In situ hybridization to tissue sections using ^{35}S- or ^{33}P-radiolabelled cRNA probes was essentially as described (Green et al., 1992). Whole-mount in situ hybridization was as described (Dickson et al., 1993). Sense and antisense *Nkx-2.5* probes were synthesized with T7 or SP6 RNA polymerase (Promega) from a subclone of Nkx-2.5a covering nucleotides 820-1593 (Fig. 2). Sense probes in all cases showed background levels of hybridization in all tissues (data not shown).

RESULTS

Cloning of Nkx-2 genes

A single gene, *Nkx-2.5*, was found amongst 14 clones derived from PCR reactions using total mouse genomic DNA and oligonucleotides complementary to the 5′ region

Fig. 1. DNA sequence of the Nkx-2.5a cDNA clone with the predicted Nkx-2.5 amino acid sequence. The start and stop codons, and the homeobox are boxed. The amino-terminal homology is underlined (see Fig. 2B). The NK2-specific domain is doubly underlined (see Fig. 2C). The region between the homeodomain and the NK2-specific domain shared by Nkx2.5 and Nkx-2.6 is indicated with dashes (Fig. 2C). A possible polyadenylation signal and mRNA instability sequence are respectively overlined and underlined in bold.

and helix 3 of the *NK4/msh-2* homeobox (see discussion of nomenclature below). *Nkx-2.5* cDNA clones were selected from an adult mouse heart cDNA library (described below). Genomic clones encompassing *Nkx-2.5* were isolated by screening a mouse genomic library at moderate stringency with an *Nkx-2.5* cDNA probe. In the genomic library screen, additional cross-hybridizing clones were detected and several strongly hybridizing examples were isolated and characterized. DNA sequencing showed that one of these isolates encoded an intact homeodomain derived from a novel gene. On the basis of sequence comparisons (see below), this gene was named *Nkx-2.6*. The predicted amino acid sequence of another isolate corresponded to the previously published Nkx-2.3 homeodomain protein (Price et al., 1992) which is highly related to Nkx-2.5 and Nkx-2.6.

Nkx-2.5 cDNA clones

RNase protection analysis indicated *Nkx-2.5* expression in the embryonic and adult heart (see below). To isolate *Nkx-2.5* cDNA clones, a mouse adult heart cDNA library was screened with the PCR-derived homeobox probe. Four clones were obtained and the DNA sequence of one of them (Nkx-2.5a) is shown in Fig. 1. The Nkx-2.5a cDNA is 1596 nucleotides in length. Northern analysis of adult and foetal heart RNA showed a single transcript of approximately 1.6-1.7 kb in length (data not shown). We conclude that Nkx-2.5a is a near full-length cDNA. A potentially variant polyadenylylation sequence (AUUAAA) (Mavilio et al., 1986) was found at position 1567 of this clone, but no poly(A) tail occurred at the 3′ end. A consensus mRNA destabilization sequence (UAUUUAU) (Shaw and Kamen, 1986) was found in the 3′-untranslated region, suggesting that the steady-state level of *Nkx-2.5* mRNA is regulated by mRNA turnover.

Homeodomain sequences

The amino acid sequences of the Nkx-2.5 and -2.6 homeodomains and their closest evolutionary relatives are shown in Fig. 2A. Of the known *Drosophila* homeodomain proteins, Nkx-2.5/2.6 are most closely related within their homeodomains to NK2 (72% and 68%, respectively) and to NK3 and NK4 (63-67%). Among all known homeodomain proteins, Nkx-2.5/2.6 are most closely related to the mammalian proteins Nkx-2.1, Nkx-2.2, Nkx-2.3 and Nkx-2.4 (77-95%), which are themselves more related to NK2 than to any other *Drosophila* homeodomain protein (Price et al., 1992). All of the known NK2-related proteins except planarian Dth-1 and Dth-2 (Garcia-Fernandez et al., 1991) contain an amino acid motif not found in other homeobox proteins (see below). Based on these observations, we have followed the nomenclature of Price et al. (1992) and assigned the new genes to the *Nkx-2* series.

Amino acid groupings within the NK-related homeodomains (arrowed and shaded in Fig. 2A) not only support the close evolutionary relationship between the *NK2*, *NK3* and *NK4* genes, but also suggest functional divergence within this group. The tyrosine at position 54, which may influence DNA-binding site specificity (Otting et al., 1990; Wolberger et al., 1991), is unique among the *Drosophila* NK2, NK3 and NK4 proteins and their close relatives (Fig. 2A). However, the distribution of amino acids at positions 4, 38, 42 and 52 (shaded in Fig. 2A) suggest a phylogenetic and perhaps functional split within the group. It may be significant that, of these amino acids, those at positions 4, 42

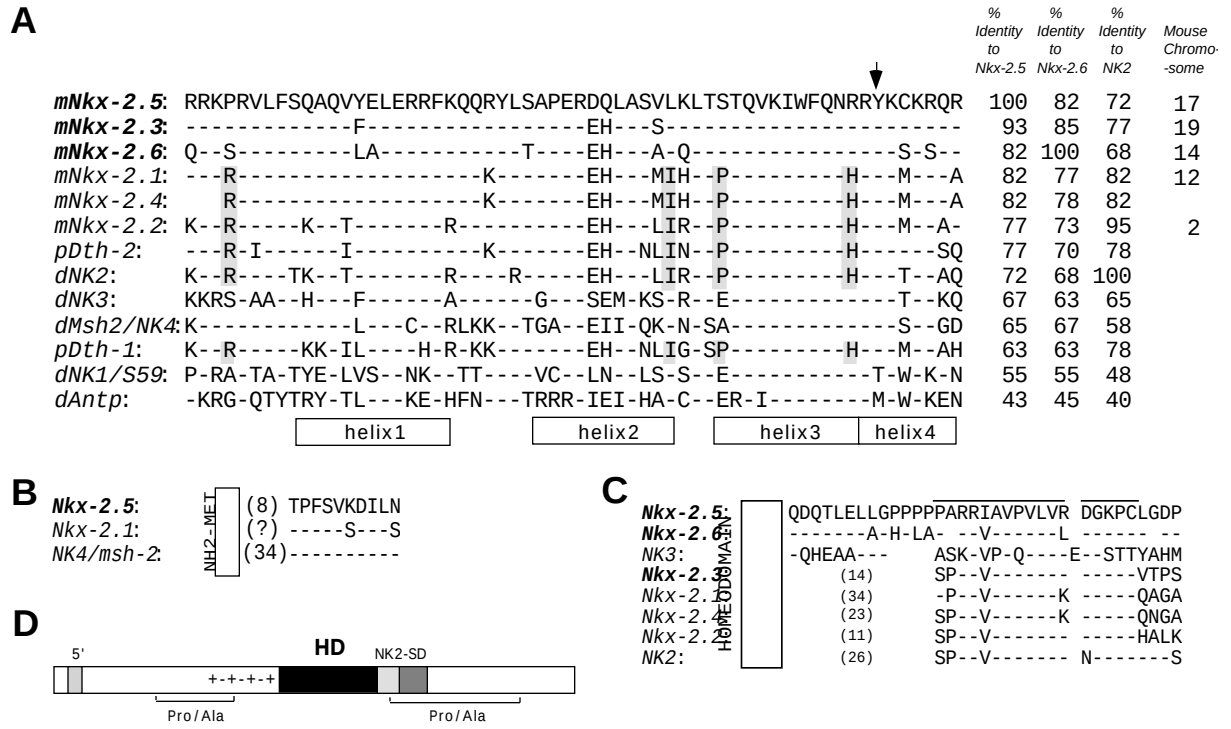

Fig. 2. (A) Comparison of the Nkx-2.5, Nkx-2.3 and Nkx-2.6 homeodomain sequences with those of closest evolutionary relatives and the antennapedia (*Antp*) sequence. Gene names are prefixed with a species abbreviation in lower case: m, mouse; d, *Drosophila*; p, planaria. All sequences are compared to the Nkx-2.5 homeodomain and identities are indicated by a dash. The 60 amino acid homeodomain is presented, as not enough published sequence data was available to include the 61st amino acid. The sequence of the Nkx-2.4 homeodomain is incomplete at its amino-terminus (Price et al., 1992). Predicted homeodomain helicies are indicated with boxes at the bottom of the figure. The percentage amino acid identity of each homeodomain to the Nkx-2.5, Nkx-2.6 and NK2 sequences is indicated on the right, along with the mouse chromosomal location of respective genes, where known (taken from Price et al., 1992 and H. Himmelbauer, N. Jenkins, N. Copeland, L. Silver and R. Harvey, unpublished data). The arrow indicates the Tyr unique to NK2, NK3 and NK4/msh-2 and their relatives. Shaded amino acids reveal possible phylogenetic subgroups within these NK-related homeobox genes (see text). (B) Amino-terminal decapeptide shared by the Nkx-2.5, Nkx-2.1 and NK4/msh-2 proteins. The N-terminal sequences of other NK2 relatives are not available. The distance in amino acids from the predicted initiator methionine (NH2-MET) of each protein is indicated in brackets. (C) The 17 amino acid NK2-specific homology domain, found carboxy-terminal to the homeodomain (overlined). The distance in amino acids from the 60 amino acid homeodomain (Fig. 2A) is indicated in brackets. An extended homology is seen in Nkx-2.5 and Nkx-2.6 and for these two proteins, all amino acids between the homeodomain and the NK2-specific domain are shown. A possible related sequence is also seen in NK3. (D) Schematic diagram depicting the positions of the homeodomain (HD), homology regions, charged domain (+-+-) and proline/alanine-rich regions (Pro/Ala) in the Nkx-2.5 protein sequence (boxed). 5′ indicates the homology region shown in Fig. 2B. NK2-SD represents the NK2-specific domain (dense stripes) and the Nkx-2.5/Nkx-2.6 extended homology (sparse stripes) shown in Fig. 2C.

and 52 lie within the N-terminal arm and helix 3 of the homeodomain, which may also be involved in DNA-binding site specificity (Lin and McGinnis, 1992). In the light of these groupings, *Nkx-2.3*, *Nkx-2.5* and *Nkx-2.6*, the genes described in this paper, may be more closely related to *NK3* and *NK4/msh-2*, than to *NK2*. This is an interesting speculation, since *Nkx-2.5* and *NK4/msh-2* are both expressed in the hearts of their respective species and may have homologous roles during heart development (see below and Discussion). Some support for this grouping was obtained from phylogenetic parsimony analysis, which showed that *Nkx-2.3*, *Nkx-2.5* and *Nkx-2.6* were more ancient than all other *NK2*-related genes and were as equally related to *NK4* as to *NK2* (data not shown).

Nkx-2.5 protein

The DNA sequence of Nkx-2.5a predicts a protein of 318 amino acids with a relative molecular mass of 34×10^3 (Fig. 1), consistent with the size of the protein produced by translation of synthetic *Nkx-2.5* mRNA in vitro (data not shown). The predicted initiator methionine codon is in a favourable context for translation initiation (Kozak, 1987). There are no other in-frame methionine codons upstream. Like several developmental regulatory genes from *Drosophila* and vertebrates, out-of-frame AUG codons are present upstream from the predicted initiator (at positions 14, 70 and 187 in Fig. 1). While we cannot presently assess whether these are involved in translational regulation, efficient translation of synthetic *Nkx-2.5* mRNA in rabbit reticulocyte lysates requires the removal of 5′ untranslated sequences (data not shown).

Two regions of the *Nkx-2.5* sequence are conserved in other homeobox proteins (Fig. 2B,C). The first, a decapeptide conserved in mouse Nkx-2.5 and Nkx-2.1, and *Drosophila* NK4/msh-2 (Fig. 2B), is located near the predicted amino-termini of these proteins. The second, pre-

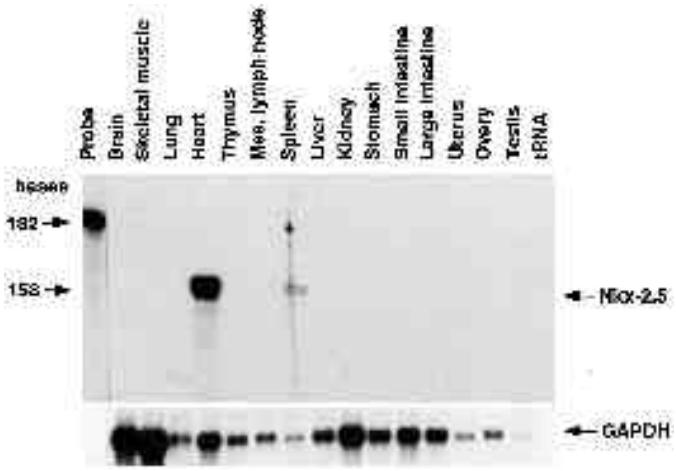

Fig. 3. RNase protection analysis of *Nkx-2.5* expression in adult mouse tissues. Sizes of the probe and protected fragments are shown on the left of the figure. 20 μg of total RNA was used for each protection. Note that the *Nkx-2.5* probe detects two closely migrating bands. This probe is synthesized from a DNA fragment derived by PCR using oligonucleotides specific for the *Drosophila NK4/msh-2* gene and so has mismatches at its ends relative the mouse *Nkx-2.5* sequence. The presence of intact RNA was confirmed by northern analysis using a *GAPDH* gene probe.

viously noted by Price et al. (1992) (Fig. 2C), is carboxy-terminal to the homeodomain and consists of a 17 amino acid motif found only in NK2 relatives. A more divergent sequence is present in NK3 (Fig. 2C). The NK2-specific domain contains a central cluster of invariant hydrophobic amino acids (VPVLV), which may have a role in protein-protein interaction. Mutagenesis of this domain in Nkx-2.1 does not influence its ability to bind to DNA in vitro (Guazzi et al., 1990).

The DNA sequence within and around the *Nkx-2.6* homeobox (not shown) predicts an amino acid sequence with 77/94 positions identical to Nkx-2.5 over a contiguous region that includes the homeodomain and the NK2-specific domain (Fig. 2C). The additional homology is not present in Nkx-2.3, the next most closely related protein (Fig. 2A).

Parts of the Nkx-2.5 protein, both amino- and carboxy-terminal to the homeodomain, are rich in proline and alanine (Fig. 2D). There is also a region of high charge density immediately N-terminal to the homeodomain (Fig. 2D). The function of these regions is not known, although they may be transcriptional activation or repression domains (Mitchell and Tijan, 1989; Han and Manley, 1993).

Distribution of *Nkx-2.5* mRNA in foetal and adult tissues

RNase protection has been used to examine the distribution of *Nkx-2.5* transcripts in embryonic and adult tissues and in cell lines (Fig. 3; summarized in Table 1). In adult tissues analysed, only the heart was strongly positive, while spleen and tongue showed a weak signal. In numerous tissue samples dissected from mid-gestation embryos, only the heart was strongly positive.

In the light of the expression in adult spleen, other haematopoietic organs and immortalized haematopoietic cell lines were examined for *Nkx-2.5* expression. Neither

Table 1. Summary of *Nkx-2.5* expression analyzed by RNase protection

Tissue type	Cell line	*Nkx-2.5* mRNA
Foetal heart		++++
Adult heart		+++
Adult spleen		+
Adult tongue		+
Adult stomach		−
Foetal brain		−
Neonatal brain (P1)		−
Postnatal brain (P5)		−
(P10)		−
(P12)		−
Adult brain		−
Fibroblasts	NIH3T3	+++
	Balb/c 3T3	−
	STO	−
Myoblasts	C3H 10T½	−
	C2C12	−
Myocytes	C3H 10T½	−
	C2C12	−
Foetal liver		−
Adult bone marrow		−
Bone marrow stroma	BAd	+
	PA6	−
	BMS	−
Adult thymus		−
Thymic stroma	SCL-19	+++
	B6T-ea	−
	SCL-5	−
	SCL-13	−
Adult lymph node		−
Lymph node stroma	TSL-1	+
	TSL-3	−

Relative levels of expression (+ to ++++) were estimated from autoradiograms of RNase protections using 20 μg of total or 1 μg of poly(A)$^+$ RNA. Cell lines are described in Materials and Methods.

mesenteric lymph nodes, thymus, foetal liver or bone marrow (Table 1), nor fifteen cell lines representing a variety of lineages and stages of haematopoietic development (data not shown) expressed the *Nkx-2.5* gene. These data suggest that the splenic stroma and not resident haematopoietic cells express *Nkx-2.5*. Several stromal cell lines derived from other haematopoietic organs were also tested for *Nkx-2.5* expression. While expression was strong in the SCL-19 (thymic) line and weak in BAd (bone marrow) and TSL-1 (lymph node) lines, no consistent pattern was observed indicating that *Nkx-2.5* expression is not common to stromal cell lines that support haematopoietic development. Expression was also seen in NIH3T3, but not Balb/c or STO fibroblasts (Table 1). No expression was found in adult skeletal muscle, or in C2C12 or C3H 10T½ myoblasts and myoctyes.

Nkx-2.1 and *Nkx-2.2*, two of the murine homeobox genes closely related to *Nkx-2.5* (Fig. 2A), are expressed in the brain (Price et al., 1992). We have likewise detected *Nkx-2.6* expression in neonatal brain (data not shown). However, *Nkx-2.5* transcripts were not found in foetal, neonatal or adult brain (Table 1).

In situ hybridization analysis of *Nkx-2.5* expression

Expression of *Nkx-2.5* in the myocardiogenic lineage

Expression of *Nkx-2.5* was examined during postimplantation development using in situ hybridization to tissue

sections and whole-mount embryos. *Nkx-2.5* transcripts were first detected in whole-mount embryos at early headfold (late primitive streak) stages. The pattern of hybridization is a symmetrically split crescent capping the extreme anterior and lateral parts of the embryo (Fig. 4A,D). Hybridization to sections at a slightly later stage indicated that both the mesodermal and endodermal components of the anterior splanchnopleura express *Nkx-2.5* (Figs 5E,F, 8).

Numerous studies have described the fate of the anterior splanchnic mesoderm to be myocardium (DeHaan, 1965; Viragh and Challice, 1977; Kaufman and Navaratnam, 1981; DeRuiter et al., 1992). During gastrulation, when the embryonic coelom forms within lateral mesoderm, the splanchnic mesoderm is squamous, but around 7.5 days post coitum (pc) it assumes the cuboidal morphology characteristic of myocardiogenic progenitor cells. Very soon after, when foregut closure begins and the anterior intestinal portal (AIP) becomes evident, a sparse layer of mesenchymal cells appears between the myocardiogenic progenitors and the subjacent endoderm. These are the progenitors of the endocardium, but it is not clear whether they are a product of the cuboidal myocardial layer itself, or migrate in from adjacent regions.

The time of *Nkx-2.5* activation with respect to cellular differentiation in the splanchnic mesoderm could not be

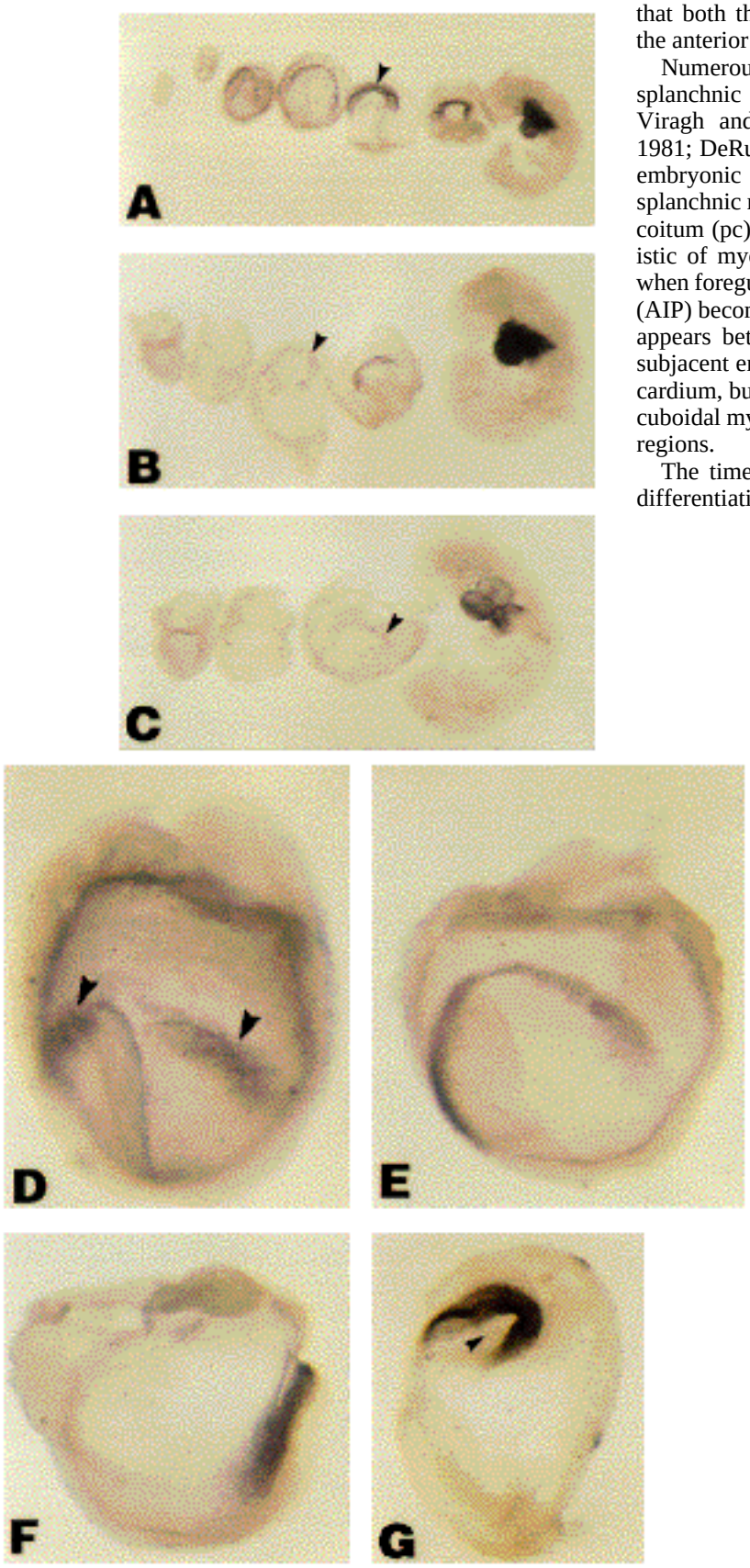

Fig. 4. In situ hybridization analysis of *Nkx-2.5* expression in whole embryos using digoxigenin-labelled probes. (A) Developmental time course of *Nkx-2.5* expression. From the left, stages are approximately: 6.0 days pc (egg cylinder), 7.0 days pc (primitive streak), 7.5 days pc (late primitive streak/early headfold), 7.8 days pc (headfold), 7.9 days pc, 8.0 days pc and 9.0 days pc. The anterior intestinal portal (AIP) is first evident in the 7.8 days pc embryo (arrowed) upon thickening of the *Nkx-2.5* expression pattern. (B) Developmental time course of α-cardiac actin gene expression. From the left, stages are approximately: 7.2 days pc (primitive streak), 7.6 days pc (early headfold), 7.9 days pc (headfold), 8.0 days pc and 10.0 days pc. The arrow indicates the weak but detectable actin expression at the time of AIP appearance. (C) Developmental time course of myosin heavy chain β gene expression. From the left, stages are approximately: 7.2 days pc (mid-primitive streak), 7.6 days pc (early headfold), 8.0 days pc and 9.0 days pc. The arrow indicates the first appearance of myosin gene expression when the AIP is clearly evident. (D) Enlargement of the 7.5 days pc embryo shown in A, indicating the first sign of *Nkx-2.5* expression in a split crescent of anterior/lateral splanchnopleural cells (arrowed). Embryo viewed from the anterior side. The apparent signal at the top edges of the embryos is artifactual probe trapping within the amniotic cavity. (E) Enlargement of the 7.8 days pc embryo shown in A, viewed from the anterior side. The split *Nkx-2.5* expression pattern has resolved into a complete crescent. The AIP is not yet visible. (F) Same embryo as in E, viewed from the lateral side to indicate the early extent of headfold formation at this stage. (G) Enlargement of the 8.0 days pc embryo shown in A, viewed from the anteriolateral side. The arrow indicates the AIP.

accurately determined after in situ hybridization. However, expression was detected two distinct stages before AIP formation (Fig. 4A) (DeRuiter et al., 1992) indicating that *Nkx-2.5* is activated around the time of appearance of the cuboidal progenitors. This is supported by a separate observation. Early cardiogenic progenitors are initially separated into bilaterally symmetrical cell groups by imposition of a wedge of yolk sac endoderm (DeRuiter et al., 1992). Split cardiac progenitor cell populations have also been described from fate mapping experiments in all vertebrates studies (DeHaan, 1965; Lawson et al., 1991). The *Nkx-2.5* expression pattern is also split in this manner (Fig. 4D), supporting our impression that initiation of expression occurs in the earliest recognisable cardiogenic precursors.

With further development and embryonic growth (Fig. 4A), the initial expression pattern resolves into a complete crescent, consistent with the known development of the promyocardium. When the AIP first appears, the crescent of expression broadens (Fig. 4G) and undergoes changes consistent with morphogenesis of the linear and S-shaped heart tube (Fig. 5A,B). Strong expression continues in myogenic cells of embryonic and foetal hearts (Figs 5C,D, 6A,B). At all stages, hybridization appears relatively even across both atrial and ventricular chambers (Fig. 6B). The bifurcated venous ports of the heart tube (the sinus venosae) are strongly positive (Fig. 5A), but no expression was found in the aorta, endocardium, or epicardium (Fig. 6B).

We have compared the expression of *Nkx-2.5* with that of the α-cardiac actin and myosin heavy chain-β genes, early markers of myogenic differentiation in the heart (Figs 4A,B,C, 5D) (Sassoon et al., 1988; Lyons et al., 1990). Activation of the α-cardiac actin gene was detected when the AIP first becomes evident and the myosin heavy chain-β gene a little later, when the AIP is more developed. Thus, *Nkx-2.5* expression is clearly seen before any molecular evidence for myogenic differentiation.

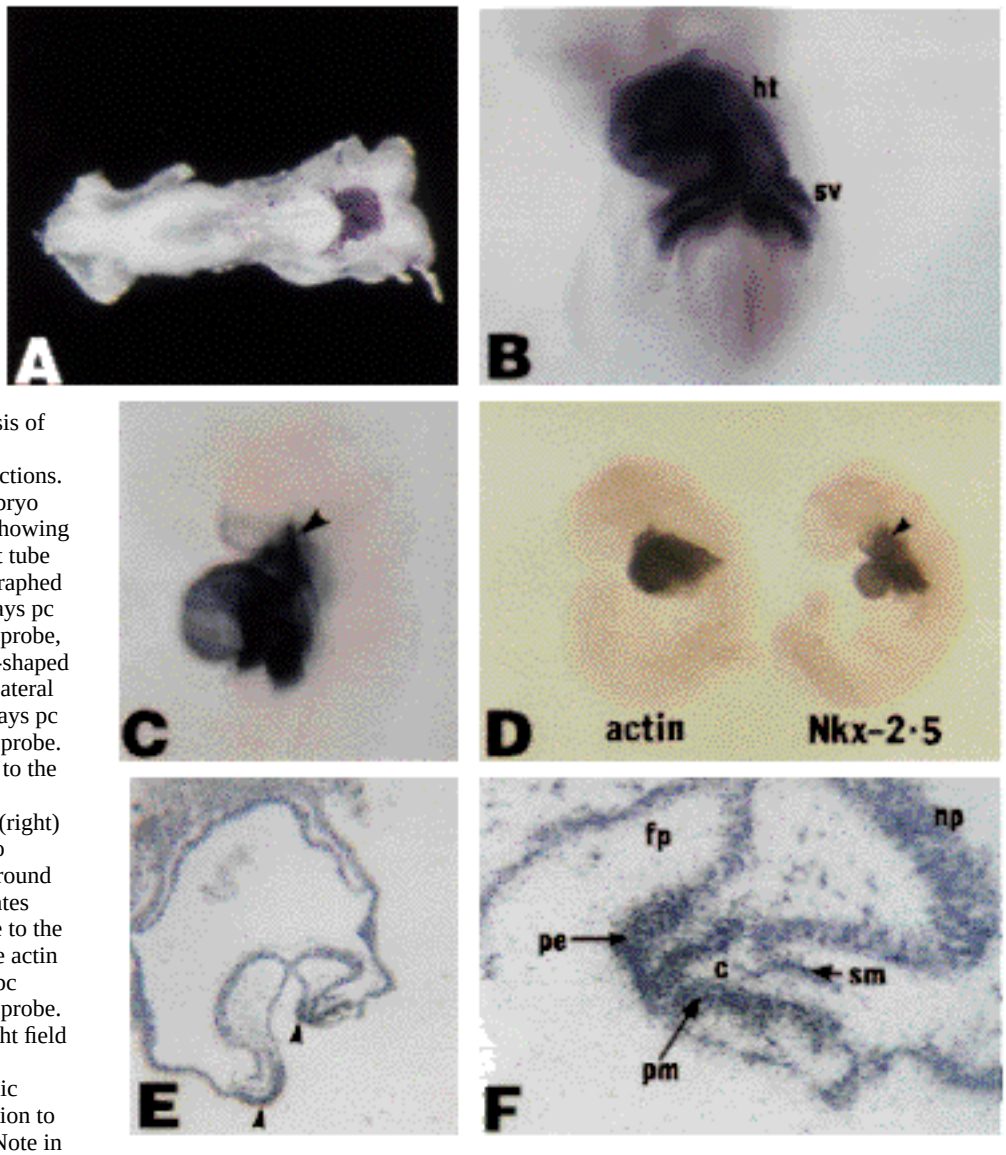

Fig. 5. In situ hybridization analysis of *Nkx-2.5* expression in whole mid-gestation embryos and in tissue sections. (A) Whole-mount 9.0 days pc embryo hybridized to the *Nkx-2.5* probe, showing a positive signal in the linear heart tube and sinus venosa. Embryo photographed in dark field. (B) Detail of a 9.5 days pc embryo hybridized to the *Nkx-2.5* probe, showing a positive signal in the S-shaped heart tube and sinus venosa. (C) Lateral view of the heart region of a 9.5 days pc embryo hybridized to the *Nkx-2.5* probe. The arrow indicates hybridization to the pharyngeal floor. (D) Comparison between hybridization of *Nkx-2.5* (right) and α-cardiac actin (left) probes to embryos of a comparable stage (around 9.0-9.5 days pc). The arrow indicates hybridization of the *Nkx-2.5* probe to the pharyngeal floor, not seen with the actin probe. (E) Section of an 8.0 days pc embryo hybridized to the *Nkx-2.5* probe. The section is photographed in light field and hybridization is indicated by accumulation of black photographic grains. Arrows indicate hybridization to the future pharyngeal endoderm. Note in the left side, hybridization to endoderm, but not mesoderm. (F) Enlargement of the right side of the embryo depicted in E showing hybridization to both cardiac mesoderm and future foregut endoderm. Abbreviations: c, embryonic coelom; fp, foregut pocket; ht, heart; np, neural plate; pe, primitive pharyngeal endoderm; pm, promyocardium; sv, sinus venosa; sm, somatopleura.

Nkx-2.5 expression in primitive pharyngeal endoderm and thyroid

At approximately 8 days pc, *Nkx-2.5* expression was seen in definitive endoderm as well as myocardiogenic progenitors (Figs 5E,F, 8). Hybridization to the endoderm was only found laterally, where in direct apposition to cardiogenic

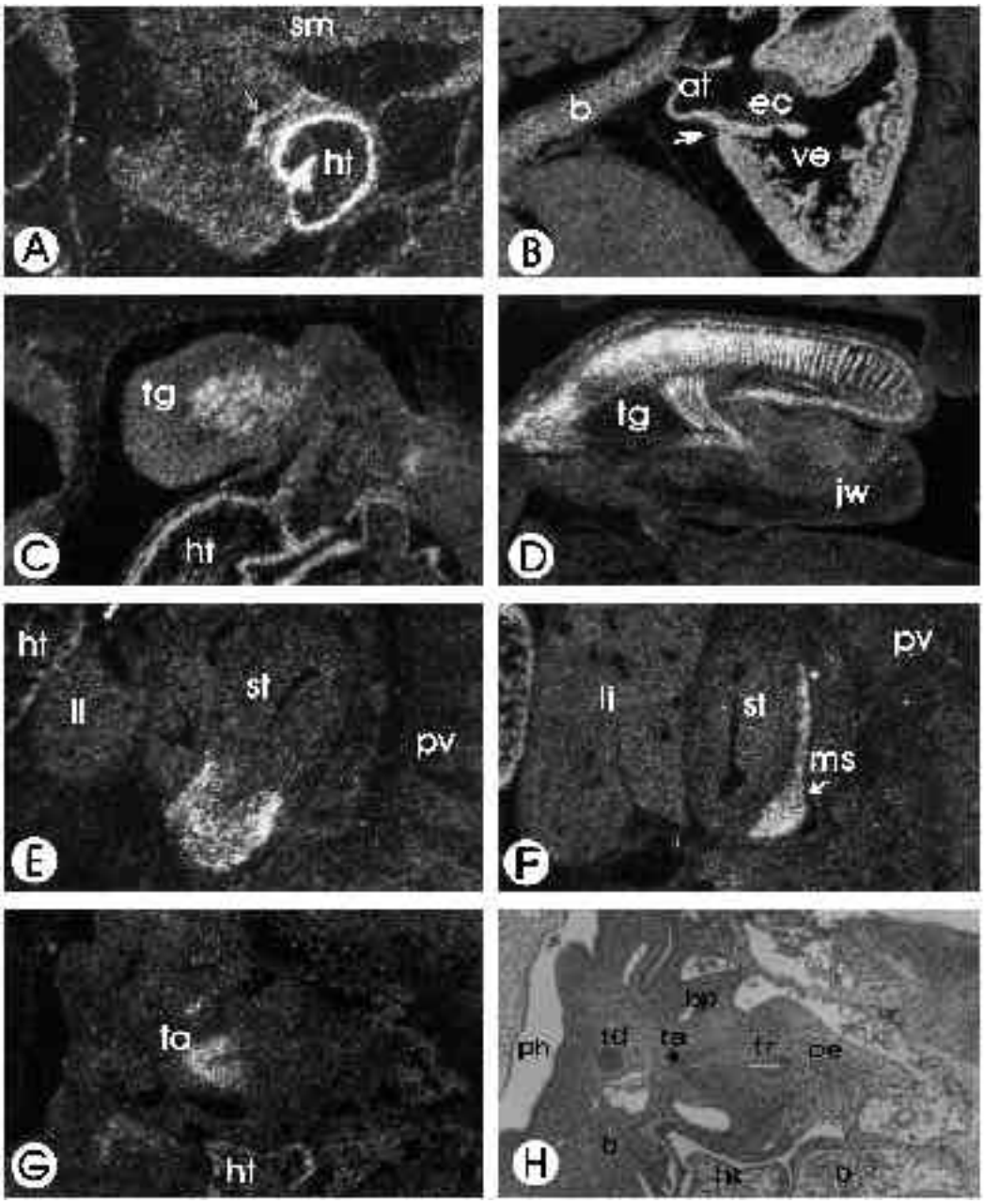

Fig. 6. In situ hybridization analysis of *Nkx-2.5* expression in heart, pharynx, tongue, stomach, spleen and thyroid. (A-G) Dark-field photomicrographs to reveal the *Nkx-2.5* hybridization signal; (H) a light-field photomicrograph of the section shown in G, stained with haemotoxylin and eosin to show tissue histology. (A) Section of an 8.5 days pc embryo showing hybridization to the developing heart tube and pharyngeal floor (arrowed). (B) Section of a 12.5 days pc embryos showing hybridization to the heart in both atrial and ventricular chambers. Arrow shows epicardial tissue that is negative for *Nkx-2.5* expression. Endocardial cushions are also negative. (C,D) Sections of 11.5 and 14.5 embryos, respectively, showing hybridization to lingual myoblasts and differentiated muscles. (E) Section through the stomach at 12.5 days pc showing hybridization to the distal end of stomach mesoderm. (F) Section through the 12.5 days pc stomach at a different level, showing hybridization to associated mesenteric tissue, probably corresponding to the spleen anlage (see text). (G,H) Adjacent sections through the branchial region of an 11.5 days pc embryo showing hybridization to the developing thyroid gland and heart. Hybridization is seen at the distal end of the thyroid diverticulum, in front of and in close association with the trachea and between the third branchial pouches. Note that the more proximal thryoid diverticulum is negative. Abbreviations: at, atrium; b, blood in which nucleated erythrocytes have reacted chemically with the photographic emulsion to give the impression of a positive hybridization signal; bp, branchial pouch; ec, endocardial cushion; ht, heart; jw, jaw; li, liver; ms, mesentery; oe, oesophagus; ph, pharynx; pv, prevertebrae; sm, somites; st, stomach; ta, thyroid anlage; tg, tongue; tr, trachea; ve, ventricle.

mesoderm (Fig. 5F). Thus the cardiogenic mesoderm and associated endoderm may receive the same regional signal to express the *Nkx-2.5* gene during development. Alternatively, the expression pattern may betray an inductive interaction between the two tissues, where *Nkx-2.5* is expressed in both inducing and responding tissues. This latter possibility is interesting since pharyngeal endoderm is regarded as the source of the heart inducer (Jacobson and Sater, 1988) (see Discussion).

Lateral endoderm in the cardiogenic region will form the floor of the pharynx after foregut closure. *Nkx-2.5* expression was found at 8.5-9.5 days pc in parts of the floor, but not the roof of the pharynx (Figs 5D, 6A, 8). At later stages, pharyngeal expression was not detected except in a midline structure that we identify as the thyroid primordium (Fig. 6G,H). The thyroid anlagen develops at the distal end of a diverticulum that descends from the floor of the pharynx beginning around 9 days pc (Rugh, 1968). We have seen *Nkx-2.5* expression in sections (Fig. 6I) and whole mounts (not shown) at a distal position along the thyroglossal duct at 11.5 to 12.5 days pc. Hybridization has not been detected at earlier or later stages, but the organ is small and may have been missed in the sections analyzed. Our identification of the hybridizing region as the developing thyroid rests on its position relative to the thyroglossal duct and its tight association with the front of the trachea (Fig. 6G,H), between the third and fourth branchial pouches and aortic arches. At both 11.5 and 12.5 days pc, the structure has a bilobal shape (Fig. 6G and data not shown) and, at 12.5 days pc, a follicular substructure, typical of thyroid acini, became evident (data not shown).

Expression of *Nkx-2.5* at other embryonic sites

In addition to the cardiogenic mesoderm and associated pharyngeal endoderm, we have detected *Nkx-2.5* expression in several other embryonic sites. First, *Nkx-2.5* expression was seen in tongue myoblasts and differentiated muscles (Fig. 6C,D). Lingual myoblasts arise in occipital somites, but no hybridization was observed to newly formed occipital somites themselves (Fig. 6A). Strong hybridization to lingual myoblasts was first seen at 11.5 days pc (Fig. 6C), although in 10.5 days pc embryos hybridized as whole mounts, three streaks of weakly positive cells could be seen in the occipital somite region and these may represent migrating myoblasts (data not shown). Expression is strongly evident in both intrinsic and extrinsic lingual muscles at 14.5 days pc (Fig. 6D) and in the adult tongue by RNase protection (Table 1). The epithelia and interspersed connective tissue do not express *Nkx-2.5*.

Also at 11.5 days pc, expression was observed in a small region of visceral mesoderm capping the distal end of the stomach (Fig. 6E). At 14.5 days pc, the expressing cells appear to ring the stomach (data not shown) and may therefore correspond to the future sphincter of pylorus. Expression was not detected in adult stomach by RNase protection (Table 1).

Nkx-2.5 expression was found in a population of cells embedded within mesenteric tissue next to the stomach and developing pancreas, beginning at 11.5 days pc (Fig. 6F). This position is consistent with that of the developing spleen (Rugh, 1968) and *Nkx-2.5* is also expressed in adult splenic stroma (see above). In some sections, the signal in the stomach and spleen anlagen appeared to be linked (data not shown), perhaps reflecting a common origin for these tissues.

Nkx-2.5 expression in ES cell-derived embryoid bodies

In vitro differentiation of embryonic stem (ES) cells into embryoid bodies containing spontaneously contracting cardiac muscle (Doetschman et al., 1985; Sanchez et al.,

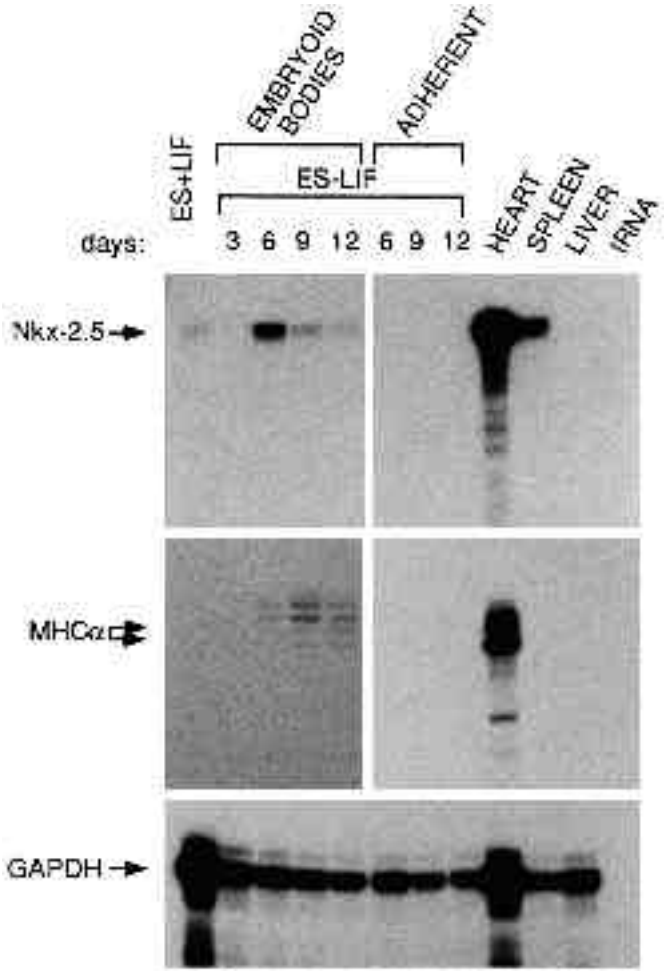

Fig. 7. RNase protection analysis of *Nkx-2.5* expression in ES cells and developing embryoid bodies. Expression is compared in undifferentiated ES cells passaged in leukaemia inhibitory factor (ES+LIF) with a time course of cells induced to differentiate by LIF withdrawal (ES-LIF), both in suspension culture (EMBRYOID BODIES) and in adherent culture (ADHERENT). Time of differentiation is indicated in days. Control RNAs (adult heart, spleen, liver and tRNA) are also analyzed. Note *Nkx-2.5* expression only in undifferentiated ES cells and embryoid bodies. A myosin heavy chain (MHC) probe was used to analyse muscle differentiation during embryoid body development. A rat MHC β probe was used that also detects MHC α transcripts (arrowed). Multiplicity of bands is presumably due to base mismatches between the rat probe and mouse target RNA. Although MHC β transcripts (top bands) are expressed in embryonic cardiac muscle, MHC α transcripts are heart-specific and the more definitive marker of heart myogenesis in this system. A GAPDH probe has been used to ascertain the presence and integrity of RNA.

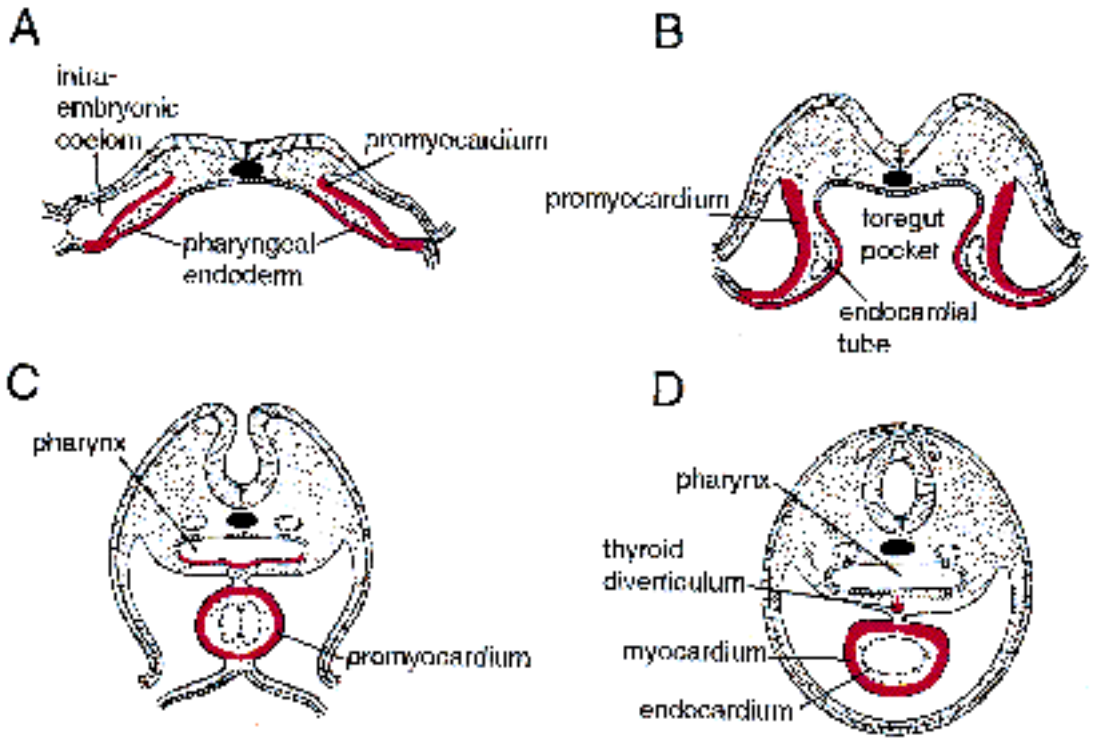

Fig. 8. Schematic representation of *Nkx-2.5* expression in the developing cardiac mesoderm, pharyngeal endoderm and developing thyroid. A-D correspond to embryos at approximately 7.8, 8.0, 8.5 and 11.5 days of development, respectively. The distribution of *Nkx-2.5* transcripts is indicated in red.

1991), represents one of few in vitro models of myocardial development. To examine whether *Nkx-2.5* is expressed in cardiac muscle induced in vitro, we assayed expression in a time course of differentiating D3 ES cells (Doetschman et al., 1985). The D3 ES cells are maintained in this laboratory on a STO fibroblast feeder layer in the presence of the differentiation inhibitor, LIF (leukaemia inhibitory factor). Before differentiation, ES cells are passaged without feeder cells in the presence of LIF. Differentiation is then activated by LIF withdrawal and cells are allowed to form free-floating aggregates. 10-40% of these aggregates form rhythmically beating foci, with contractions beginning 8-10 days after removal of LIF. Samples were processed at three day intervals for RNase protection analysis using probes specific for *Nkx-2.5* and *GAPDH*, and a myosin heavy chain probe that detects both α (*MHC*α) and β (*MHC*β) isoform mRNAs (Fig. 7).

In undifferentiated D3 ES cells, *Nkx-2.5* is weakly expressed, but this expression diminishes to undetectable levels by 3 days of differentiation. Subsequently, transcript levels peak at 6 days, then decrease several fold thereafter. ES cells that were prevented from forming free aggregates by allowing them to adhere to gelatinized culture plates, did not produce beating muscle foci or express *Nkx-2.5* or *MHC* transcripts at detectable levels. This indicates that patterning within embryoid bodies is a prerequisite for formation of myocardial cells.

Nkx-2.5 expression is detected during in vitro differentiation before *MHC*-α transcripts, the predominant *MHC* mRNA isoform in mouse adult heart (Fig. 7) and an early marker of cardiac myogenesis in this system (Sanchez et al., 1991). *Nkx-2.5* expression also appears at least 2 days before contractile activity is first observed. Although it is not known to what extent cardiomyogenesis in embryoid bodies models that in normal development, the *Nkx-2.5* gene is expressed with similar developmental kinetics in both systems. These data support the idea that *Nkx-2.5* expression plays an early and important role in cardiac myogenesis.

DISCUSSION

We have isolated three murine homeobox genes, *Nkx-2.3*, *Nkx-2.5* and *Nkx-2.6*, that are related to the *Drosophila NK* genes, as well as to the mammalian genes *Nkx-2.1*, *Nkx-2.2* and *Nkx-2.4*. The *Nkx-2.3* gene isolated in this screen has been described independently by Price et al. (1992).

Throughout development, the *Nkx-2.5* gene is expressed in the heart, pharynx, thyroid, tongue, stomach and spleen. The first and most striking feature of the expression pattern is the presence of *Nkx-2.5* transcripts in early myocardiogenic progenitor cells. It is difficult to determine the precise time of onset of *Nkx-2.5* expression because of variability in the developmental stage of embryos within the same litter. Anatomical landmarks are more relevant, although individual embryonic structures, eg. heart and somites, do not necessarily develop synchronously (Kaufman and Navaratnam, 1981). Our estimation of the time of onset of expression (around 7.5 days pc) places it close to the appearance of the cuboidal cardiogenic epithelia at late primitive streak/early headfold stages. Expression predates the appearance of myogenic differentiation markers by several hours. During

differentiation and morphogenesis, expression is found across the whole myocardium of the heart, without obvious differences between atrial or ventricular compartments. This pattern persists thoroughout embryonic and foetal development and transcripts can also be detected in the adult mouse heart. *Nkx-2.5* expression is also induced in ES-cell-derived embryoid bodies that form beating myocardial tissue.

These expression characteristics suggest that the *Nkx-2.5* gene acts in commitment to and/or differentiation of the myocardial lineage. Analysis of *Nkx-2.5* expression and function may therefore provide an early entry point for experimental dissection of the molecular basis of cardiac myogenesis. This is a particularly intriguing problem, since the myogenic factors that occupy nodal positions in commitment to and differentiation of skeletal muscle (Weintraub et al., 1991) are not expressed in mammalian cardiac muscle, despite the fact that cardiac and skeletal muscle coexpress many genes. Different HLH factors may act in this capacity in cardiac muscle, since functional HLH-binding sites are found within cardiac-expressed gene promoters (Sartorelli et al., 1992) and HLH proteins distinct from the known myogenic factors can be immunodetected in early chick hearts (Litvin et al., 1993). The regulatory relationship between the Nkx-2.5 protein and these implied heart HLH factors warrants investigation.

Nkx-2.5 is also expressed in the future pharyngeal floor endoderm (Fig. 8), strengthening the idea that this gene acts early in myocardial commitment or differentiation. Pharyngeal endoderm is regarded in the classical embryology literature as the source of the cardiac inducer (Jacobson and Sater, 1988). Most studies addressing this issue have been performed in amphibia, but the general conclusions are supported by experiments on chick embryos and observation of human endodermal defects (Jacobson and Sater, 1988). *Nkx-2.5* expression is found in lateral primitive foregut endoderm, where it lies in apposition to promyocardium, and in the floor of the pharynx after foregut closure. This pattern suggests that the mesodermal and endodermal tissues that express *Nkx-2.5* are in communication. Induction of cardiac mesoderm by endoderm (Jacobson and Sater, 1988) and of hepatic endoderm by cardiac mesoderm (Le Douarin, 1975), independently indicates that there is reciprocal communication between these tissues, and *Nkx-2.5* may be a mediator of this communication. A similar situation occurs in *Xenopus*, where primitive endoderm of the blastula, the source of mesoderm-inducing factors, expresses the homeobox gene *Mix1,* as well as induces its expression in ectoderm during mesoderm formation (Rosa, 1989).

An alternative interpretation of the early *Nkx-2.5* expression pattern is that it reflects a prior regionalization within the presumptive endomesoderm. Much of the definitive endoderm, in addition to mesoderm, is derived from the primitive ectoderm during gastrulation (Tam and Beddington, 1992). Primitive endoderm is displaced into the extraembryonic region by the definitive endoderm, which either moves through the steak with the mesoderm, or delaminates from the ectoderm directly. Little is known about the cellular movements and mesoderm/endoderm neighbour relationships during this process. However, it seems possible that the observed *Nkx-2.5* expression pattern could result from a regionalizing signal in the ectoderm that is propagated in a tight spatial relationship to both its mesodermal and endodermal derivatives during gastrulation.

Members of the TGFβ family of growth factors are likely cardiac-inducer molecules. TGFβ1 can induce the formation of partially differentiated cardiomyocytes from rabbit cardiac fibroblasts (Eghbali et al., 1991) and beating heart structures in axolotl mesodermal explants cultured without endoderm (Muslin and Williams, 1991). TGFβ1 and β2 mRNAs or protein have been localized to the vicinity of the developing murine heart. TGFβ1 mRNA is limited to primitive cardiac endothelial cells and the endocardium (Akhurst et al., 1990). TGFβ2 mRNA, on the other hand, is expressed in tissues strikingly similar to those expressing *Nkx-2.5*, including paired anterior/lateral regions of late primitive streak embryos, promyocardium and associated primitive endoderm, pharyngeal floor itself and the thyroid diverticulum (Dickson et al., 1993). Expression is also seen in somatic mesoderm and the septum transversum where *Nkx-2.5* transcripts are not detected. Nevertheless, these expression data show a strong correlation between TGFβ2 and *Nkx-2.5* expression in the cardiogenic region and *Nkx-2.5* mRNA may be an early product of TGFβ2 induction in the region of the developing heart.

Nkx-2.5 is also expressed in the developing thyroid, a derivative of the pharyngeal floor (Fig. 8). *Nkx-2.1*, a close relative of *Nkx-2.5*, is also expressed during thyroid development and regulates the expression of the thyroid-specific genes thyroglobulin and thyroid peroxidase (Guazzi et al., 1990; Mizuno et al., 1991). Since *Nkx-2.1* is expressed some days before thyroid differentiation, it could function in thyroid commitment (Lazzaro et al., 1991); however, its expression has only been reported in the thyroid diverticulum and not generally in the floor of the pharynx. Thus, *Nkx-2.5* is expressed earlier in thyroid development and may have a separate role. Consistent with this the Nkx-2.5 protein can bind in vitro to promoter fragments of the thyroglobulin gene, but with an affinity distinctly lower than that of Nkx-2.1 (data not shown). In the light of the intimate association between the mesodermal and endodermal expression domains of *Nkx-2.5*, it is an interesting possibility that the thyroid, like the liver, is dependent for its development upon communication with cardiac mesoderm.

Amongst known *Drosophila* homeodomain proteins, the Nkx-2.5 homeodomain has the highest degree of amino acid identity with NK2. However, parsimony analysis and particular amino acid groupings within the homeodomains of the NK2-related subfamily, indicate that the *Nkx-2.5* gene may be at least as closely related to *Drosophila NK4/msh-2*, the gene from which we designed the PCR oligonucleotides used in the mouse screen. It is interesting that *NK4/msh-2* is expressed in the *Drosophila* heart and is required for the formation of heart muscle progenitors (Bodmer et al., 1990), but, since it is not known whether insect and mammalian hearts are homologous, it remains to be seen whether this is a chance or significant observation. The phylogenetic uncertainties notwithstanding, we propose that *Nkx-2.5* participates in commitment to, or differentiation of, the mammalian myocadial lineage. Gene disruption experiments in mice should make it possible to address this issue.

Finally, *Nkx-2.5* is expressed in lingual myoblasts and differentiated muscles, but not in other skeletal muscles. In situ hybridization suggests that expression in lingual myoblasts is initiated during their migration from the occipital somite region and is well established by the time myoblasts populate the tongue. Skeletal myofibre diversity is, in part, attributable to the genesis of distinct skeletal myoblast lineages during development (Stockdale, 1992; Smith et al., 1993) and the need to understand the developmental basis of this diversity has recently been highlighted (Stockdale, 1992). The experiments presented in this paper demonstrate that lingual myoblasts can be distinguished from all other skeletal myoblasts on the basis of *Nkx-2.5* expression. The developmental influences that establish this apparent diversity can now be addressed.

We thank K. Davis and T. Huebert for excellent technical assistance. We are also indebted to Jane Andrews for subclone mapping, Bev Faulkner-Jones for technical advice, Ivor Mason for help with whole-mount in situ hybridization, Lindsay Williams for the D3 ES cells and ES cell RNA samples, John Allen for haematopoietic cell line RNAs, Andreas Strasser for stromal cell lines, George Muscat for C2C12 cells, Patrick Tam and Felix Beck for examining in situ hybridization slides, Margaret Buckingham for the MHC probe, Edna Hardeman for the cardiac actin probe and Matthijs Smith and Doug Hilton for critique of the manuscript. This research was supported by funds from The National Heart Foundation of Australia, the National Health and Medical Research Council (Australia) and The National Institute of Child Health and Human Development, DHHS (USA) (HD26024-01). T. J. L. holds a CSFP scholarship from the Australian government, L. P. holds a National Heart Foundation of Australia Postgraduate Science Research Scholarship and R. H. holds an Australian Reseach Council Research Fellowship.

REFERENCES

Akhurst, R. J., Lehnert, S. A., Faissner, A. and Duffie, E. (1990). TGF beta in murine morphogenetic processes: the early embryo and cardiogenesis. *Development* **108,** 645-656.

Awgulewitsch, A. and Jacobs, D. (1992). Deformed autoregulatory element from *Drosophila* functions in a conserved manner in transgenic mice. *Nature* **358,** 341-344.

Bodmer, R., Jan, L. Y. and Jan, Y. N. (1990). A new homeobox-containing gene, *msh-2*, is transiently expressed early during mesoderm formation in *Drosophila*. *Development* **110,** 661-669.

Chen, L., Krause, M., Draper, B., Weintraub, H. and Fire, A. (1992). Body-wall muscle formation in *Caenorhabditis elegans* embryos that lack the MyoD homolog *hlh-1*. *Science* **256,** 240-243.

DeHaan, R. L. (1965). In *Morphogenesis of the Vertebrate Heart* (ed. R. L. DeHaan and H. Ursprung), pp. 377-419. New York: Holt, Rinehart & Winston.

Dekel, I., Magal, Y., Pearson-White, S., Emerson, C. P. and Shani, M. (1992). Conditional conversion of ES cells to skeletal muscle by an exogenous MyoD1 gene. *New Biol.* **4,** 217-224.

DeRuiter, M. C., Poelmann, R. E., VanderPlas-de Vries, I., Mentink, M. M. T. and Gittenberger-de Groot, A. C. (1992). The development of the myocardium and endocardium in mouse embryos. *Anat. Embryol.* **185,** 461-473.

Dickson, M. C., Slanger, H. G., Duffie, E., Mummery, C. L. and Akhurst, R. J. (1993). RNA and protein localisations of TGFβ2 in the early mouse embryo suggest an involvement in cardiac development. *Development* **117,** 625-639.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morph.* **87,** 27-45.

Dressler, G. R. and Douglass, E. C. (1992). Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor. *Proc. Natl. Acad. Sci. USA* **89,** 1179-1183.

Eghbali, M., Tomek, R., Woods, C. and Bhambi, B. (1991). Cardiac fibroblasts are predisposed to convert into myocyte phenotype: Specific effects of transforming growth factor β. *Proc. Natl. Acad. Sci. USA* **88,** 795-799.

Garcia-Fernandez, J., Baguna, J. and Salo, E. (1991). Planarian homeobox genes: Cloning, sequence analysis, and expression. *Proc. Natl. Acad. Sci. USA* **88,** 7338-7342.

Guazzi, S., Price, M., De Felice, M., Damante, G., Mattei, M.-G. and Di Lauro, R. (1990). Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *EMBO J.* **9,** 3631-3639.

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

Hatano, M., Roberts, C. W. M., Minden, M., Crist, W. M. and Korsmeyer, J. J. (1991). Deregulation of a homeobox gene, HOX11, by the t(10:14) in T cell leukemia. *Science* **253,** 79-82.

Hayashi, S. and Scott, M. (1990). What determines the specificity of action of *Drosophila* homeodomain proteins. *Cell* **63,** 883-894.

Hunt, P. and Krumlauf, R. (1991). Deciphering the Hox code: clues to patterning branchial regions of the head. *Cell* **66,** 1075-1078.

Ingraham, H. A., Chen, R., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E., Lin, C. R., Simmons, D. M., Swanson, L. and Rosenfeld, M. G. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell* **55,** 519-529.

Jacobson, A. G. and Sater, A. K. (1988). Features of embryonic induction. *Development* **104,** 341-359.

Jegalian, B. G. and De Robertis, E. (1992). Homeotic transformations in the mouse induced by overexpression of a human *Hox3.3* transgene. *Cell* **71,** 901-910.

Kamps, M. P., Murre, X.-M., Sun, D. and Baltimore, D. (1990). A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell* **60,** 547-555.

Kaufman, M. H. and Navaratnam, V. (1981). Early differentiation of the heart in mouse embryos. *J. Anat.* **133,** 235-246.

Kenyon, C. and Wang, B. (1991). A cluster of *Antennapedia*-class homeobox genes in a non-segmented animal. *Science* **253,** 516-517.

Kessel, M., Balling, R. and Gruss, P. (1990). Variations of cervical vertebrae after expression of a *Hox-1.1* transgene in mice. *Cell* **61,** 301-308.

Kim, Y. and Nirenberg, M. (1989). *Drosophila* NK-homeobox genes. *Proc. Natl. Acad. Sci. USA* **86,** 7716-7720.

Kozak, M. (1987). An analysis of 5′ untranslated sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15,** 8125-8148.

Krieg, P. A. and Melton, D. A. (1988). In vitro synthesis with SP6 RNA polymerase. *Methods Enzymol.* **155,** 397-415.

Lassar, A. B., Paterson, B. M. and Weintraub, H. (1986). Transfection of a cDNA locus that mediates the conversion of 10T½ fibroblasts to myoblasts. *Cell* **47,** 649-656.

Lawson, K. A., Meneses, J. J. and Pederson, R. A. (1991). Clonal analysis of epiblast during germ layer formation in the mouse embryo. *Development* **113,** 891-911.

Lazzaro, D., Price, M., Felice, M. and Di Lauro, R. (1991). The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. *Development* **113,** 1093-1104.

Le Douarin, N. M. (1975). An experimental analysis of liver development. *Medical Biology* **53,** 427-455.

Le Mouellic, H., Lallemand, Y. and Brulet, P. (1992). Homeosis in the mouse induced by a null mutation in the *Hox-3.1* gene. *Cell* **69,** 251-264.

Li, C. L. and Johnson, G. R. (1985). Stimulation of multipotential, erythroid and other murine haematopoietic progenitor cells by adherent cell lines in the absence of detectable multi-CSF (IL-3). *Nature* **316,** 633-636.

Lin, C., Lin, S.-C., Chang, C.-P. and Rosenfeld, M. G. (1992). Pit-1-dependent expression of the receptor for growth hormone releasing factor mediates pituitary cell growth. *Nature* **360,** 765-767.

Lin, L. and McGinnis, W. (1992). Mapping the functional specificity in the Dfd and Ubx homeo domains. *Genes Dev.* **6,** 1071-1081.

Litvin, J., Montgomery, M. O., Goldhamer, D. J., Emerson, C. P. and Bader, D. M. (1993). Identification of DNA-binding protein(s) in the developing heart. *Dev. Biol.* **156,** 409-417.

Lyons, G. E., Schiaffino, S., Sassoon, D., Barton, P. and Buckingham, M. (1990). Developmental regulation of myosin gene expression in mouse cardiac cells. *J. Cell Biol.* **111,** 2427-2436.

Malicki, J., Cianetti, L. C., Peschle, C. and McGinnis, W. (1992). A human HOX4B regulatory element provides head-specific expression in *Drosophila* embryos. *Nature* **358,** 345-347.

Malicki, J., Schughart, K. and McGinnin, W. (1990). Mouse *Hox 2.2* specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell* **63,** 961-967.

Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, C. and Boncinelli, E. (1986). Differential and stage-related expression in embryonic tissues of a new human homeobox gene. *Nature* **324,** 664-668.

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

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

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

Miyazawa, Y., Sato, C., Hiai, H., Nishi, Y. and Matsuyama, M. (1980). Establishment of a reticuloepithelial-like cell line from mouse thymuses and its feeder capacity for the growth of bone marrow cells. *Cell Str. Funct.* **5,** 305-314.

Mizuno, K., Gonzalez, F. J. and Kimura, S. (1991). Thyroid-specific enhancer-binding protein (T/EBP): cDNA cloning, functional characterization, and structural identity with thyroid transcription factor TTF-1. *Mol. Cell. Biol.* **11,** 4927-4933.

Muslin, A. J. and Williams, L. T. (1991). Well-defined growth factors promote cardiac development in axolotl mesodermal explants. *Development* **112,** 1095-1101.

Nishikawa, S., Ogawa, M., Nishikawa, S., Kunisada, T. and Kodama, H. (1988). B lymphocyte stromal cell clone: stromal cell clones acting on different stages of B cell differentiation. *Eur. J. Immunol.* **18,** 1767-1771.

Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J. and Wuthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* **9,** 3085-3092.

Pietrangeli, C. E., Hayashi, S. and Kincade, P. W. (1988). Stromal cell lines which support lymphocyte growth: characterization, sensitivity to radiation and responsiveness to growth factors. *Eur. J. Immunol.* **18,** 863-872.

Pollock, R. A., Jay, G. and Bieberich, C. J. (1992). Altering the boundaries of *Hox3.1* expression: evidence for antipodal gene regulation. *Cell* **71,** 911-923.

Price, M., Lazzaro, D., Pohl, T., Mattei, M.-G., Ruther, U., Olivo, J.-C., Duboule, D. and Di Lauro, R. (1992). Regional expression of the homeobox gene *Nkx-2.2* in the developing mammalian forebrain. *Neuron* **8,** 241-255.

Robertson, E. J. (1987). *Teratocarcinomas and Embryonic Stem Cells: a Practical Approach*. Practical approach series. Oxford: IRL Press Limited.

Rosa, F. M. (1989). *Mix1*: a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of Xenopus embryos. *Cell* **57,** 965-974.

Rugh, R. (1968). *The Mouse: its Reproduction and Development*. Minneapolis: Burgess Publishing Company.

Ruiz i Altaba, A. and Melton, D. A. (1990). Axial patterning and the establishment of polarity in the frog embryo. *Trends Genet.* **6,** 57-64.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, Second Edition.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Sanchez, A., Jones, W. K., Gulick, J., Doetschman, T. and Robbins, J. (1991). Myosin heavy chain gene expression in mouse embryoid bodies. *J. Biol. Chem.* **266,** 22419-22426.

Sartorelli, V., Hong, N. A., Bishopric, N. H. and Kedes, L. (1992). Myocardial activation of the human cardiac alpha-actin promoter by helix-loop-helix proteins. *Proc. Natl. Acad. Sci. USA* **89,** 4047-4051.

Sassoon, D. A. (1993). Myogenic regulatory factors: dissecting their role and regulation during vertebrate embryogenesis. *Dev. Biol.* **156,** 11-23.

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

Schummer, M., Scheurlen, I., Schaller, C. and Galliot, B. (1992). HOM/HOX homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration. *EMBO J.* **11,** 1815-1823.

Scott, M. P., Tamkun, J. W. and Hartzell, G. W. (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta* **989,** 25-48.

Shashikant, C. S., Utset, M. F., Violette, S. M., Wise, T. L., Einat, P., Einat, M., Pendleton, J. W., Schughart, K. and Ruddle, F. H. (1991). Homeobox genes in mouse development. *Eukaryotic Gene Expression* **1,** 207-245.

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

Smith, T. H., Block, N. E., Rhodes, S. J., Konieczny, S. F. and Boon Miller, J. (1993). A unique pattern of expression of the four muscle regulatory factor proteins distinguishes somitic from embryonic, fetal and newborn mouse myogenic cells. *Development* **117,** 1125-1133.

Song, K. N., Wang, Y. Q. and Sassoon, D. (1992). Expression of *Hox-7.1* in myoblasts inhibits terminal differentiation and induces cell transformation. *Nature* **360,** 477-481.

Stockdale, F. E. (1992). Myogenic cell lineages. *Dev. Biol.* **154,** 284-298.

Strasser, A., Harris, A. W., Bath, M. L. and Cory, S. (1990). Novel primitive lymphoid tumours induced in transgenic mice by cooperation between *myc* and *bcl-2. Nature* **348,** 331-333.

Tabin, C. J. (1992). Why we have (only) five fingers per hand: Hox genes and the evolution of paired limbs. *Development* **116,** 289-296.

Tam, P. P. L. and Beddington, R. A. (1992). In *Establishment and organization of germ layers in the gastrulating mouse embryo.* Ciba Foundation Symposium (ed. D. J. Chadwick J. Marsh), pp. 27-40. New York: John Wiley and Sons, Inc.

Tronche, F. and Yaniv, M. (1992). HNF1, a homeoprotein member of the hepatic transcription regulatory network. *Bioessays* **14,** 579-587.

Viragh, S. and Challice, C. E. (1977). Origin and differentiation of cardiac muscle cells in the mouse. *J. Ultrastruct. Res.* **42,** 1-24.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T. K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y. and Lassar, A. (1991). The *myoD* gene family: nodal point during specification of the muscle cell lineage. *Science* **251,** 761-766.

Weintraub, H., Tapscott, R. L., Davis, R. L., Thayer, M. J., Adam, M. A., Lassar, A. B. and Miller, A. D. (1989). Activation of muscle specific genes in pigment, nerve, fat, liver and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci. USA* **86,** 5434-5438.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. and Pabo, C. O. (1991). Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67,** 517-528.

(Accepted 22 July 1993)

Note added in proof

The *Drosophila NK4/msh-2* gene has been renamed *tinman* (see Bodmer (1993) *Development* **118**, 719-729).