*Mechanisms of Development*, 33 (1991) 27–38
Elsevier Scientific Publishers Ireland, Ltd.



MOD 00003

# The murine paired box gene, *Pax7*, is expressed specifically during the development of the nervous and muscular system

Birgit Jostes, Claudia Walther and Peter Gruss

*Department of Molecular Cell Biology, Max Planck Institute of Biophysical Chemistry, Göttingen, F.R.G.*

(Received 1 August 1990)
(Accepted 20 August 1990)

**Eight murine paired box-containing (*Pax*) genes have been isolated so far. The gene described here, *Pax7*, contains not only a paired box, but also an octapeptide and a paired-type homeobox. As shown by Northern and in situ analysis, *Pax7* is expressed from day 8 to 17 p.c. during embryogenesis. At early stages *Pax7* transcripts are present in a subset of cells throughout the entire brain, but later in development expression is limited to the mesencephalon. In the developing neural tube *Pax7* is restricted to the dorsal ventricular zone along the entire antero-posterior axis, suggesting a role for *Pax7* in the formation of certain parts of the CNS. Additionally *Pax7* expression can be followed during myogenesis from the dermamyotome of the somites to the skeletal muscle tissues.**

**Paired box; *Pax*; Homeobox; Mouse embryology; Neurogenesis; Myogenesis**

## Introduction

The complex series of events which unfolds during embryogenesis is governed by genetic mechanisms which seem to have been highly conserved during evolution. Many *Drosophila* developmental control genes act as transcription factors based on DNA-binding activity of conserved protein domains to regulatory elements. Consequently many vertebrate genes could be isolated based on sequence similarities to *Drosophila* DNA motifs. One of the best studied examples of a conserved sequence during evolution is the homeobox. It is present not only in developmental control genes of *Drosophila*, but also in highly divergent species such as mouse, man, chicken, zebrafish, *C. elegans* and even in yeast (for review see Scott et al., 1989). The homeoboxes can be subdivided into 11 classes so far. Among them is the paired-type homeobox class, which includes the three *Drosophila* segmentation genes *paired* (*prd*), *gooseberry distal* (*gsb-d*) and *gooseberry proximal* (*gsb-p*) (Frigerio et al., 1986; Baumgartner et al., 1987). These three genes as well as two tissue-specific *Drosophila* genes, *Pox meso* and *Pox neuro* contain a second conserved sequence termed the paired box (Bopp et al., 1986, 1989). This motif seems to be as widely distributed in evolution as the homeobox, since it is found in at least three human and eight murine genes and is present in the genomes of turtle, *Xenopus*, chicken, zebrafish and nematodes (Burri et al., 1989; Deutsch et al., 1988; Dressler et al., 1988, 1990, Plachov et al., 1990; and Walther et al., unpublished data).

The *Drosophila prd, gsb-d* and *gsb-p* genes are known to play a crucial role in segmentation of the embryo. Although the adult mouse is not an obviously segmented organism like *Drosophila*, segmentation is an important part of murine embryogenesis (Hogan et al., 1985). This is most obvious in somitogenesis, in which metameric units, the somites, are formed along the rostrocaudal axis lateral to the neural tube. Each somite differentiates into dermatome, myotome and sclerotome and ultimately into skin, skeletal muscles and axial skeleton, respectively. Here one of the paired box containing genes (*Pax*) genes, namely *Pax1*, is expressed in the sclerotome of the somites and in later stages in the intervertebral disks (Deutsch et al., 1988). The pronephros of the embryonic excretory system is another structure which clearly shows a segmented character. Interestingly two other *Pax* genes, *Pax2* (Dressler et al., 1990) and *Pax8* (Plachov et al., in press) are expressed

*Correspondence:* P. Gruss, Department of Molecular Cell Biology, Max Planck Institute of Biophysical Chemistry, 3400 Göttingen, F.R.G.

in this tissue. Thus all murine paired box containing (*Pax*) genes described so far show expression in segmented structures of the mouse embryo.

Several *Drosophila* segmentation genes show a second regulatory function later in development in neurogenesis (Patel et al., 1989). For instance *gsb*$^-$ mutants exhibit numerous alterations in the pattern of neurons within the central nervous system (CNS). A regulatory role in neurogenesis seems to have also been conserved for the murine *Pax* genes. *Pax2* (Nornes et al., 1990) and *Pax8* (Plachov et al., in press) are expressed in cell subsets of the developing nervous system, compatible with a role in neuron determination. At least three other *Pax* genes are also expressed in neuronal tissues (this report; Goulding et al., unpublished data; Walther et al., unpublished data). All show a distinct expression in the neural tube, generating a variety of overlapping or complementary expression patterns.

Here we describe the expression pattern and part of the genomic organisation of the murine paired box containing gene *Pax7*. This gene was isolated based on sequence homology to the paired box. It is, however, the first *Pax* gene and the first murine gene overall that contains in addition a paired-type homeobox. Apart from these two motifs it also contains another conserved sequence of unknown function, the octapeptide, which was first described in the two gooseberry genes and in two human paired box genes (Burri et al., 1989).

*Pax7* is expressed during somitogenesis. *Pax7* transcripts are present in the myotome of the somites and throughout their differentiation into skeletal muscles. During the development of the nervous system *Pax7* shows a distinct spatial and temporal expression pattern in brain as well as in neural tube where it is restricted to the dorsal part of the ventricular zone, which generates a stem cell population for the entire CNS. Hence *Pax7* contains not only two or three conserved protein domains as paired and the gooseberry genes, respectively, but shows also expression during development of segmented structures and neurogenesis.

```
1   AGTGTCCACCCCTCTTGGCCAAGGCCGGGTCAATCAGCTTGGTGGGGTCTTCATCAACGGT   60
    ValSerThrProLeuGlyGlnGlyArgValAsnGlnLeuGlyGlyValPheIleAsnGly

61  CGACCCCTGCCGAACCACATCCGTCACAAGATAGTGGAAATGGCCCACCATGGCATCCGG   120
    ArgProLeuProAsnHisIleArgHisLysIleValGluMetAlaHisHisGlyIleArg

121 CCCTGCGTCATCTCCCGTCAGCTCCGTGTTTCCCATGGTTGTGTCTCCAAGATTCTATGC   180
    ProCysValIleSerArgGlnLeuArgValSerHisGlyCysValSerLysIleLeuCys

181 CGATATCAGGAGACTGGGTCCATCCGGCCCGGGGCTATCGGAGGCAGCAAGCCCAGACAG   240
    ArgTyrGlnGluThrGlySerIleArgProGlyAlaIleGlyGlySerLysProArgGln

241 GTGGCGACTCCGGATGTGGAGAAAAAGATTGAGGAGTATAAGAGAGAGAACCCCGGGATG   300
    ValAlaThrProAspValGluLysLysIleGluGluTyrLysArgGluAsnProGlyMet

301 TTCAGCTGGGAAATCCGGGACCGGCTGCTGAAGGACGGTCACTGCGACCGAAGCACGGTG   360
    PheSerTrpGluIleArgAspArgLeuLeuLysAspGlyHisCysAspArgSerThrVal

361 CCCTCAGTGAGTTCGATTAGCCGAGTGCTCAGAATCAAGTTCGGGAAGAAAGAGGACGAC   420
    ProSerValSerSerIleSerArgValLeuArgIleLysPheGlyLysLysGluAspAsp

421 GAGGAAGGAGACAAGAAAGAAGAAGATGGCGAGAAGAAAGCCAAACACAGCATCGATGGC   480
    GluGluGlyAspLysLysGluGluAspGlyGluLysLysAlaLysHisSerIleAspGly

481 ATCCTGGGCGACAAAGGGAACCGTCTGGATGAGGGCTCAGGTGTGGAATCAGAACCCGAC   540
    IleLeuGlyAspLysGlyAsnArgLeuAspGluGlySerGlyValGluSerGluProAsp

541 CTCCCCCTGAAGCGCAAGCAGCGCCGCAGTCGGACCACGTTCACAGCCGAGCAGCTGGAG   600
    LeuProLeuLysArgLysGlnArgArgSerArgThrThrPheThrAlaGluGlnLeuGlu

601 GAGCTGGAGAAGGCCTTCGAGAGGACCCACTACCCGGACATCTACACCCGGGAGGAGCTG   660
    GluLeuGluLysAlaPheGluArgThrHisTyrProAspIleTyrThrArgGluGluLeu

661 GCACAGAGGACCAAGCTCACGGAGGCACGCTTCCAGGTCTGGTTCAGTAACCGGCGTGCC   720
    AlaGlnArgThrLysLeuThrGluAlaArgPheGlnValTrpPheSerAsnArgArgAla

721 CGCTGGCGTAAGCAGGCAGGAGCTAACCAGCTGGCCGCCTTCAACCACCTTCTGCCGGGA   780
    ArgTrpArgLysGlnAlaGlyAlaAsnGlnLeuAlaAlaPheAsnHisLeuLeuProGly

781 GGTTTCCCACCCACCGGCATGCCCACGCTGCCACCCTACCAGCTGCCGGACTCTACCTAC   840
    GlyPheProProThrGlyMetProThrLeuProProTyrGlnLeuProAspSerThrTyr

841 CCCACCACCACCACCATCTCCCAAGATGGAGGCAGCACAGTACACAGGCCCCAGCCCCTT   900
    ProThrThrThrThrIleSerGlnAspGlySerThrValHisArgProGlnProLeu

901 CCGCCATCAACCC                                                  917
    ProProSerThr
```

Fig. 1. DNA sequence and predicted protein sequence for a *Pax7* region, including the paired box (underlined), the octapeptide (bold letters) and the paired-type homeobox (double underlined). Positions of introns are marked by arrows.

## Results

*Genomic organisation of Pax7*

Eight murine paired box containing genes have been identified so far, numbered according to their isolation. As described elsewhere (Walther et al., unpublished data) genomic clones of *Pax7* were isolated from a Balb/c mouse library (Clontech) and cDNA clones were obtained by screening an embryonic λgt10 C57BL/6 mouse day 8.5 p.c. library (Fahrner et al., 1987). So far we did not succeed in isolating a full-length cDNA, although three cDNA libraries were screened with different probes.

The longest open reading frame as deduced from the available sequence starts five amino acids upstream and proceeds 170 amino acids downstream of the paired domain (shown in Fig. 1). The paired domain is located close to the amino terminus in all *Pax* genes analyzed so far. Taken together with the high overall homology in structure and sequence of *Pax7* to *Pax3* (Goulding, unpublished data) it can be expected that only 10 to 21 amino acids are missing at the 5′ end, depending on the initiation codon used. Similarly it can be estimated that the stop codon is approximately 150 amino acids downstream of the 3′ end of the available sequence. *Pax7* contains the paired box and in addition two other conserved motifs, the octapeptide (Burri et al., 1989) and a paired-type homeobox (Frigerio et al., 1986).

Thus *Pax7* is the first murine member of the paired-type homeobox class.

As in all other known murine *Pax* genes, with the exception of *Pax1*, the paired box of *Pax7* is encoded on three exons (Walther et al., unpublished data). It is separated from the preceding exon by an intron five amino acids upstream of the box. The highly conserved first part of the box (up to box-codon 74) is separated by another intron from the more divergent second part of the box, which is interrupted by an intron within box-codon 117. The 3′ end of the paired box and a highly charged region followed by the octapeptide downstream of it are encoded on the same exon. The homeobox and flanking sequences are located on a different exon.

Computer analysis (Gribskov and Devereux, 1986; Jameson and Wolf, 1988) of the secondary structure of the protein predicts an α-helix between amino acids 22 to 31 in the first part and a helix-turn-helix motif between amino acids 80 to 106 in the second part of the paired domain.

*The octapeptide*

The octapeptide was first defined by Burri et al. (1989) as a conserved motif present in the two *gooseberry* genes of *Drosophila* and the two human genes HuP1 and HuP2. As indicated in Fig. 2 the octapeptides of *Pax7*, HuP1 and *gsb-d* are identical (Burri et al., 1989). The octapeptide of the other human gene HuP2 shows one amino acid exchange, whereas *gsb-p* is more diverged (5/8) as are the other Pax genes *Pax1* (5/8), *Pax2* (6/8) and *Pax8* (5/8) (Burri et al., 1989; Deutsch, personal communication; Dressler et al., 1990; Nornes et al., 1990; Plachov et al., 1990). Therefore all *Pax* genes described so far contain at least two conserved motifs, the paired box and the octapeptide. The *Drosophila* paired gene does not show any related sequence to the octapeptide.

*A murine paired-type homeobox*

For most of the homeobox classes, members have been identified in *Drosophila*, mouse, man and other species, but so far the class of paired-type homeoboxes consisted only of three *Drosophila* genes – namely *prd*, *gsb-d* and *gsb-p* (Bopp et al., 1986, 1989; Baumgartner et al., 1987). *Pax7* is the first murine gene containing a paired-type homeobox. The homeobox, 61 amino acids long, of *Pax7* exhibits 84% homology to *prd* (Frigerio et al., 1986; Scott et al., 1989), or 93% if conservative exchanges are neglected (S-C, T-S, E-D, K-R, F-I, W-L). Interestingly, the *Pax7* homeobox shows a higher homology to the *gsb-p* (83%) than to the *gsb-d* (74%) homeobox, contrary to the situation for the octapeptide (*gsb-p* 62.5%; *gsb-d* 100%). The homology is extended at least 10 amino acids beyond the amino end of the box and the spacing between the octapeptide and the homeobox is in the range of 20 to 25 amino acids (Fig. 2). No homeobox sequences are known so far for HuP1 and HuP2, whereas *Pax1*, *Pax2* and *Pax8* certainly do not contain a homeobox (Deutsch, personal communication; Dressler et al., 1990; Plachov et al., 1990).

*Pax7 and HuP1 – equivalent genes in mouse and man?*

Strikingly, the paired domains and the octapeptides of *Pax7* and HuP1 (Burri et al., 1989) are identical. Moreover, the highly charged sequence of alternating clusters of basic and acidic amino acids between these two domains shows only four conservative and one non-conservative exchanges. In addition they exhibit the same intron/exon structure in the available sequences. So far, no sequence data downstream of the octapeptide where a homeobox would be expected are available for HuP1. Although the complete sequences are not yet identified, the extensive homologies between *Pax7* and HuP1 strongly suggest that they represent equivalent genes in mouse and man.

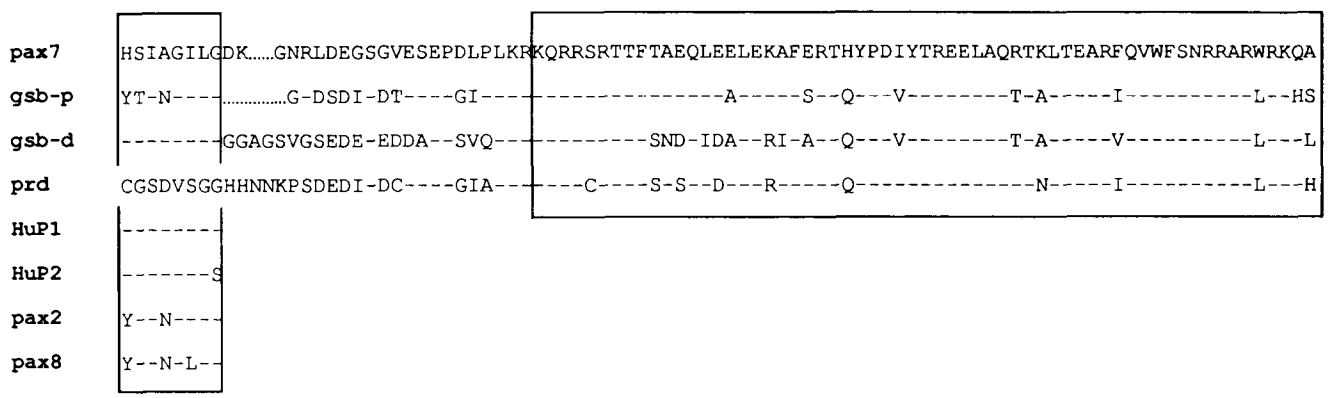



Fig. 2. Comparison of the octapeptide (left box) and the paired-type homeobox (right box) of mouse (*Pax2*, *Pax7*, *Pax8*), man (HuP1, HuP2) and fly (*gsb-p*, *gsb-d*, *prd*). Horizontal lines indicate amino acid identity to *Pax7*. Dotted lines mark missing amino acids compared to the longest sequence of *gsb-d*. The octapeptide box is interrupted, indicating the non-existence of such a box in the *prd* gene. So far a paired-type homeobox is not known for the human genes, whereas *Pax2* and *Pax8* definitely do not contain one.

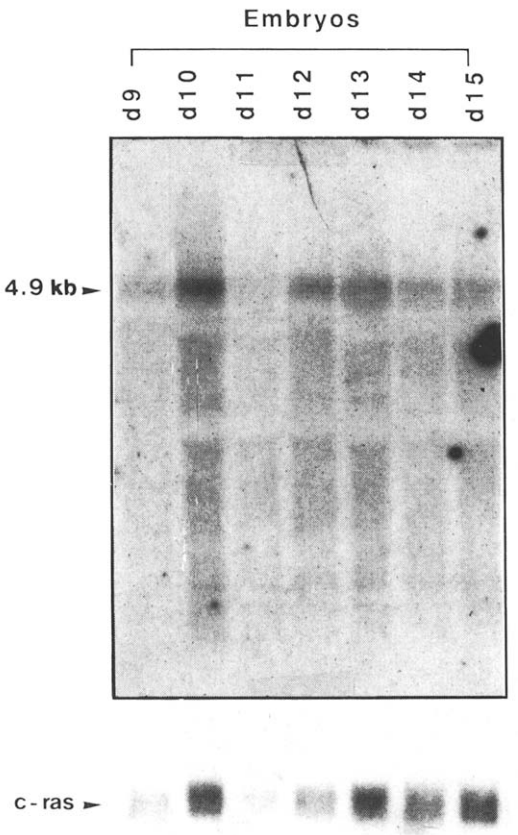



Fig. 3. Expression of *Pax7* in total embryonic tissue. During stage 9 to 15 p.c. a single 4.9 kb transcript can be detected, as indicated by an arrow. For the high stringency hybridization a 291 bp long *Sal*I-*Hin*dIII fragment was used. Because of a weak signal this blot was exposed for 13 days. c-*ras* as control was exposed for o/n.

*Pax7 is expressed in embryonic tissues*

The expression pattern of *Pax7* was first investigated by Northern analysis. No expression was detectable in adult tissues including spleen, uterus, lung, intestine, stomach, thymus, liver, kidney, heart, cerebrum, cerebellum and total brain (data not shown). Adult brain was also tested by in situ analysis, but no signals were detectable. *Pax7* mRNA is present in embryonic tissues, as documented in Fig. 3. A single transcript of about 4.9 kb shows up at all stages tested from day 9 to 15 post coitum (p.c.). Taking the control mRNA into account the amount of transcript is equal during day 9 to 13 p.c. After day 14 p.c. there is a slight drop in the level of expression.

In situ hybridization was done for day 8 to 16 p.c. embryos and sagittal and cross-sections were examined for all stages. Although the probe used contains the 3′ end of the paired box and the 5′ end of the homeobox, it shows less than 70% homology at the nucleotide level to the most closely related gene, *Pax3*, and does not cross-hybridize. *Pax7* is expressed both in ectoderm- and mesoderm-derived tissues (Fig. 4). The ectodermal derivatives expressing *Pax7* are the neural tube, the brain and the nasal olfactory epithelium. *Pax7* expression in the mesoderm begins during myogenesis in the dermamyotome of the somites and persists during differentiation into skeletal muscles of the trunk and limbs.

*Specific expression during neurogenesis*

All known murine *Pax* genes with the exception of *Pax1* (Deutsch et al., 1988) show expression in the neural tube (Nornes et al., 1990; Plachov et al., 1990; Goulding et al., unpublished data; Walther et al., unpublished data). *Pax7* is expressed as soon as the neural tube is closed at day 9 p.c. along the entire anteroposterior axis. The neural epithelium, now called the ventricular zone, is mitotically active and generates stem cell populations for the central nervous system. As shown in Fig. 4 E-J, *Pax7* expression is limited to the dorsal half of the neural tube, but it is not detectable in the most dorsal region of the ventricular zone. According to Altman and Bayer (1984) this region corresponds to the dorsal intermediate plates, but does not include the alar and roof plates. At day 12 p.c. of development a second cell layer, the intermediate zone, has been established. In the ventricular zone there is a clear signal, whereas the intermediate zone and later also the marginal layer are definitely negative (Fig. 4I,J). Only the ventricular zone contains mitotically active cells and the signal decreases in proportion to the reduction of the ventricular zone. Ultimately, this zone is converted into the ependyma, a columnar epithelium surrounding the neural canal and at that stage *Pax7* signals are no longer detectable.

As mentioned above, *Pax7* is not restricted to the neural tube. In early stages it is expressed along the entire anteroposterior axis and has a specific temporal expression pattern during brain development: At day 8 p.c. *Pax7* transcripts can be detected in the forebrain throughout the anterior neuroepithelium (Fig. 5A/B). In this section there is an apparent gap of expression, but this is due to the fact that *Pax7* is only expressed in the marginal zones of the neural epithelium, which is not always in the plane of the section. On day 9 and 10 p.c. (Fig. 5C,D) *Pax7* is highly active in all brain vesicles. Beginning on day 11 (Fig. 5E,F) the signal in the tel- and diencephalon is declining, whereas the mesencephalon shows strong expression in all stages tested. As seen in Fig. 5G,H *Pax7* is expressed both in the roof and in the basal part of the mesencephalon. In the roof the signal can be observed in the whole ventricular zone as in the neural tube, whereas in the basal part the signal is limited to an area close to the boundary of the intermediate zone. Quite obvious on this slide (Fig. 5G,H) is another signal in the nose (first seen on day 12), which shows up in the nasal olfactory epithelium. It is easier to define on day 14 (Fig. 6C,D) where it coincides perfectly with the nasal sinus. *Pax7* is not expressed in any other sense organ.

*Pax7 is transcribed during myogenesis*

*Pax7* is not only expressed in ectoderm- but also mesoderm-derived structures. Fig. 4C,D,G–J and Fig. 6A–D demonstrate clearly the expression pattern of *Pax7* during myogenesis. It can be followed from the somitic mesoderm to the skeletal muscle masses of the

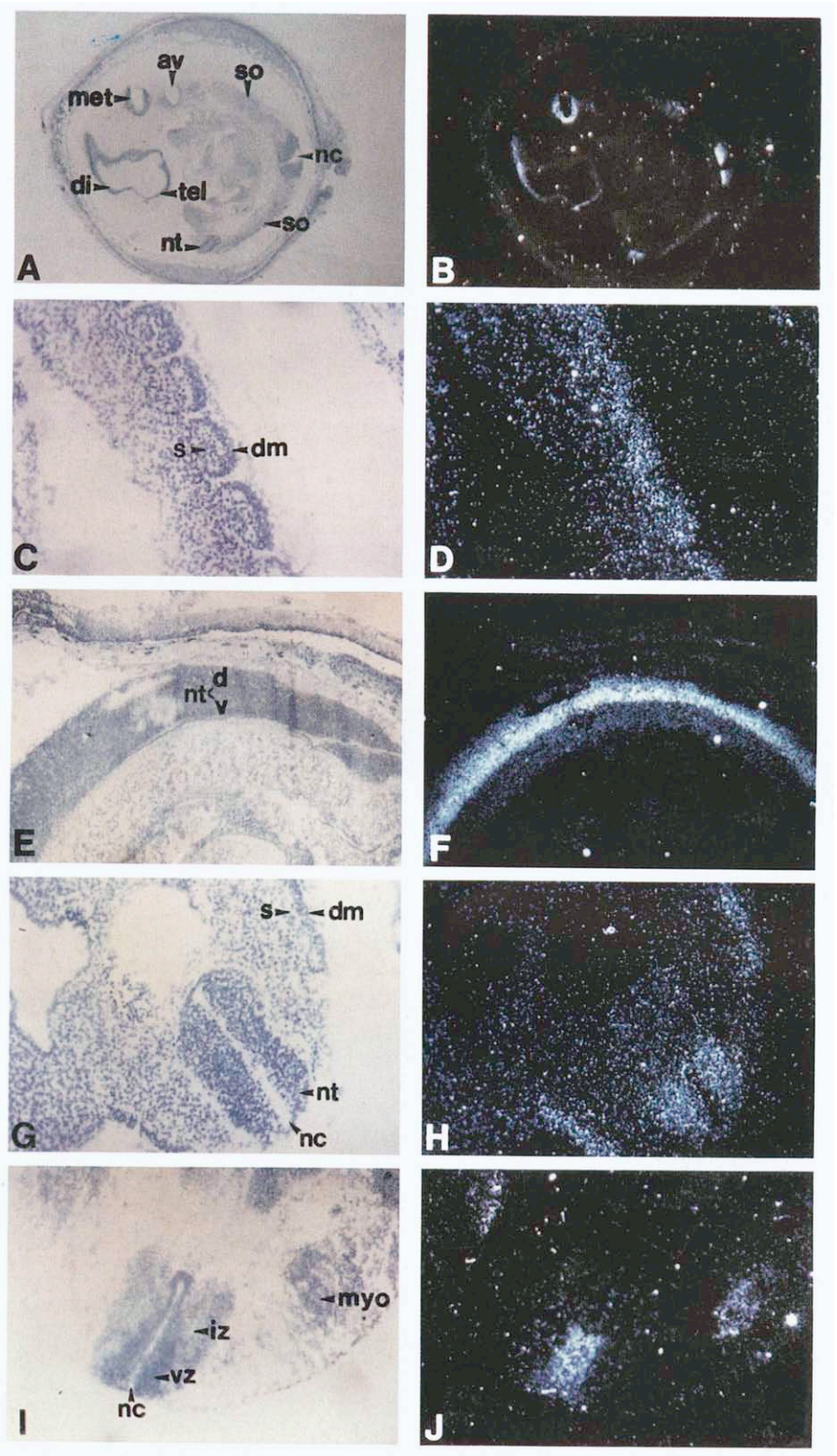
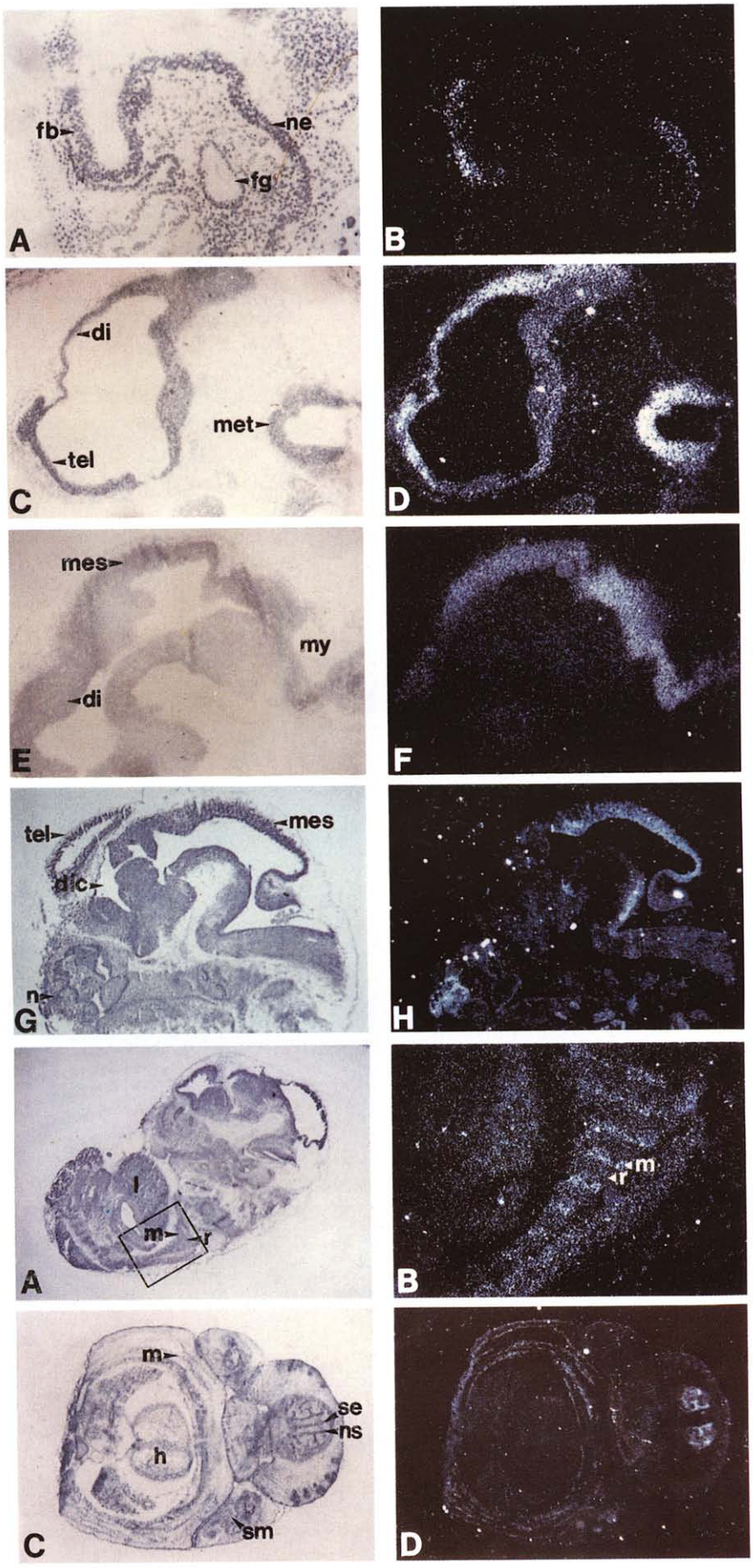



Fig. 4. Expression pattern of *Pax7* in a day 10 p.c. embryo (approx. 3.6 mm). Sagittal sections A-F, cross-sections G, H. Panels I, J are cross-sections of a day 12 p.c. embryo. Brightfield and darkfield images of each section are shown side by side. av, auditory vesicle; d, dorsal; di, diencephalon; dm, dermamyotome; iz, inner zone; met, metencephalon; myo, myotome; nc, neural canal; s, sclerotome; so, somite; tel, telencephalon; v, ventral; vz, ventricular zone.


fb
ne
fg
A
B
di
met
tel
C
D
mes
my
di
E
F
tel
mes
dic
n
G
H
l
m
r
A
m
r
B
m
se
ns
h
sm
C
D

trunk and limbs. *Pax7* mRNA is first detectable on day 9 p.c. in the dermamyotome of the somites. At this stage the dermatome and the myotome are not distinguishable, but later *Pax7* is never found in any dermatome-derived tissue layers. On Fig. 4I,J, a day 12 p.c. cross-section, the expanding myotome exhibits *Pax7* activity. Shown in the day 13 p.c. embryo (Fig. 6A,B) are the positively stained intercostal muscles and on a day 14 p.c. cross section (Fig. 6C,D) the skeletal muscular system of the trunk and the shoulder muscles in the forelimbs. In this cross-section it can also be seen that the cardiac muscles, like all smooth muscles, are negative.

## Discussion

*Molecular structure of Pax7*

The murine *Pax7* gene was originally isolated by virtue of its paired box, a sequence present in DNA-binding proteins (Scott et al., 1989; Treisman et al., 1989, Desplan, personal communication; Goulding, personal communication) which are important in directing embryonic development in *Drosophila* (Akam, 1987; Bopp et al., 1986; Baumgartner et al., 1987; Patel et al., 1989). In fact, *Pax7* contains three conserved protein domains, the paired domain, the octapeptide and the paired-type homeodomain. The paired box was initially identified in the *Drosophila* segmentation genes *paired, gsb-d* and *gsb-p* (Bopp et al., 1986). Furthermore, *gsb-d* and *gsb-p*, but not *paired*, contain the octapeptide which was first defined for these two genes and the human HuP1 and HuP2 (Bürri et al., 1989). The paired domains and the octapeptides of HuP1 and *Pax7* are identical. Additionally, the flanking sequences are highly conserved and *Pax7* and HuP1 exhibit the same intron-exon boundaries. Although the homeobox has not yet been identified for HuP1 the extensive homologies of the available sequences strongly suggest that *Pax7* and HuP1 represent equivalent genes in mouse and man.

So far only three genes have been described which belong to the paired-type homeobox class – namely the *Drosophila* segmentation genes, *prd, gsb-d* and *gsb-p* (Frigerio et al., 1987). In this report we describe for the first time a murine member of this class. All other published *Pax* genes do not contain a paired-type homeobox (*Pax1*, Deutsch, personal communication; *Pax2*, Dressler et al., 1990; *Pax8*, Plachov et al., 1990). However, *Pax3* and *Pax6* also show this conserved sequence (Goulding et al., unpublished data; Walther et al., unpublished data).

Specific DNA-binding activity for the paired-type homeobox of *prd* has been shown by Treisman et al. (1989). They also demonstrated that amino acid 9 of the recognition helix is crucial for the specificity of this binding. The paired-type homeobox has a serine at this position and recognizes the $e_5$ motif, first defined in the *even-skipped* (*eve*) promoter. *fushi tarazu* (*ftz*) and *zerknüllt* (*zen*) are examples of class I homeoboxes, which have a glutamine at amino acid 9 of the recognition helix and recognize the synthetic $NP_6$ consensus sequence (Desplan et al., 1988; Hoey and Levine, 1988). If the serine of the paired-type homeobox is changed to a glutamine, binding to $NP_6$ is now possible. Similarly a change of amino acid 9 to a glutamine confers binding ability to the *bcd* target sequences within the hunchback promoter (Driever and Nüsslein-Volhard, 1989). Surprisingly, these mutants still bind to the $e_5$ sequence in the *eve* promoter, which is in contradiction to a unique determinative function of amino acid 9. A mutant with three exchanges at absolutely conserved positions in the homeobox was still able to bind $e_5$. This implies the existence of another DNA-binding domain and it has been shown recently that this binding activity is due to the paired box (Desplan, personal communication; Goulding, personal communication).

*Segmentation processes during murine development*

The *Drosophila* paired box genes are part of the hierarchy of developmental control genes and are involved in segmentation of the embryo. Although the adult mouse is not an obviously segmented organism like *Drosophila*, segmentation plays an important role during murine embryogenesis. All *Pax* genes studied so far are expressed in segmented tissues. *Pax1* and *Pax7* transcripts are detectable during somitogenesis (Deutsch et al., 1988; this report), whereas *Pax2* and *Pax8* are known to be expressed in the segmented

Fig. 5. (see opposite page, top panels A–H) Expression of *Pax7* during brain development. Sagittal sections of day 8 (A, B), day 10 (C, D), day 11 (E, F) and day 13 p.c. (G, H). di, diencephalon; dic, diocoel; fb, forebrain; fg, foregut; mes, mesencephalon; met, metencephalon; my, myelencephalon; n, nose; ne, neuroepithelium; tel, telencephalon.

Fig. 6. (see opposite page, bottom panel A–D) (A, B) Expression of *Pax7* in the skeletal muscles in between the ribs of a day 13 p.c. embryo sagittal section (approx. 10 mm). The box in the brightfield photograph indicates the area of higher magnification shown in the darkfield image. l, liver; m, muscles; r, ribs. (C, D) Expression of *Pax7* in a day 14 p.c. cross-section through the head and the trunk at the level of the heart. h, heart; m, muscles; ns, nasal sinus; se, nasal septum; sm, shoulder muscles.

pronephros, the early embryonic excretory system (Dressler et al., 1990; Plachov et al., in press).

Somitogenesis is the best example of segmentation in the mouse embryo. The somites are metameric units along the rostrocaudal axis lateral to the neural tube, which subsequently differentiate into dermatome, myotome and sclerotome. Ultimately they generate the skin, skeletal muscles and axial skeleton. *Pax7* is expressed in the myotome and later in the skeletal muscles, whereas *Pax1* has been shown to be expressed in the sclerotome and at later stages in the intervertebral disks (Deutsch et al., 1988).

The myotome of the somites gives rise to all skeletal muscles of the trunk and to the limb muscles (Carlson, 1988). *Pax7* expression begins with the loose somitic mesoderm and continues throughout its conversion into myoblasts, the formation of myotubes and the organization into muscle fibers. *Pax7* signals cannot be detected in the cardiac muscles or any smooth muscles, which do not derive from somitic mesoderm. Much of the smooth muscles are descendants from the splanchnic mesoderm where *Pax7* is absent. This shows the specificity of *Pax7* for somitic mesoderm-derived muscle tissues, the skeletal muscular system of the trunk and limbs.

*Myogenin* and *MyoD1*, two murine skeletal muscle-specific genes, show a similar expression pattern to *Pax7* (Davis et al., 1987; Edmondson and Olsen, 1989; Wright et al., 1989; Sassoon et al., 1989). *Myogenin* and *Pax7* are first detected at day 9 p.c. in somitic mesoderm, but *Pax7* seems to be active in a broader range of cells. *MyoD1* is not expressed at this early stage, but starting on day 10.5 p.c. of development a signal of increasing intensity can be observed. All three genes exhibit the same pattern of expression in the intercostal myoblasts. In addition, the skeletal muscles of the trunk and the limbs maintain all three mRNAs in high abundancy. For *MyoD1* homologous genes have been identified in man (*myf3*), chick and *Xenopus* (Braun et al., 1989b; Lin et al., 1989; Hopwood et al., 1989). The human homologue to *myogenin* is *myf4* (Braun et al., 1989b) and there are others, *myf5* and *myf6*, which are only known in man so far (Braun et al., 1989a, 1990). All the human myf gene products, as well as *myogenin* and *MyoD1* are able to convert the myogenic C3H 10T1/2 fibroblasts into myotubes, so each of these genes is able to induce the myogenic program. Therefore all of them can act as muscle specific regulatory factors. A helix-loop-helix (HLH) structure, originally designated the myc-like domain (Davis et al., 1987), is conserved in all the muscle-specific genes described above. This HLH motif is necessary and sufficient for the generation of the muscle phenotype in 10T1/2 cells by the *MyoD1* protein (Tapscott et al., 1988). *Pax7* does not contain this motif, although it cannot be excluded that it is encoded on the not yet available portion of the coding sequence. It also remains to be shown, if *Pax7* can act as a muscle-specific regulatory factor and induce the myogenic program as the genes described above.

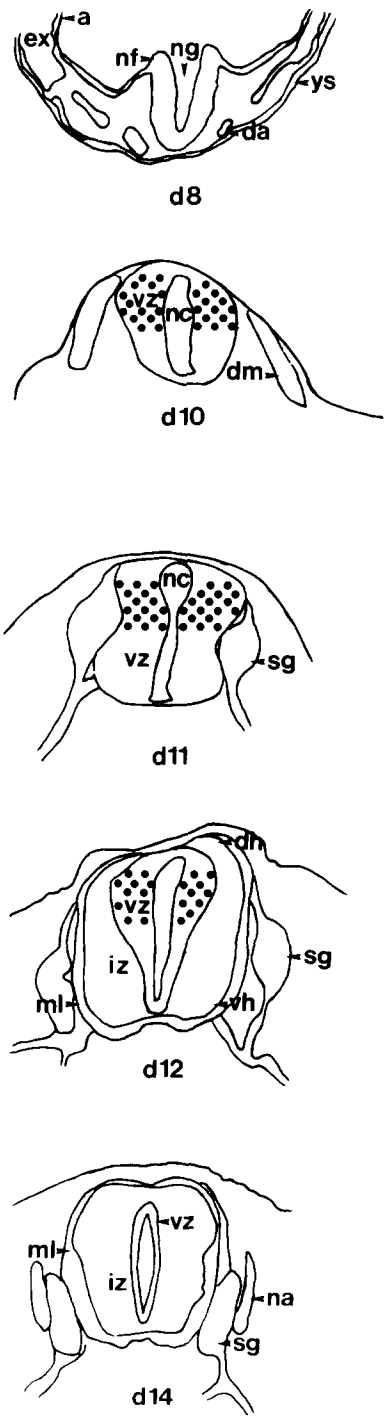



Fig. 7. Schematic illustration of the developing neural tube from day 8 to 14 p.c. The regions where *Pax7* is expressed in the neural tube are marked by dots. a, amnion; da, dorsal aorta; dh, dorsal horns; dm, dermamyotome; ex, exocoelom; iz, intermediate zone; ml, marginal layer; na, neural arch; nc, neural canal; nf, neural fold; ng, neural groove; sg, spinal ganglia; vh, ventral horn; vz ventricular zone; ys, yolk sac.

*Pax7 and neurogenesis*

The neural tube is primarily divided: (a) antero-posterior at the mesencephalon-rhombencephalon border and (b) dorsal-ventral at the sulcus limitans. *Pax7* shows no restriction concerning the antero-posterior axis; however, it is clearly limited to the dorsal half of the neural tube, except in the mesencephalon.

*Pax7* is first detectable at day 8 p.c. in the headfold, which is not yet differentiated into pro-, mes- and rhombencephalon. Surprisingly *Pax7* is not expressed in the more posterior regions until the neural tube is closed (Fig. 7). The proto-oncogene *int-1*, the murine homologue to the *Drosophila wingless*-gene, exhibits a similar expression pattern at this stage (Wilkinson et al., 1987). It is also active only in the headfold, but compared to *Pax7* its expression extends slightly more medial in the neural plate, and *Pax7* is more restricted to the lateral tips. It is unlikely that *Pax7* or *int-1* are involved in the first induction processes of neurulation, because they are not yet expressed along the entire longitudinal axis.

Later at day 9 p.c., *Pax7* is expressed throughout the entire rostrocaudal axis in the dorsal part of the neural tube. At this stage, the neural tube consists only of one cell layer, the ventricular zone. The cells close to the central canal are fixed by tight junctions to a particular position and undergo mitosis, thus producing stem cell populations for the whole central nervous system (Sauer, 1935, 1959; Sidman et al., 1959; Lyser, 1968; Langman and Haden, 1970; Nornes and Carry, 1978; Holley, 1982; Altman and Bayer, 1984). Starting at day 11 p.c., stem cells in the ventral part of the tube lose their connection to the luminal side and migrate radially in a specific order with regard to their neighbouring cells (Nornes and Das, 1974; Altman and Bayer, 1984). This maturation proceeds from ventral to dorsal. In the newly formed intermediate zone and marginal layer, differentiation of the postmitotic neuroblasts takes place and the signal of *Pax7* is correspondingly reduced. In these cell layers *Pax7* exhibits no activity. The ventricular zone is reduced as the intermediate and marginal zones expand. At day 14 *Pax7* expression is no longer detectable. This correlates with the fact that the ventricular zone has stopped mitosis and now represents the ependyma, a columnar epithelium surrounding the central canal. The intermediate zone and the marginal layer give rise to the gray and white matter of the central nervous system, respectively. The neuroblasts in the zones lateral to the *Pax7* expressing stem cells give rise to sensory neurons (Altman and Bayer, 1984).

Altman and Bayer (1984) subdivided the neural tube into five epithelial zones, of which only three components of the lateral plates produce neurons. In a dorsal to ventral order these are the alar, intermediate and basal plates, which are the source of microneurons (i.e., interneurons of the dorsal horns), mesoneurons (i.e., relay neurons of the intermediate gray) and macroneurons (i.e., motor neurons of the ventral horns), respectively. The roof and floor plates are the source of neuroglial cell types rather than of neurons. Based on this subdivision of the neural tube *Pax7* is only expressed in the intermediate plates. *Pax3* is also known to be active in the intermediate plates and additionally in the alar and the roof plate (Goulding, personal communication), whereas *int-1* shows activity only in the roof plate (Wilkinson et al., 1987). Ventral to the sulcus limitans *Pax6* is expressed in the basal plate of the ventricular zone (Walther, personal communication). The murine homologue to the *Drosophila even-skipped* segmentation gene, *Evx1*, is active in the ventricular zone just ventral of the sulcus limitans (Bastian and Gruss, 1990) in approximately the same region as *Engrailed-1* (*En-1*; Davis and Joyner, 1988) and partially overlapping the *Pax6* expressing cells. *Pax2* (Nornes et al., 1990) and *Pax8* (Plachov et al., 1990) exhibit a similar expression pattern in the ventricular zone. However, *Pax2* like *Evx1* and *En1* is not restricted to this layer of the neural tube, but is also expressed in a subset of cells in the intermediate zone. All these genes generate a number of restricted but overlapping or complementary expression patterns in the neural tube. Thus in a concerted action they may play a role in the determination and specification of neuronal and glial cell subtypes during the development of the neural tube.

## Experimental procedures

*DNA sequencing*

Genomic fragments were cloned into the plasmid vector Bluescript KS (Stratagene), cDNAs into M13mp18 and M13mp19. Sequencing was done by the dideoxy method (Sanger et al., 1977) using commercial sequencing kits (Sequenase, US Biochemicals; T7, Pharmacia).

*RNA isolation and Northern-blot analysis*

Tissues were obtained from adult NMRI females and embryos from natural matings of female NMRI and male C57BL6 mice for RNA isolation. The morning of the vaginal plug was designated as day 0 p.c. Total RNA was isolated by homogenization of tissues and embryos in guanidium isothiocyanate according to Chirgwin et al. (1979) and poly(A)$^+$ RNA was isolated by oligo(dT)-celullose chromatography (Ausubel et al., 1989). Approximately 10 $\mu$g of poly(A)$^+$ RNA were electrophoresed through a 1% agarose gel in 3.7% formaldehyde and 1 × Mops buffer (20 mM morpholinepropanesulfonic acid, 50 mM sodium-acetate, 10 mM EDTA, pH 7.0) and blotted onto Nylon membranes (Hybond-N) with 10 × SSC. The blot was hybridized o/n at 42°C in 5 × SSC, 50% formamide, 5 ×

Denhardts, 0.5% SDS, washed twice for 5 min at room temperature and then 30 min at 65 °C in 2 × SSC, 0.1% SDS, followed by another wash at 65 °C in 0.1 × SSC, 0.1% SDS.

*In situ hybridization*

Basically, the protocol of Hogan et al. (1986) was used with some modifications. A 319 bp *Sma*I-*Stu*I fragment was subcloned into the Bluescript KS vector (Stratagene). The resulting plasmid was linearized by restriction endonucleases. Single-stranded high specific activity RNA probes were transcribed in vitro using 100 μCi $^{35}$S-UTP and T3 or T7 polymerases (Promega Biotech.). After DNAse I digestion, the RNA was precipitated with 10% trichloroacetic acid and collected on nitrocellulose filters. The probes were eluted at 65 °C in 20 mM EDTA pH 8.0, 0.1% SDS and ethanol precipitated. Partial degradation was done on ice with 0.2 N NaOH for 30 min, followed by neutralization with 1 M acetic acid. After ethanol precipitation the probes were resuspended in 50% formamide, 10 mM DTT.

Embryos were removed from all surrounding tissues in cold phosphate-buffered saline (PBS) and frozen immediately on dry ice. Embryos of day 8 to 10 were kept in decidua. After embedding in OCT medium (Miles Laboratory) the embryo was cut in a cryostat at −20 °C into 8 μm slices and transferred onto gelatine-chromealaun subbed slides. The sections were dried at 55 °C for 10 min, fixed in 4% paraformaldehyde (PFA) for 20 min, washed for 5 min in 3 × PBS and twice in 1 × PBS and dehydrated in graded ethanol. Slides were stored at −20 °C.

Before hybridization the slides were dipped into PBS and incubated for 30 min in 2 × SSC at 70 °C. All further treatments were done at room temperature. Slides were dipped again in PBS, digested by 0.125 mg · ml$^{-1}$ pronase for 10 min, dipped for 30 s in 0.2% glycine and PBS. After refixation in 4% PFA for 20 min, the slides were rinsed in PBS and then acetylated in 0.1 M triethanolamine with 1/400 vol. acetic anhydride for 10 min. Following another rinse with PBS, the slides were dehydrated in graded ethanol and dried for 2 h.

The RNA probe was diluted to $5 \times 10^4$ cpm/μl in hybridization buffer (50% formamide, 10% dextran sulphate, 10 mM Tris, pH 7.5, 10 mM NaPi, pH 6.8, 5 mM EDTA, 10 mM DTT, 10 mM β-mercaptoethanol, 1 mM ADP[βS], 0.1 mM UTP, 10 μM ATP [γS], 2 × SSC, 150 μg/ml salmon sperm DNA, 150 μg/ml yeast tRNA). Approximately 8 μl of this mix were used for two sections. For hybridization the slides were placed in a humidified chamber with wash buffer (50% formamide, 2 × SSC, 10 mM β-mercaptoethanol) and incubated o/n at 50 °C. The slides were washed in this prewarmed buffer for 2 h at 37 °C, digested with RNAse A (20 μg/ml) in 0.5 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 10 mM β-mercaptoethanol and again washed o/n. After dehydration in graded ethanol, slides were immersed in Kodak NTB-2 for autoradiography and exposed for 7 days at 4 °C. Development was done in Kodak D19 developer for 3 min, 1% acetic acid for 30 s and 30% sodium thiosulphate for 3 min. Following an extended rinse with water, slides were stained with Giemsa for 20 min and photomicrographs were taken with a Leitz Labovert bright-field/dark-field microscope.

## Acknowledgements

We thank H. Jäckle and S. Coté for providing gooseberry clones, and K. Fahrner, B. Hogan and G. Martin for their generous gifts of the cDNA libraries. We are grateful to U. Deutsch and D. Plachov for the cDNA clones of *Pax7*, which have been used for sequencing and preparation of in situ probes. We would also like to thank A. Püschel, U. Deutsch, H. Bastian, M. Goulding and in particular C. Lobe for helpful discussions and comments on the manuscript and R. Altschäffel for photographic work. C.W. is supported by a Boehringer Ingelheim Fond Fellowship. This project was supported by the Max Planck Society and the Deutsche Forschungsgemeinschaft (Grant No. En 84/18-4).

## References

Akam, M. (1987) Development, 101, 1–22.

Altman, J. and Bayer, S.A. (1984) The development of the rat spinal cord. In Advances in Anatomy, Embryology and Cell Biology, Vol. 85. New York, Springer-Verlag.

Bastian, H. and Gruss, P. (1990) EMBO J., 9, 1839–1852.

Baumgartner, S., Bopp, D., Burri, M. and Noll, M. (1987) Genes Dev., 1, 1247–1267.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986) Cell, 47, 1033–1040.

Bopp, D., Jamet, E., Baumgartner, S., Burri, M. and Noll, M. (1989) EMBO J., 8, 3447–3457.

Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. and Arnold, H.H. (1989a) EMBO J., 8, 701–709.

Braun, T., Bober, E., Buschhausen-Denker, G., Kotz, S., Grzeschik, K.-H. and Arnold, H.H. (1989b) EMBO J., 8, 3617–3625.

Braun, T., Bober, E., Winter, B., Rosenthal, N. and Arnold, H.H. (1990) EMBO J., 9, 821–831.

Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G. and Noll, M. (1989) EMBO J., 8, 1183–1190.

Carlson, B.M. (1988) Patten's foundation of embryology. McGraw-Hill Book Company, pp. 222–249 and 327–354.

Davis, C.A. and Joyner, A.L. (1988) Genes Dev., 2, 1736–1744.

Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) Cell 51, 987–1000.

Desplan, C., Theis, J. and O'Farrell, P.H. (1988) Cell 54, 1081–1090.

Deutsch, U., Dressler, G.R. and Gruss, P. (1988) Cell 53, 617–625.

Dressler, G.R., Deutsch, U., Balling, R., Simon, D., Guenet, J.-L. and Gruss, P. (1988) Murine genes with homology to *Drosophila* seg-

mentation genes. In Mechanisms of Segmentation. Development, 104 (Suppl.) 181–186.
Dressler, G.R., Deutsch, U., Chowdhury, K., Nornes, H.O. and Gruss, P. (1990) Development, 109, 787–795.
Driever, W. and Nüsslein-Volhard, C. (1989) Nature, 337, 138–143.
Edmondson, D.G. and Olsen, E.N. (1989) Genes Dev. 3, 628–640.
Fahrner, K., Hogan, B.L.M. and Flavell, R.A. (1987) EMBO J., 6, 1269–1271.
Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986) Cell, 47, 735–746.
Gribskov, M. and Devereux, J. (1986) Nucl. Acids Res., 14, 327–334.
Hoey, T. and Levine, M. (1988) Nature, 332, 858–861.
Hogan, B., Holland, P. and Schofield, P. (1985) TIG, 3, 67–74.
Holley, J.A. (1982) J. Comp. Neurol., 205, 371–382.
Jameson and Wolf (1988) CABIOS 4, 181–186.
Langman, J. and Haden, C.C. (1970) J. Comp. Neurol., 138, 419–431.
Lewis, E.B. (1978) Nature, 276, 565.
Lyser, K.M. (1968) Dev. Biol., 17, 117.
Nornes, H.O. and Das, G.D. (1974) Brain Res., 73, 121–138.
Nornes, H.O. and Carry, M. (1978) Brain Res., 159, 1–16.
Nornes, H.O., Dressler, G.R., Knapik, E.W., Deutsch, U. and Gruss, P. (1989) Development, 109, 797–809.
Patel, N.H., Schafer, B., Goodman, C.S. and Holmgren, R. (1989) Genes Dev., 3, 890–904.
Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J.-L. and Gruss, P. (1990) Development, in press.
Sassoon, D., Lyons, G., Wright, W.E., Lin, V., Lassar, A., Weintraub, H. and Buckingham, M. (1989) Nature, 341, 303–307.
Sauer, F.C. (1935) J. Comp. Neurol., 62, 377–405.
Sauer, M.E. (1959) Anat. Rec., 133, 456.
Scott, M.P., Tamkun, J.W. and Hartzell, J.W. (1989) Biochim. Biophys. Acta 989, 25.
Sidman, R.L., Miale, I.L. and Feder, N. (1959) Expl. Neurol. 1, 322–333.
Tapscott, S.J., Davis R.L., Thayer, M.J., Cheng, P.F., Weintraub, H. and Lassar, A.B. (1988) Science, 242, 405–411.
Treisman, J., Gönczy, P., Vashishta, M., Harris, E. and Desplan, C. (1989) Cell, 59, 553–562.
Wilkinson, D.G., Bailes, J.A. and McMahon, A.P. (1987) Cell, 50, 79–88.
Wright, W., Sassoon, D.A. and Lin, V.K. (1989) Cell, 56, 607–617.