Development 116, 611-623 (1992)
Printed in Great Britain © The Company of Biologists Limited 1992



# *PAX8*, a human paired box gene: isolation and expression in developing thyroid, kidney and Wilms' tumors

ANDREJ POLEEV[1], HELMUT FICKENSCHER[2], STEFAN MUNDLOS[3], ANDREAS WINTERPACHT[3], BERNHARD ZABEL[3], ANDREW FIDLER[4], PETER GRUSS[5] and DIMITRIJ PLACHOV[1]

[1]*Institute for Molecular Biology, Billrothstrasse 11, A-5020 Salzburg, Austria*
[2]*Institute for Clinical and Molecular Virology, Loschgestrasse 7, D-8520 Erlangen, Germany*
[3]*Universitäts-Kinderklinik, D-6500 Mainz, Germany*
[4]*Molecular Carcinogenesis Laboratory, Department of Biochemistry, University of Otago, Dunedin, New Zealand*
[5]*Max Planck Institute for Biophysical Chemistry, D-3400 Göttingen, Germany*

## Summary

**Recent evidence indicates a crucial role for paired box genes in mouse and human embryogenesis. The murine *Pax8* gene encodes a sequence-specific transcription factor and is expressed in the developing secretory system as well as in the developing and adult thyroid. This restricted expression pattern suggested involvement of the Pax8 gene in the morphogenesis of the above organs and prompted us to investigate the *PAX8* gene in humans. In this report, we describe the isolation and characterization of *PAX8* cDNAs from a human adult kidney cDNA library. An open reading frame of 450 amino acids contains the 128 amino acid paired domain at its amino-terminal end. The predicted human and mouse Pax8 proteins show 97.8% conservation and are identical in their paired domains. Two independent cDNA clones reveal differential splicing of the *PAX8* transcripts resulting in the removal of a 63 amino acid serine-rich region from the carboxy end of the predicted Pax8 protein. The truncated Pax8 protein becomes more similar to the predicted murine Pax2 protein, that is also expressed during kidney development and lacks the serine rich region. RNAase protection analysis shows the presence of both *PAX8* transcripts in human thyroid, kidney and five Wilms' tumors. No truncated *Pax8* transcripts could be detected in mouse kidney. In situ hybridization to sections of human embryonic and fetal kidney showed expression of *PAX8* in condensed mesenchyme, comma-shaped and S-shaped bodies. In contrast, *PAX2* expression was present mainly in the very early stages of differentiation, in the induced, condensing mesenchyme. This restricted expression pattern suggests a specific role for both genes during glomeruli maturation. Using somatic cell hybrids, we assigned the *PAX8* gene to human chromosome 2. Comparative mapping data suggest a localization on distal 2q.**

Nomenclature: Pax, vertebrate paired box genes (species not defined); *Pax*, mouse paired box gene; *PAX*, human paired box gene.

Key words: paired box, human *PAX8*, human chromosome 2q, Wilms' tumor.

## Introduction

A number of mammalian developmental control genes have recently been isolated and characterized by reversed genetics (for review see Kessel and Gruss, 1990). This approach has greatly enhanced our understanding of the molecular processes that take place during mammalian development. Most of these developmental control genes, such as homeobox (Hox), paired box (Pax) and POU genes, encode transcription factors (Gehring et al., 1990; Treisman et al., 1991; Chalepakis et al., 1991; Rosenfeld, 1991). It is believed that these genes are capable of modulating the expression of corresponding dependent downstream target genes, thus providing positional information and/or controlling organogenesis (Kessel and Gruss, 1990).

The paired box was first identified in the *Drosophila* segmentation genes *paired* (Frigerio et al., 1986), *gooseberry-distal* and *gooseberry-proximal* (Baumgartner et al., 1987; Côté et al., 1987), as a conserved domain of 128 amino acids (Bopp et al., 1986). To date, eight paired box genes from mouse (Walther et al., 1991) and three such genes from man (Burri et al., 1989) have been isolated on the basis of their homology to the paired box sequences of *Drosophila*. The restricted temporal and spatial pattern of the expression of *Pax* genes during mouse embryogenesis suggests a regulatory role in mammalian development.

For example, the *Pax1* gene is expressed in the sclero-

tome cells which form the intervertebral disks of the vertebral column (Deutsch et al., 1988). *Pax2* is expressed in particular cells along the entire developing spinal cord, in the optic and otic vesicles, in the Wolffian duct and nephric cord and subsequently in the ureter and condensed metanephric mesenchyme (Dressler et al., 1990; Nornes et al., 1990; Dressler and Douglass, 1992). Another murine paired box gene, *Pax3*, contains a paired type homeodomain in addition to the paired domain and has been shown to bind DNA in a sequence-specific manner (Goulding et al., 1991). The *Pax3* gene is expressed exclusively during embryogenesis, and its expression is limited to the dorsal part of the developing spinal cord as well as to distinct regions in the brain (Goulding et al., 1991). Preliminary characterization of a further paired box gene with an additional homeobox, *Pax7*, reveals its expression in the developing nervous system and in myoblasts (Jostes et al. 1990).

The role of Pax genes in mammalian development is further substantiated by a recent finding that certain congenital malformations in mice and man are associated with mutations in Pax genes. Mutations in the Pax3 gene are associated with *splotch* phenotype in mice (Epstein et al., 1991) and with the Waardenburg's syndrome in man (Tassabehji et al., 1992; Baldwin et al., 1992). Mutations in the Pax6 gene cause a semidominant mutation *small eye* (*Sey*) in mice (Hill et al., 1991; Walther and Gruss, 1991) and aniridia disorder in humans (Ton et al., 1991).

The last isolated member of the murine paired box gene family, *Pax8*, is highly homologous to *Pax2*, and both genes share an overlapping, but distinct, pattern of expression during murine embryogenesis (Plachov et al., 1990). *Pax8* is expressed during initial stages in induced parts of the developing excretory system, i.e. in the nephric cord and metanephric condensations (Plachov et al., 1990); regions undergoing intensive differentiation during mesenchyme-epithelium transition (Saxén, 1987). In the developing spinal cord, the expression of *Pax8* is transient, coinciding with that of *Pax2* during days 11-12 of embryogenesis. *Pax8* is also expressed during very early stages of thyroid differentiation (Plachov et al. 1990) and in adult thyroid (Zannini et al., 1992). Binding sites for *Pax8* within the promoters of two thyroid-specific genes have been extensively characterized leading to the suggestion that *Pax8* could function as a transcription control factor (Zannini et al., 1992).

In this report, we describe the isolation and sequence of the human *PAX8* gene, with the aim of investigating its expression during human embryogenesis and its possible involvement in congenital malformations. In addition, the possible involvement of *PAX8* in Wilms' tumor, an embryonic tumor of the kidney (Mierau et al., 1987), was evaluated. We have found a high degree of conservation between the amino acid sequences of the human and mouse Pax8 proteins. Furthermore, we describe evidence for differential splicing of the human *PAX8* transcript that results in removal of a 63 amino acid serine-rich region at the carboxy end of the *PAX8* protein. These spliced *PAX8* transcripts are homologous to the sequence of the *Pax2* gene in this domain. We demonstrate the presence of such transcripts in normal human kidney, five Wilms' tumors and human thyroid by RNAase protection assays. Interestingly, these alternatively spliced transcripts are not detectable in the mouse.

Using somatic cell hybrids and comparative mapping data, we mapped the human *PAX8* gene to the distal part of human chromosome 2q.

## Materials and methods

### *cDNA library screening*

The mouse c2A cDNA (Plachov et al., 1990) fragment encompassing most of the paired box was labeled by random priming (Feinberg and Vogelstein, 1983) and used as a probe to screen a human adult kidney cortex cDNA library in λgt10 (Bell et al., 1986) under low-stringency hybridization conditions (42% formamide, 5×SSPE, 5×Denhardt's, 0.5% SDS at 37°C; washing in 2×SSPE, 0.1%SDS at 37°C). In total $1.2\times10^6$ plaques were screened.

### *DNA sequencing*

Overlapping restriction fragments of the human *PAX8* cDNA were cloned into Bluescript KS (Stratagene) plasmid and sequenced by the dideoxy termination method (Sanger et al., 1977) using a T7 sequencing kit (Pharmacia).

### *RNA isolation and RNAase protection analysis*

Total cellular RNA was prepared by the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987) from human tissues (kidney, thyroid and Wilms' tumors), mouse tissues (kidney, brain and liver) and from the human cervix carcinoma cell line HeLa. RNAase protection experiments were carried out according to Stamminger et al. (1990). Total cellular RNA (40 μg) was hybridized with $2.5\times10^5$ counts/minute of $^{32}$P-labelled riboprobe at 37°C and subjected to digestion by RNAase A (6 μg/ml) and RNAase T1 (12 U/ml). Fragments were separated on a 6% denaturating polyacrylamide gel and visualized after exposure to X-ray film (Kodak).

### *Tissue in situ hybridization*

Human fetuses of 9-12 weeks gestational age were obtained after therapeutic abortions using prostaglandin pessaries or suction termination. The tissue was fixed in 4% formalin in PBS for 12-24 hours, placed in 0.5 M sucrose for 6-12 hours to prevent freezing artifacts and then snap frozen in liquid nitrogen. Processing of the tissues was as described earlier (Mundlos et al., 1991). In brief, cross-sections of the developing kidney were cut on a Leitz cryostat and mounted on triethoxyethanolamine-treated microscopic slides. After prehybridization, 30-50 μl of the labelled RNA-probes ($10^5$ counts/minute/μl) in a buffer containing 50% formamide, 10% dextran sulfate, 0.3 M NaCl, 10 mM Tris pH 7, 10 mM $NaHPO_4$ pH 5, 5.5 mM EDTA, 0.02% Denhardt's, and t-RNA 0.5 mg/ml were applied to the slides. Hybridization was carried out at 47°C overnight. Slides were washed at 50°C for 2 hours in 50% formamide/SSC, followed by incubation with 20 μg/ml RNAase (Sigma, Munich, FRG.) for 15 minutes and washed with 1×SSC three times for 30 minutes at room temperature. After dehydration, slides were dipped in Kodak NTB2 photoemulsion diluted 1:1 with water and air dried. Autoradiography was performed at 4°C in a dry chamber. Exposure time varied from 10 to 20 days. Sense probes served as negative controls.

The following cDNA-fragments were used as probes: the *Pvu*II/*Stu*I fragment of human *PAX8* and a 550 bp cDNA-fragment of human *PAX2*. These two clones, which correspond to the sequences outside of the paired box, show little homology to each

other and to other Pax clones. To study the expression pattern of the *WT1* gene, a 610 bp *Hinc*II/*Eco*RI-fragment was used. The probe is derived from the 3′ region of the WT33 cDNA (Call et al., 1990) outside the zinc finger domain to avoid cross-hybridization with corresponding sequences in other genes. The cDNAs were subcloned in Bluescript transcription vectors and transcribed with high efficiency from the T3 or T7 promoters in the presence of $^{35}$S-UTP, 800 μCi/mmol (Amersham, Poole, UK), to generate labelled sense and antisense RNA probes.

### *Mapping studies*

The human-mouse somatic cell hybrid lines used for chromosomal localization of the human *PAX8* gene have been isolated and described previously (Zabel et al., 1985). The mapping panel consisted of these somatic cell hybrids derived from cells from several unrelated individuals and different rodent strains. The hybrids were characterized by PCR amplification of a 329 bp fragment from the *PAX8* cDNA (nt 1622-1950). The PCR reaction was performed using AmpliTaq (Perkin Elmer Cetus) as described by the manufacturers (94°C, 10 seconds; 60°C, 10 seconds; 72°C, 10 seconds; 35 cycles). Oligonucleotides for PCR screening were primer 1 = 5′-AAGTCCAGCATTGCGGCACA-3′ and primer 2 = 5′-GAGGGAAGTGCTTATGGTCC-3′.

## Results

### *Isolation and structure of the human* PAX8 *cDNAs*

Using the mouse Pax8 cDNA c2A, which includes most of the paired box sequence (Plachov et al., 1990), as a probe, we isolated twenty clones from a human adult kidney cortex cDNA library (Bell et al., 1986) in a low-stringency screen. All eleven clones characterized so far represented the human *PAX8* gene. Expression of another paired box gene, *Pax2*, that is highly homologous to the *Pax8* gene, can be detected in mouse adult kidney upon RNAase protection analysis (H. Fickenscher, unpublished data). However, in our search, we were not able to identify any human *PAX2* cDNAs. Four representative *PAX8* cDNA clones are depicted in Fig. 1. Two of them, H8 and H29, contain the entire coding sequence of *PAX8.* All the clones contain at their 3′ ends the polyadenylation sequence AATAAA (Proudfoot and Brownlee, 1976), followed by a long stretch of adenosine residues, twelve nucleotides downstream, in the case of clones H2 and H12 (Fig. 2). The region around the polyadenylation sequence is highly conserved in mouse and human Pax8 genes (data not shown) and probably represents the authentic 3′ end of the Pax8 mRNA. One *PAX8* clone, H14, contains five additional nucleotides, ATCTC, in front of the poly (A) tail, which are absent in the other Pax8 clones (data not shown).

Two cDNA plasmids, H12 and H19, display a deletion of a 189 bp fragment between positions 1059 and 1247 (Figs 1 and 2). This deletion probably represents a differential splicing event in the *PAX8* primary transcript, rather than being a cloning artifact, because the clones H12 and H19 were isolated independently. Furthermore, we identified one mouse *Pax8* cDNA clone, Z8, presenting the same deletion (data not shown).

A comparison of the mouse and human Pax8 coding sequences reveals two small deletions in the human *PAX8* gene that do not change the reading frame. Three nucleotides, AAC, at the murine positions 738-740 and 18 nucleotides at the murine positions 1173-1190 (Plachov et al., 1990) are absent from the human transcript. Apart from these deletions the coding regions from both species show a high degree of homology: of the 1350 compared nucleotides 91.3% are conserved at their corresponding positions. The mouse and human Pax8 cDNAs are similar in size (2528 and 2712 nt, respectively).

### *The open reading frame of the human* PAX8 *gene*

Two *PAX8* cDNA clones contain a large open reading frame predicting a human *PAX8* protein of 450 amino acids (aa) (Fig. 2). The first AUG at position 161 is not surrounded by any Kozak's consensus sequence (Kozak, 1986). However, 16 nucleotides upstream and 11 nucleotides downstream of this AUG are identical in mouse and man. This sequence conservation around the first AUG suggests similar mechanisms of Pax8 translation in both species. The absense of strong homology between the murine and human sequences 5′ upstream of the human position 140 indicates that the AUG at position 161 is the start codon of the open reading frame. This open reading frame contains the 128 aa paired domain implicated in sequence-specific DNA binding (Zannini et al., 1992) at its amino terminus and, 41 aa further downstream, an octapeptide conserved in several other genes containing a paired box, YSINGLLG, the function of which is still unknown (Burri et al., 1989; Plachov et al., 1990). There are four

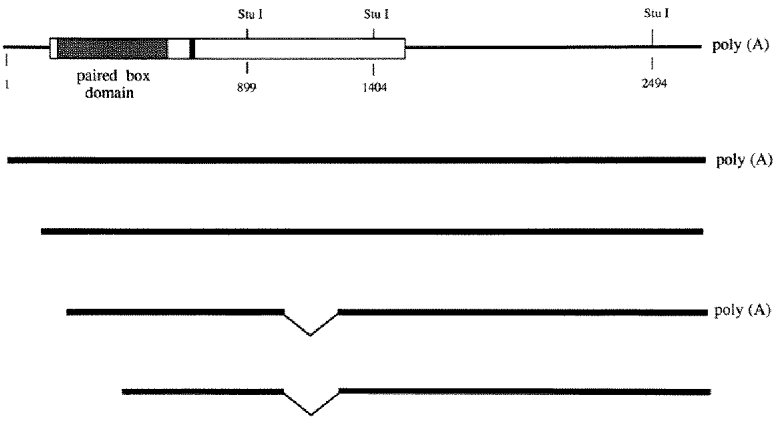


**Fig. 1.** The structure of human *PAX8* cDNA. The coding region is indicated by an open box. The paired box domain and the adjacent octapeptide are indicated by a dotted and filled box, respectively. The *PAX8* cDNA clones are shown underneath as black bars. The presence of the poly(A) tails and the deletion of a sequence encoding a serine-rich region are indicated by corresponding cDNA clones.

leucine residues at aa positions 264, 271, 278, 285 that could contribute to a 'zipper' structure (Landschulz et al., 1988). Leucines are conserved at these positions in mouse *Pax8* and *Pax2* conceptual proteins.

The deletion of 189 bp in plasmids H12 and H19 would result in loss of a 63 aa domain. This region is rich in hydroxylated amino acids, there are 14 serine and 3 threonine residues. Being a multiple of 3, the removal of this 189 bp region retains the reading frame. One clone, H26, shows a deviation in the nucleotide sequence of this region

```
   1  TTCAGAAGGAGGAGAGACACCGGGCCCAGGGCACCCTCGCGGGCGGGCGGACCCAAGCAGTGAGGGCCTGCAGCCGGCCGGCCAGGGCAG  90
  91  CGGCAGGCGCGGCCCGGACCTACGGGAGGAAGCCCCGAGCCCTCGGCGGGCTGCGAGCGACTCCCCGGCGATGCCTCACAACTCCATCAG 180
   1                                                                    M  P  H  N  S  I  R   7

 181  ATCTGGCCATGGAGGGCTGAACCAGCTGGGAGGGGCCTTTGTGAATGGCAGACCTCTGCCGGAAGTGGTCCGCCAGCGCATCGTAGACCT 270
   8   S  G  H  G  G  L  N  Q  L  G  G  A  F  V  N  G  R  P  L  P  E  V  V  R  Q  R  I  V  D  L  37

 271  GGCCCACCAGGGTGTAAGGCCCTGCGACATCTCTCGCCAGCTCCGCGTCAGCCATGGCTGCGTCAGCAAGATCCTTGGCAGGTACTACGA 360
  38   A  H  Q  G  V  R  P  C  D  I  S  R  Q  L  R  V  S  H  G  C  V  S  K  I  L  G  R  Y  Y  E  67

 361  GACTGGCAGCATCCGGCCTGGAGTGATAGGGGGCTCCAAGCCCAAGGTGGCCACCCCCAAGGTGGTGGAGAAGATTGGGGACTACAAACG 450
  68   T  G  S  I  R  P  G  V  I  G  G  S  K  P  K  V  A  T  P  K  V  V  E  K  I  G  D  Y  K  R  97

 451  CCAGAACCCTACCATGTTTGCCTGGGAGATCCGAGACCGGCTCCTGGCTGAGGGCGTCTGTGACAATGACACTGTGCCCAGTGTCAGCTC 540
  98   Q  N  P  T  M  F  A  W  E  I  R  D  R  L  L  A  E  G  V  C  D  N  D  T  V  P  S  V  S  S  127

 541  CATTAATAGAATCATCCGGACCAAAGTGCAGCAACCATTCAACCTCCCTATGGACAGCTGCGTGGCCACCAAGTCCCTGAGTCCCGGACA 630
 128   I  N  R  I  I  R  T  K  V  Q  Q  P  F  N  L  P  M  D  S  C  V  A  T  K  S  L  S  P  G  H  157

 631  CACGCTGATCCCCAGCTCAGCTGTAACTCCCCCGGAGTCACCCCAGTCGGATTCCCTGGGCTCCACCTACTCCATCAATGGGCTCCTGGG 720
 158   T  L  I  P  S  S  A  V  T  P  P  E  S  P  Q  S  D  S  L  G  S  T  Y  S  I  N  G  L  L  G  187

 721  CATCGCTCAGCCTGGCAGCGACAAGAGGAAAATGGATGACAGTGATCAGGATAGCTGCCGACTAAGCATTGACTCACAGAGCAGCAGCAG 810
 188   I  A  Q  P  G  S  D  K  R  K  M  D  D  S  D  Q  D  S  C  R  L  S  I  D  S  Q  S  S  S  S  217

 811  CGGACCCCGAAAGCACCTTCGCACGGATGCCTTCAGCCAGCACCACCTCGAGCCGCTCGAGTGCCCATTTGAGCGGCAGCACTACCCAGA 900
 218   G  P  R  K  H  L  R  T  D  A  F  S  Q  H  H  L  E  P  L  E  C  P  F  E  R  Q  H  Y  P  E  247

 901  GGCCTATGCCTCCCCCAGCCACACCAAAGGCGAGCAGGGCCTCTACCCGCTGCCCTTGCTCAACAGCACCCTGGACGACGGGAAGGCCAC 990
 248   A  Y  A  S  P  S  H  T  K  G  E  Q  G  L  Y  P  L  P  L  L  N  S  T  L  D  D  G  K  A  T  277

 991  CCTGACCCCTTCCAACACGCCACTGGGGCGCAACCTCTCGACTCACCAGACCTACCCCGTGGTGGCAGATCCTCACTCACCCTTCGCCAT 1080
 278   L  T  P  S  N  T  P  L  G  R  N  L  S  T  H  Q  T  Y  P  V  V  A  D  P  H  S  P  F  A  I  307

1081  AAAGCAGGAAACCCCCGAGGTGTCCAGTTCTAGCTCCACCCCTTCCTCTTTATCTAGCTCCGCCTTTTTGGATCTGCAGCAAGTCGGCTC 1170
 308   K  Q  E  T  P  E  V  S  S  S  S  S  T  P  S  S  L  S  S  S  A  F  L  D  L  Q  Q  V  G  S  337

1171  CGGGGTCCCGCCCTTCAATGCCTTTCCCCATGCTGCCTCCGTGTACGGGCAGTTCACGGGCCAGGCCCTCCTCTCAGGGCGAGAGATGGT 1260
 338   G  V  P  P  F  N  A  F  P  H  A  A  S  V  Y  G  Q  F  T  G  Q  A  L  L  S  G  R  E  M  V  367

1261  GGGGCCCACGCTGCCCGGATACCCACCCCACATCCCCACCAGCGGACAGGGCAGCTATGCCTCCTCTGCCATCGCAGGCATGGTGGCAGG 1350
 368   G  P  T  L  P  G  Y  P  P  H  I  P  T  S  G  Q  G  S  Y  A  S  S  A  I  A  G  M  V  A  G  397

1351  AAGTGAATACTCTGGCAATGCCTATGGCCACACCCCCTACTCCTCCTACAGCGAGGCCTGGGGCTTCCCCAACTCCAGCTTGCTGAGTTC 1440
 398   S  E  Y  S  G  N  A  Y  G  H  T  P  Y  S  S  Y  S  E  A  W  G  F  P  N  S  S  L  L  S  S  427

1441  CCCATATTATTACAGTTCCACATCAAGGCCGAGTGCACCGCCCACCACTGCCACGGCCTTTGACCATCTGTAGTTGCCATGGGGACAGTG 1530
 428   P  Y  Y  Y  S  S  T  S  R  P  S  A  P  P  T  T  A  T  A  F  D  H  L  end                  450

1531  GGAGCGACTGAGCAACAGGAGGACTCAGCCTGGGACAGGCCCCAGAGAGTCACACAAAGGAATCTTTATTTATTACATGAAAAATAACCA 1620
1621  CAAGTCCAGCATTGCGGCACACTCCCTGTGTGGTTAATTTAATGAACCATGAAAGACAGGATGACCTTGGACAAGGCCAAACTGTCCTCC 1710
1711  AAGACTCCTTAATGAGGGGCAGGAGTCCCAGGGAAAGAGAACCATGCCATGCTGAAAAAGACAAAATTGAAGAAGAAATGTAGCCCCAGC 1800
1801  CGGTACCCTCCAAAGGAGAGAAGAAGCAATAGCCGAGGAACTTGGGGGGATGGCGAATGGTTCCTGCCCGGGCCCAAGGGTGCACAGGGC 1890
1891  ACCTCCATGGCTCCATTATTAACACAACTCTAGCAATTATGGACCATAAGCACTTCCCTCCAGCCCACAAGTCACAGCCTGGTGCCGAGG 1980
1981  CTCTGCTCACCAGCCACCCAGGGAGTCACCTCCCTCAGCCTCCCGCCTGCCCCACACGGAGGCTCTGGCTGTCCTCTTTCCTCCACTCCA 2070
2071  TTTGCTTGGCTCTTTCTACACCTCCCTCTTGGATGGGCTGAGGGCTGGAGCGAGTCCCTCAGAAATTCCACCAGGCTGTCAGCTGACCTC 2160
2161  TTTTTCCTGCTGCTGTGAAGGTATAGCACCACCCAGGTCCTCCTGCAGTGCGGCATCCCCTTGGCAGCTGCCGTCAGCCAGGCCAGCCCC 2250
2251  AGGGAGCTTAAAACAGACATTCCACAGGGCCTGGGCCCCTGGGAGGTGAGGTGTGGTGTGCGGCTTCACCCAGGGCAGAACAAGGCAGAA 2340
2341  TCGCAGGAAACCCGCTTCCCCTTCCTGACAGCTCCTGCCAAGCCAAATGTGCTTCCTGCAGCTCACGCCCACCAGCTACTGAAGGGACCC 2430
2431  AAGGCACCCCCTGAAGCCAGCGATAGAGGGTCCCTCTCTGCTCCCCAGCAGCTCCTGCCCCCAAGGCCTGACTGTATATACTGTAAATGA 2520
2521  AACTTTGTTTGGGTCAAGCTTCCTTCTTTCTAACCCCCAGACTTTGGCCTCTGAGTGAAATGTCTCTCTTTGCCCTGTGGGGCTTCTCTC 2610
2611  CTTGATGCTTCTTTCTTTTTTTAAAGACAACCTGCCATTACCACATGACTCAATAAACCATTGCTCTTCAAAAAAAAAAAAAAAAAAAAA 2700
2701  AAAAAAAAAAAA                                                                              2712
```

**Fig. 2.** Sequence of human *PAX8*. The paired box domain and the polyadenylation signal are in bold type. The octapeptide is underlined. The borders of a deletion in H12 and H19 cDNA clones are indicated by triangles.

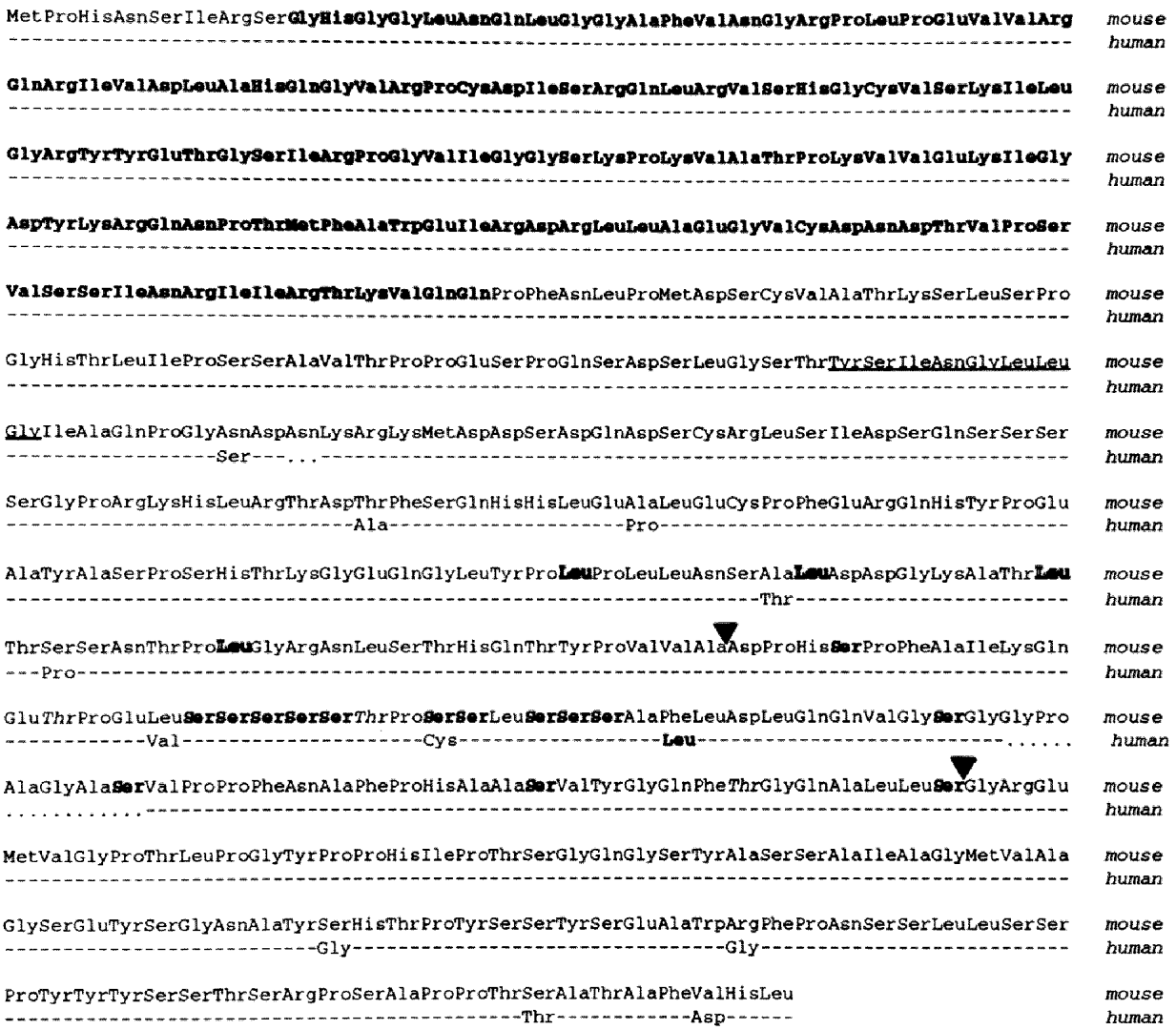


**Fig. 3.** Amino acid comparison of mouse and human Pax8. The paired domain at the amino terminus as well as the serine residues of the region subjected to differential splicing are in bold type. The spliced region is bordered by triangles. The octapeptide is underlined. The leucine residues in 6 aa spacing are outlined. The amino acid residues of the human PAX8 protein identical to those of the mouse protein at corresponding positions are indicated by dashes. The amino acid residues missed in the human protein as compared to the mouse protein are indicated by dots.

(C at position 1145 compared to T in other clones) leading to an amino acid change from Phe to Leu. This difference could either represent a polymorphism in the *PAX8* gene or, alternatively, be a consequence of a mistake by reverse transcriptase.

### *Comparison of the conceptual murine and human Pax8 proteins*

A comparison of the predicted murine and human Pax8 amino acid sequences reveals a very high degree of conservation (Fig. 3). Both proteins are identical from their amino terminus down to the octapeptide. There is a deletion of an Asn residue in the human *PAX8* gene corresponding to the mouse aa position 195. Furthermore, there is a deletion of six amino acid residues in the human *PAX8* gene corresponding to the mouse *Pax8* aa positions 340-346 (Fig. 3). We do not yet know whether this last deletion is due to the differential splicing or a genomic deletion. Of the remaining 450 aa of the human *PAX8* protein, there are only 10 aa substitutions, three of them conservative, as compared to the corresponding positions of the mouse *Pax8* protein (Fig. 3).

### *The chromosomal localization*

The *PAX8* probe (H29) hybridized to a 2 kb fragment in *Hin*dIII digested human DNA, not present in mouse DNA. Analysis of 12 hybrid cell lines demonstrated consistent *PAX8* gene segregation only with human chromosome 2 (Table 1).

### *Expression of the* PAX8 *gene in human tissues and Wilms' tumors*

In order to investigate the distribution of the differentially spliced human *PAX8* transcripts in tissues known to express Pax8 in mouse (Plachov et al., 1990), we used an RNAase protection analysis. The 505 bp *Stu*I fragment from the H26

**Table 1.** *Segregation of the* PAX8 *gene with human chromosome 2*

| Human chromosome | (+/+) | (+/−) | (−/+) | (−/−) | % Discordancy |
|---|---|---|---|---|---|
| 1 | 3 | 4 | 1 | 6 | 36 |
| 2 | 7 | 0 | 0 | 7 | 0 |
| 3 | 3 | 4 | 1 | 6 | 36 |
| 4 | 3 | 4 | 2 | 5 | 43 |
| 5 | 4 | 3 | 1 | 6 | 29 |
| 6 | 4 | 3 | 1 | 6 | 29 |
| 7 | 3 | 4 | 1 | 6 | 36 |
| 8 | 4 | 3 | 3 | 4 | 43 |
| 9 | 2 | 5 | 1 | 6 | 43 |
| 10 | 5 | 2 | 2 | 5 | 29 |
| 11 | 2 | 5 | 0 | 7 | 36 |
| 12 | 4 | 3 | 4 | 3 | 50 |
| 13 | 2 | 5 | 1 | 6 | 43 |
| 14 | 4 | 3 | 1 | 6 | 29 |
| 15 | 3 | 4 | 2 | 5 | 43 |
| 16 | 1 | 6 | 2 | 5 | 57 |
| 17 | 3 | 4 | 5 | 2 | 64 |
| 18 | 4 | 3 | 2 | 5 | 36 |
| 19 | 2 | 5 | 1 | 6 | 43 |
| 20 | 3 | 4 | 3 | 4 | 50 |
| 21 | 3 | 4 | 5 | 2 | 64 |
| 22 | 1 | 6 | 2 | 5 | 57 |
| X | 3 | 4 | 3 | 4 | 50 |

Tabulation of the presence or absence of the human *PAX8* gene in cell hybrids. The detection of *PAX8* is correlated with the presence of each human chromosome in the somatic cell hybrids. Discordancy represents the presence of the gene in the absence of the chromosome (+/−) or the absence of the gene despite the presence of the chromosome (−/+); the sum of these numbers divided by total hybrids examined (×100) represents percentage discordancy.

clone (positions 899-1404), designated H26Stu4, was chosen as a probe. Its sequence is identical to that of H8 and H29 in this region. This fragment encompasses a region in the human *PAX8* gene where the differential splicing occurs (Fig. 1). The sequencing of the cDNAs depicted in Fig. 1 predicts that the antisense riboprobe corresponding to the H26Stu4 fragment would protect a 505 base fragment corresponding to the full *PAX8* sequence and two additional fragments each 157 bases long arising from hybridization to an mRNA corresponding to the sequence represented by the H12 and H19 clones. Indeed, the RNAase protection analysis shows the protected fragments of the expected lengths with total RNA from human kidney, and five different Wilms' tumors. No protection is observed with RNA from human HeLa cells as negative control (fragments fl and B, Fig. 4). Interestingly, the analysis reveals an additional specifically protected fragment of approximately 260 bases that cannot be accounted for by the presented sequencing data (fragment A, Fig. 4). These data might indicate the presence of a third form of *PAX8* transcripts generated by differential splicing in this region of the *PAX8* gene. We are currently investigating this possibility. Furthermore, this experiment shows a particularly strong *PAX8* expression in Wilms' tumors, when compared with the transcript level from normal adult kidney.

The RNAase protection analysis of total RNA from fetal human thyroid with the same H26Stu4 riboprobe provides similar results: three specifically protected fragments can be detected in the thyroid RNA sample (fragments fl, A and

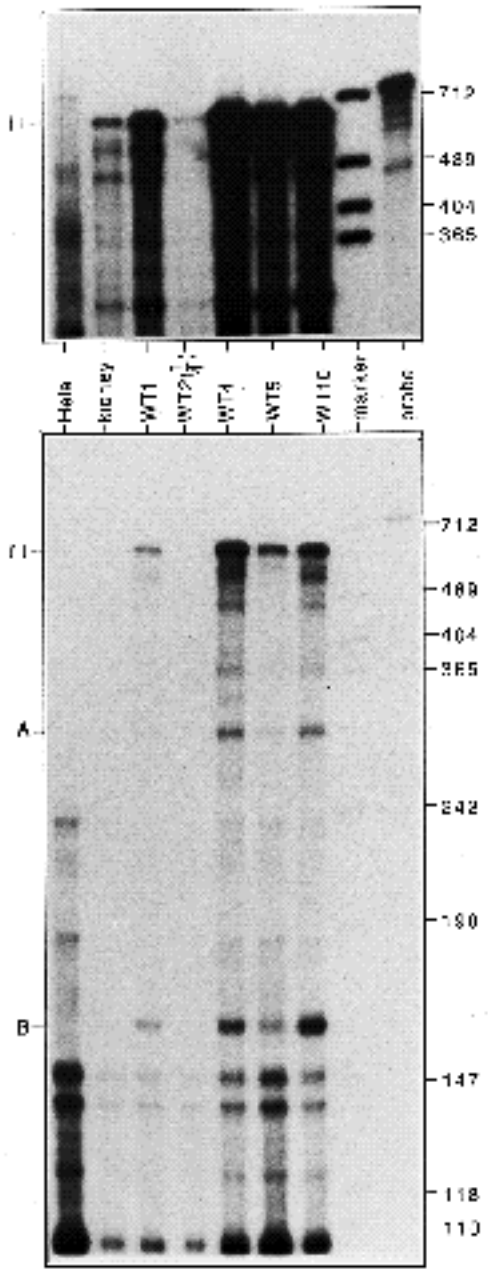


**Fig. 4.** *PAX8* expression in human kidney and Wilms' tumors. Total cellular RNA from HeLa cells, human kidney or from Wilms' tumors (WT) were analyzed by RNAase protection experiments. A riboprobe of high specific activity was prepared from the cDNA plasmid H26Stu4 after linearizing with restriction enzyme *Hin*dIII (probe H26Stu4). After RNAase digestion, specific protected fragments were observed from kidney and Wilms' tumor RNA. The largest band (fl; 505 bases) corresponds to the full-length protected internal fragment, while two predominant smaller fragments (A,B: 260 and 157 bases, respectively) seem to result from alternate spliced messages. 40 μg of total cellular RNA were used from each tissue, except from WT2, where only 10 μg of RNA were used, supplemented with 30 μg of yeast tRNA. Two different exposures are shown, the upper one represents a four-day exposure, while the lower panel shows activities after 12 hours. On the right a DNA marker and the undigested riboprobe are shown.

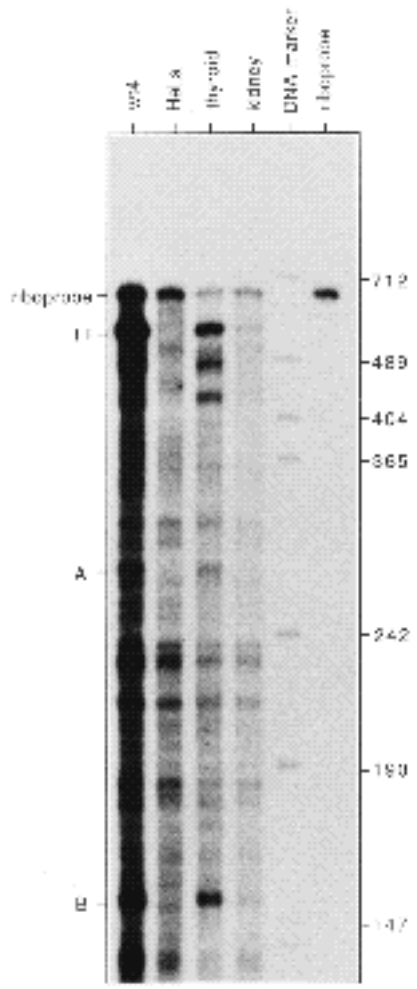


**Fig. 5.** *PAX8* expression in human fetal thyroid. *PAX8* transcripts were visualized from human fetal thyroid tissue using the H26Stu4 riboprobe. RNA from Wilms' tumor number 4 was used as a strong positive, kidney RNA as a weak positive, and HeLa RNA as a negative control. After RNAase digestion, the same specific fragments appeared as in Fig. 4.

B; Fig. 5). Taken together, these data indicate that at least three different types of PAX8 transcripts are expressed in human fetal thyroid, kidneys and five investigated Wilms' tumors.

To investigate the spatial pattern of the *PAX8* expression in the developing human kidney by in situ hybridization, we used a *Pvu*II(648)/*Stu*I(899) (Fig. 2) 251 bp fragment from plasmid H26 as a probe (designated H26PS3). RNAase protection analysis with this probe showed that only one fragment corresponding to the full-length H26PS3 is protected specifically in the RNA samples containing *PAX8* transcripts (H. Fickenscher, unpublished data). It was concluded that no differential splicing occurs in the *PAX8* transcript flanked by positions 648-899 and H26PS3 probably detects all three differentially spliced *PAX8* transcripts.

The fetal kidney (metanephros) provides a unique tissue for the study of developmental processes since all stages of kidney development are present in one section. The early stages of differentiation take place in the cortical parts of the metanephros where the ureteric bud derived epithelium induces undifferentiated mesenchyme. Induced mesenchyme condenses and develops into comma-shaped and S-shaped structures. This is followed by the ingrowth of vessels, the folding of glomeruli and the elongation of tubules in the more central part of the kidney. The formation of new glomeruli is an ongoing process during fetal life which continues up to the 34$^{th}$ week of gestation.

In our experiments, no expression of either *PAX8* or *PAX2* was found in the undifferentiated, non-condensed mesenchyme cells. After induction by the branching ureteric buds, these cells condense and subsequently express high levels of *PAX2* transcripts. This is demonstrated in Fig. 6C,D. The strongly positive structures in the cortical layers have a pairwise appearance caused by the separation through the ureteric bud. The ureteric bud itself showed significantly less expression when compared to the condensed mesenchyme. At this stage, *PAX8* transcripts were present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, *PAX2* expression declined and *PAX8* transcripts were detected at high levels (Fig. 6A,B). *PAX8* expresson was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither *PAX2* nor *PAX8* expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, *PAX2* expression is found mainly in induced, condensing mesenchyme, whereas *PAX8* expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms' tumor gene *WT1* is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, *WT1* expression increases. Within the S-shaped structures, the *WT1*-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.

### *Only one form of Pax8 transcripts can be detected in mouse kidney*

Evidence that differential splicing occurs in a serine-rich region of the human *PAX8* gene prompted us to investigate this region of the mouse *Pax8* transcripts by a RNAase protection analysis. The probes designated c27B/xx22/1 (a 297 bp *Xho*I/*Pst*I fragment of the mouse Pax8 cDNA, positions 855 and 1152, respectively; Plachov et al., 1990) and c27BDde3 (a 107 bp *Pst*I/*Dde*I fragment of the mouse Pax8 cDNA, positions 1152 and 1259, respectively; Plachov et al., 1990) were used. If a differential splicing resulting in the removal of a serine-rich region occurred in the mouse *Pax8* gene, then the c27B/xx22/1 probe would protect two fragments corresponding to the two types of transcripts. A 297 base protected fragment would correspond to the full-length probe, indicating the transcripts coding for the serine-rich region, while a 201 base protected fragment would indicate differential splicing of this region. Only one

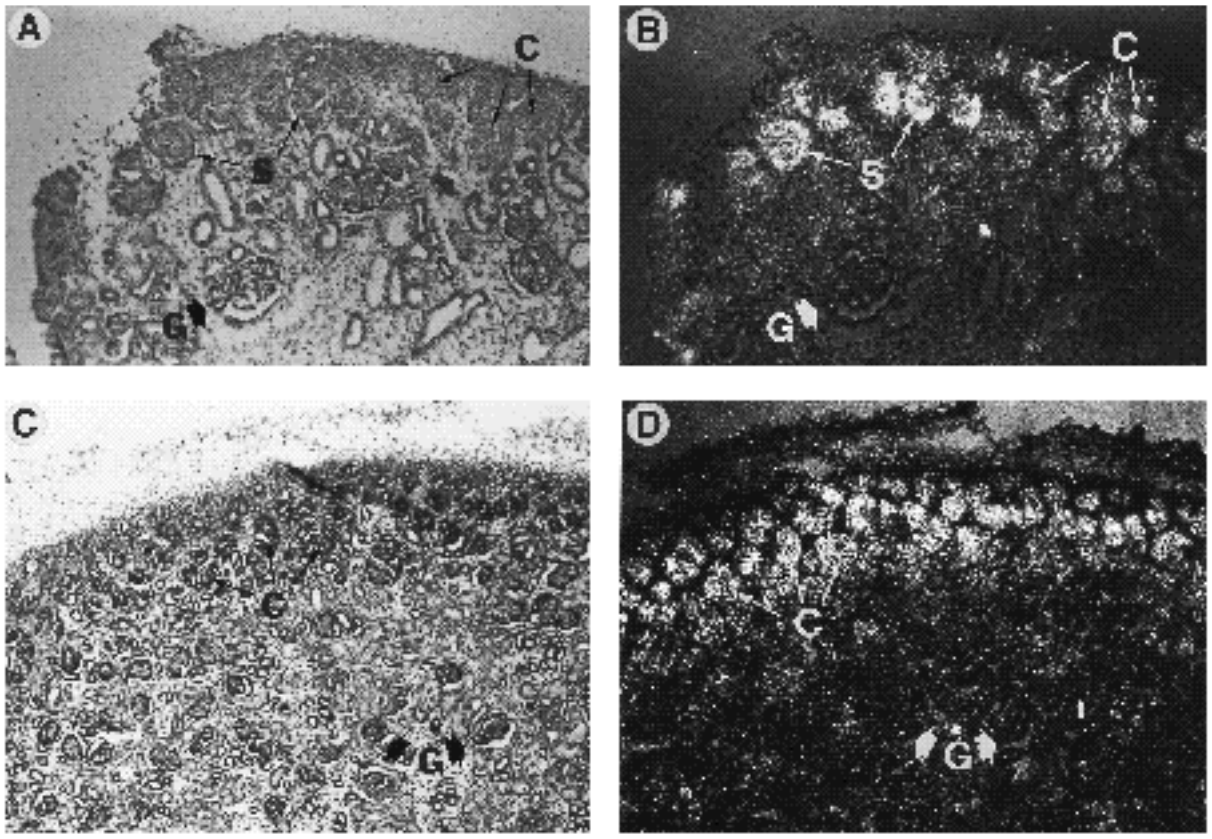


**Fig. 6.** Gene expression of *PAX8* (B) and *PAX2* (D) in section through a 10-week fetal kidney with bright-field images given in (A) and (C). Note expression in condensed mesenchyme (*PAX2*, *PAX8*) and S-bodies (*PAX8*) but not in glomeruli or collecting tubules. G, glomerulus; C, condensed mesenchyme; S, S-shaped bodies; magnification ×70 (A,B) and ×52 (C,D).

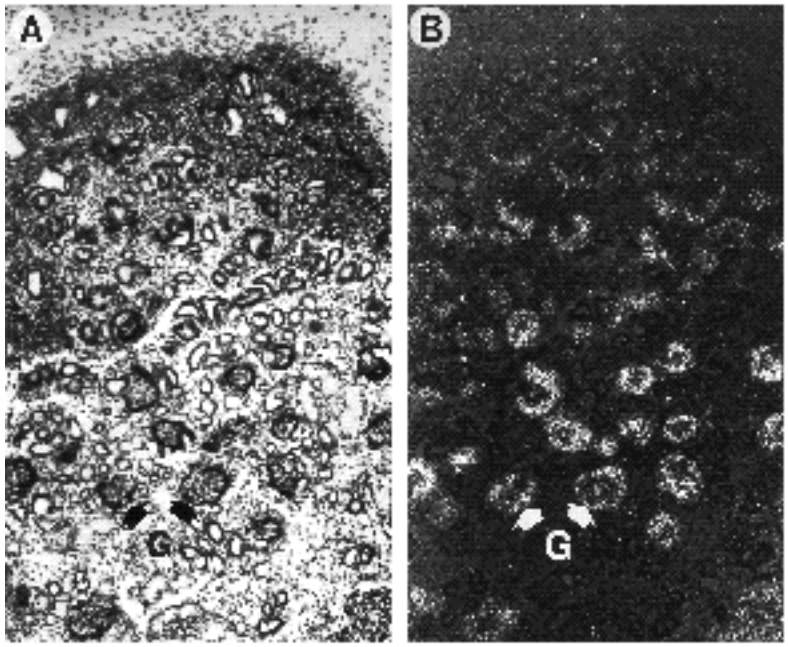


**Fig. 7.** Gene expression of WT1 in a section through a 10-week fetal kidney; (A) bright field, (B) dark field images. Note expression in condensed mesenchyme, S-bodies and in the outer cell layer of glomeruli. G, glomerulus; magnification ×70.

protected fragment in the RNA sample from mouse kidney corresponding to the 297 bp of the full-length c27B/xx22/1 probe can be detected in our analysis (Fig. 8). These data show that only *Pax8* transcripts including the serine-rich region can be detected in the RNA from mouse kidney. The presence of the sequence coding for the serine-rich region

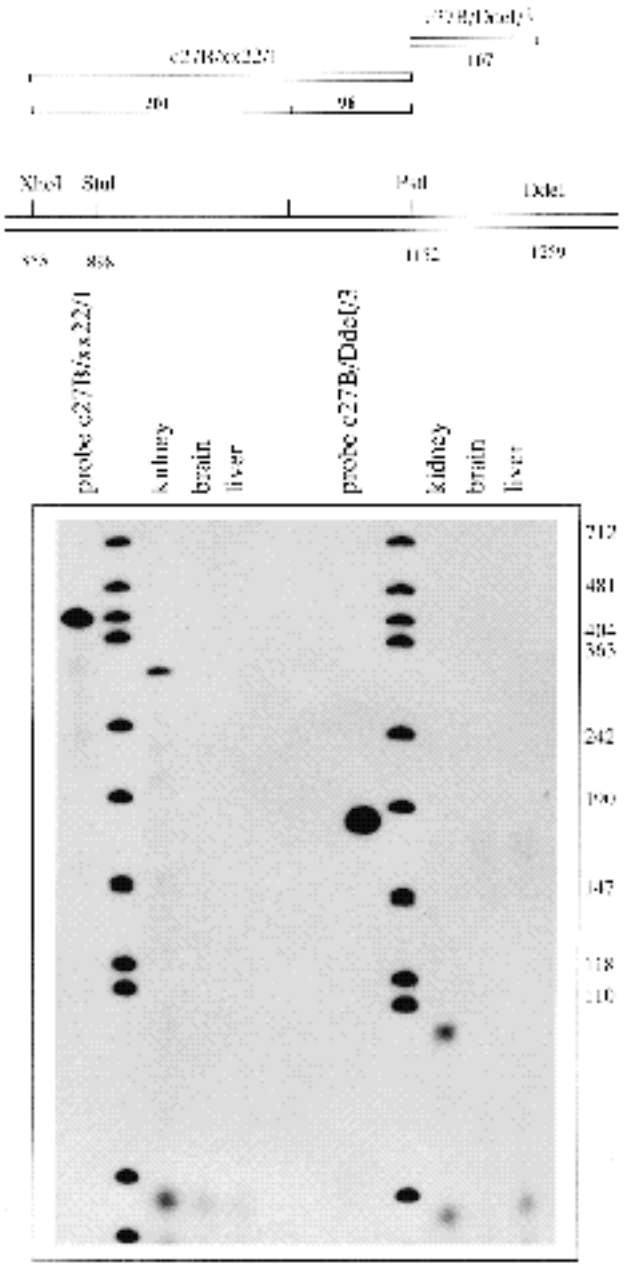


**Fig. 8.** Murine *Pax8* transcripts. Using mouse cDNA plasmids c27B/xx22/1 and c27B/DdeI/3, riboprobes were synthesized and hybridized to total cellular RNA from mouse kidney, brain or liver. In both cases only single protected fragments could be found. Their sizes suggest the absence of alternate splicing for the mouse *Pax8* gene in contrast to human *PAX8*.

in the mouse *Pax8* transcripts is confirmed by the protection of 107 base fragment by the c27BDde3 probe that represents a part of this sequence (Fig. 8).

## Discussion

In this report, we describe the isolation and sequence of cDNA clones representing the human *PAX8* gene. A comparison of the predicted human and mouse Pax8 proteins reveals a very high degree of conservation: there are only 10 amino acid substitutions, three of them conserved, out of 450 amino acids. Such evolutionary conservation may reflect an important function for Pax8 during mammalian development. However, the longest open reading frame, represented by human cDNAs H8 and H29 is seven amino acids shorter than its mouse counterpart. This seven aa truncation results from deletions at two different positions. First, the human *PAX8* protein lacks an asparagine that is present at position 195 of the mouse *Pax8* protein (Plachov et al., 1990). This gap is probably caused by a deletion of three nucleotides at the corresponding position of the human *PAX8* gene because this region of the protein is encoded by a distinct exon in the mouse gene. Second, the *PAX8* protein lacks six amino acids corresponding to the positions 340-345 aa of the mouse protein. Both proteins are identical in their paired domains. This fact suggests that DNA sequences that bind the Pax8 protein in the promoters of putative Pax8 target genes are conserved in mouse and man. Two thyroid-specific candidate target genes, thyroglobulin and thyroperoxidase, have been proposed (Zannini et al., 1992).

It has been shown that the murine Pax8 protein can function as a transcriptional activator (Zannini et al., 1992). The serine-rich region encompassing positions 302-369 aa of the mouse and 301-362 aa of the human Pax8 protein could be a candidate for an activation region (Ptashne, 1988). It has been shown in functional assays that CTF/NF-1 (Mermod et al., 1989), bicoid (Struhl et al., 1989), GHF-1/Pit-1 (Theill et al., 1989), OTF-2/oct-2 (Gerster er al., 1990; Tanaka and Herr, 1990) and LFB1 (Nicosia et al., 1990) contain a serine-rich activation region. Also a paired type homeobox gene, S8, contains such a region (Opstelten et al., 1991). The serine-rich region of the Pax8 protein might be a target for phosphorylation, which plays a crucial functional role for many transcription factors (Lee et al., 1990; Sheng et al., 1991; Ghosh and Baltimore, 1990; Kapiloff et al., 1991; Binetruy et al., 1991; Smeal et al., 1991; Pulverer et al., 1991). Currently we are investigating phosphorylation status of the Pax8 protein. The function of this serine-rich region in the Pax8 protein remains to be determined. Interestingly, this domain is removed in two independent *PAX8* cDNA clones presumably due to differential splicing. Indeed, both transcript variants can be found by RNAase protection experiments in human thyroid, kidneys and five investigated Wilms' tumors. A preliminary characterization of the genomic organisation of *PAX8* indicates that the boundaries of the spliced region correspond to exon acceptor/donor sites (not shown). Differential splicing producing different forms of a transcription factor is a well-documented event (e.g. a homeotic gene Ultrabithorax in *Drosophila*; Lopez and Hogness, 1991). Interestingly, a protein form with a truncated serine-rich domain arises due to a differential splicing in the Oct2 gene (Hatzopoulos et al., 1990). In some cases, differential splicing produces a dominant negative form of a transcription factor (Nakabeppu and Nathans, 1991; Mumberg et al., 1991; Roman et al., 1991).

Similarities between murine *Pax8* and *Pax2* genes have been observed previously (Plachov et al., 1990). Comparison of the published conceptual amino acid sequences of

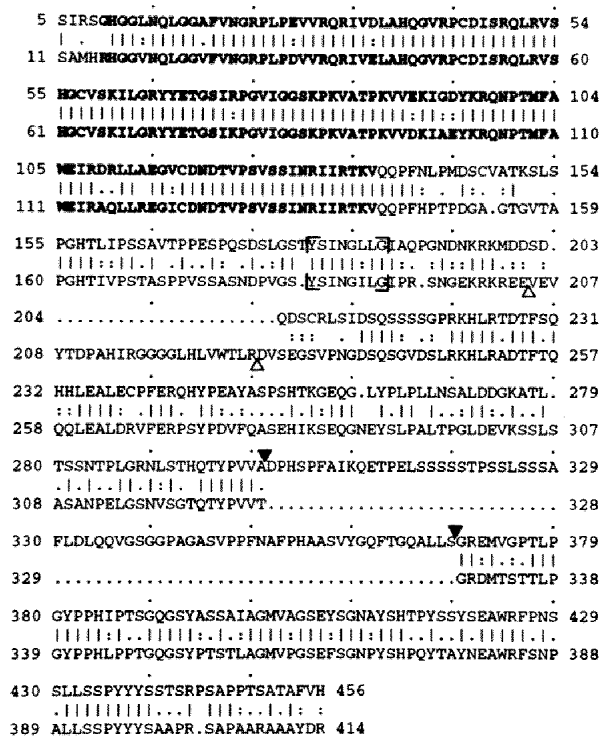

**Fig. 9.** Amino acid comparison of mouse *Pax8* and *Pax2*. The amino acid residues of the murine *Pax8* (upper lane) and *Pax2* (bottom lane) protein are shown in a one letter code. The identical residues of both proteins are connected with dashes, conservative exchanges with two dots and semiconservative with one dot, respectively. The gaps in homology are indicated by dotted lines. The borders of a serine rich region removed in the human *PAX8* protein due to the differential splicing are shown with filled triangles. The borders of a peptide removed in the mouse *Pax2* protein probably due to a differential splicing (Dressler et al., 1990) are indicated with open triangles. The octapeptides in the Pax8 and Pax2 protein are boxed.

the murine Pax8 and Pax2 proteins indicates that these proteins are highly homologous not only in their paired domains but also in their downstream sequences (Fig. 9). However, the homology is interrupted in two regions. Amino acid residues between positions 302 and 369 of the Pax8 protein, encompassing a serine-rich region, are absent in Pax2. Interestingly, exactly this region is altered by alternative splicing in the human *PAX8* gene as discussed above. Thus, alternative splicing in the human *PAX8* gene probably produces a protein form that closely resembles the murine Pax2 protein in structure and possibly in function. It would be interesting to investigate the amino acid sequence of the human PAX2 protein in this domain. Conversely, amino acid residues between positions 207 and 229 of the Pax2 protein are absent in Pax8. Of two described *Pax2* cDNA clones, one shows a deletion of 23 aa in the predicted protein that coincides in size with this insertion and is shifted by two amino acids only (Dressler et al., 1990). Thus, a Pax2 protein form could be generated that is similar to the Pax8 protein in this domain. Moreover, position 203 aa of the Pax8 protein that flanks the gap of homology exactly corresponds to the exon/intron boundary in the *Pax8* gene (Plachov, unpublished data). Currently, we are investigating the precise genomic organisation of murine and human Pax8 genes in this region.

The implications of that molecular mimicry between the two genes are not clear. It could simply reflect a functional redundancy with one gene backing up the other. However, the very high sequence conservation of the Pax8 protein in mouse and man could argue against that possibility. Alternatively, the mimicry could indicate a complex regulative interplay, especially if we assume that the differential splicing in the *PAX8* gene modulates the function of its products as has been shown for other transcription factors (Nakabeppu and Nathans, 1991; Mumberg et al., 1991; Roman et al., 1991). In this respect, it would be interesting to determine whether Pax8 and Pax2 proteins are capable of binding to common DNA targets modulating transcription from downstream genes in a manner similar to the WT1 (Wilms' tumor gene 1) and EGR-1 proteins (Madden et al., 1991). The targeted inactivation of both *Pax8* and *Pax2* genes in transgenic mice would clarify these questions.

The expression patterns of *PAX8* and *PAX2* in the human metanephros are similar but not identical. As shown in Fig. 6, *PAX2* expression is limited to the very early stages of glomeruli development, i.e. the condensed mesenchyme, whereas *PAX8* expression was mainly present in the following, more differentiated stages, i.e. comma-shaped and S-shaped bodies. Neither gene showed expression in the glomeruli or any other structures of the metanephros. These expression patterns are very similar to those reported in the mouse.

During the process of induction, mesenchymal condensations and the formation of S-bodies, profound changes in the expression of extracellular matrix genes take place. Induced cells no longer produce collagens type I and III but switch to collagen IV, osteonectin, laminin, fibronectin and proteoglycans (Ekblom et al., 1981; Mundlos et al., 1992; Laurie et al., 1989). Being putative transcription factors, PAX2 and PAX8 may be involved in the transcriptional regulation of these matrix genes.

One important objective behind the isolation of the human *PAX8* gene was to investigate its function in Wilms' tumors, embryonal kidney tumors (Mierau et al., 1987). At least three genomic regions have been implicated in the genesis of Wilms' tumors (van Heyningen and Hastie, 1992). A deleted region on chromosome 11p13 is associated with the WAGR syndrome of high risk Wilms' tumor, aniridia, genitourinary anomalies and mental retardation (reviewed in Glaser et al., 1989). The Beckwith-Wiedemann syndrome is also associated with a high risk of Wilms' tumor and the corresponding gene has been mapped to 11p15 (Koufos et al., 1989; Reeve et al., 1989; Dowdy et al., 1991). A third Wilms' tumor predisposing gene, not linked to 11p, has been implicated by genetic linkage analysis (Grundy et al., 1988; Huff et al., 1988; Schwartz et al., 1991). The genes on 11p possibly represent 'tumor suppressor' genes (Weissman et al., 1987; Dowdy et al., 1991). The molecular mechanisms leading to the development of Wilms' tumor seem to be complex, any of at least three different genes could be involved separately or jointly. A gene, *WT1*, bearing the properties of a Wilms' tumor suppressor gene has been isolated from the 11p13 region (Call

Table 2. *Expression of* PAX2, PAX8 *and* WT1 *genes in human metanephros*

| gene | condensed mesenchyme | S-bodies | glomeruli |
|---|---|---|---|
| *PAX2* | +++ | – | – |
| *PAX8* | ++ | +++ | – |
| *WT1* | + | +++ | +++ |

et al., 1990; Gessler et al., 1990) and mutations were found in some Wilms' tumors (Haber et al., 1990; Pelletier et al. 1991a,b). In agreement with these findings it has been shown that big DNA deletions at the 11p13 locus are rare in Wilms' tumors (Royer-Pokora et al., 1991). It has been suggested that the *WT1* gene would not be involved in the familial cases of Wilms' tumors due to infertility of the carriers of mutated *WT1* (Pritchard-Jones et al., 1990; van Heyningen et al., 1990). As demonstrated in Fig. 7, *WT1* expression in the human metanephros was present throughout the different stages of glomeruli maturation. Low levels of expression were found in condensed mesenchyme followed by strong expression in the cells of the podocyte layer of S-bodies and glomeruli. These results are in agreement with the pattern of expression published earlier (Pritchard-Jones et al., 1990). Thus, *WT1* expression is comparable to *PAX2* and *PAX8* in the early stages of development, but persists in the glomeruli where *PAX2* and *PAX8* expression is not present anymore (Table 2).

Because Wilms' tumors are believed to arise as a consequence of the failure of the metanephrogenic mesenchyme to differentiate, any gene potentially involved in controlling mesenchymal differentiation is also a potential 'Wilms' tumor gene'. Somatic mutation of both alleles of such a developmental gene could result in failure of the metanephrogenic mesenchyme to differentiate, the possible result being a Wilms' tumor. In this sense, the *PAX8* gene could potentially be involved in the genesis of at least some Wilms' tumors. As a first step in defining possible *PAX8* functions in Wilms' tumors, we have investigated its expression in five different tumors by RNAase protection. *PAX8* is strongly expressed in all five tumors as compared to its expression in adult human kidneys. This fact is probably a reflection of an embryonic origin of the Wilms' tumor. The pattern of differential splicing in the *PAX8* gene described above is similar in all investigated tumors. We are currently characterizing the genomic organization of *PAX8* to perform a mutation analysis of DNA from tumoral and normal tissues of patients with Wilms' tumors.

It is interesting to note that a mouse mutant, Dickie's Small-eye (S*ey*$^{Dey}$), heterozygously deleted for one allele of the *WT1* gene does not develop nephroblastomas as would be expected by analogy with the human WAGR syndrome (Buckler et al., 1991; Glaser et al., 1990). This suggests that there may be differences in murine and human kidney development that are relevant to the initiation of nephroblastoma development. In this respect, it is noteworthy that differential splicing of the serine-rich region in murine Pax8 transcripts could not be detected by RNAase protection although this splicing was easily detectable among human *PAX8* transcripts. This fact could indicate that the ratio between two variants of Pax8 transcripts would be quite different in mouse and man, respectively. However, we did also manage to isolate one mouse cDNA clone with a removed serine-rich region (data not shown).

Thyroid is another site of *PAX8* expression. As in kidney, both types of *PAX8* transcripts can be detected in human thyroid where their corresponding protein products could act in concert with another tissue-specific transcription factor, TTF-1 (Guazzi et al., 1990; Zannini et al., 1992). It has been reported that proto-oncogenes could be activated in thyroid tumors by gene rearrangement to thyroid-specific genes (Grieco et al., 1990). *PAX8* would be a candidate to mediate such thyroid-specific expression. Furthermore, the reports of association of thyroid malignancies with kidney cancers (Gault et al., 1974) should be reevaluated in the light of our results.

Using a somatic cell hybrids system, we have mapped the human *PAX8* gene to chromosome 2q. This finding is in agreement with the mapping of the murine *Pax8* gene to the centromeric region of chromosome 2 (Plachov et al., 1990; Walther et al., 1991), in close vicinity to the surfeit gene cluster (*Surf*), loci marked by defects like the *Danforth's short tail* mutation (*Sd*), the HOX-5 gene cluster and the proto-oncogene Abl. Some of these genes have been identified as part of a region of synteny with the human chromosome 2q (Yon et al. 1989; Stubbs et al., 1990).

The present data make it possible to search for mutations in the human *PAX8* gene. The expected phenotype of such a mutation would be similar to a developmental mutation of the mouse called *Danforth's short tail* mutation (*Sd*; Dunn et al., 1940; Gluecksohn-Waelsch and Rota, 1962) which was proposed to be a candidate for murine *Pax8* mutations (Plachov et al., 1990; Hastie, 1991). The *Sd* mutation is characterized by abnormalities in the axial skeleton and the reduction or absence of kidneys. A comparable pattern of human malformation can be found in subtypes of Klippel-Feil syndrome, which is characterized by vertebral fusion and/or hemivertebrae. More than 59% of patients show urogenital malformations, mostly unilateral or bilateral renal agenesis (Helmi and Pruzansky, 1980).

We thank G. Bell for the human kidney cDNA library and R. Bleackley for the human genomic library. We thank S. Naylor for providing hybrid cell DNAs and D. Housman for the WT 33 cDNA probe. We thank U. Deutsch for the initial Pax8 and Pax2 sequence comparison. We thank R. Di Lauro, P. van der Saag and A. Reeve for transmission of unpublished results. Many thanks go to Andreas Weber and Ralf Altschäffel for the excellent photographic work. The valuble suggestions of R. Di Lauro on the manuscript are gratefully acknowledged. We thank S. Naylor for providing hybrid cell DNAs. We thank D. Housman for providing the WT33 cDNA probe. We thank C. Abbott for providing sequences for *PAX8* primers. Special thanks to F. Wendler, E. Gurtner, and K. Kratochwil. A.P. was supported by a fellowship from the Austrian Academy of Sciences. D.P. was supported by the EMBO long-term fellowship and the austrian FWF grant (P 8682 MOB). H.F. was supported by the Deutsche Forschungsgemeinschaft (Fi 424/1-1).

## References

**Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O. and Milunsky, A.** (1992). An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* **355**, 637-638.

**Baumgartner, S., Bopp, D., Burri, M. and Noll, M.** (1987). Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during Drosophila embryogenesis. *Genes Dev*. **1**, 1247-1267.

**Bell, G. I., Fong, N. M., Stempien, M. M., Wormsted, M. A., Caput, D., Ku, L., Urdea, M. S., Rall, L. B. and Sanchez-Pescador, R.** (1986). Human epidermal growth factor precursor: cDNA sequence, expression in vitro and gene organization. *Nucl. Acid Res.* **14**, 8427-8446.

**Binetruy, B., Smeal, T. and Karin, M.** (1991). Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. *Nature* **351**, 122-127.

**Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M.** (1986). Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell* **47**, 1033-1040.

**Buckler, A. J., Pelletier, J., Haber, D., Glaser, T. and Housman, D. E.** (1991). The murine Wilms' tumor gene (WT1): isolation, characterisation, and expression during kidney development. *Mol. Cell. Biol.* **11**, 4707-4712.

**Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G. and Noll, M.** (1989). Conservation of the paired domain in metazoans and its structure in three isolated human genes. *EMBO J.* **8**, 1183-1190.

**Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C. and Housman, D. E.** (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* **60**, 509-520.

**Chalepakis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M. and Gruss, P.** (1991). The molecular basis of the *undulated/Pax-1* mutation. *Cell* **66**, 873-884.

**Chomczynski, P. and Sacchi, N.** (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.* **162**, 156-159.

**Côté, S., Preiss, A., Haller, J., Schuh R., Kienlin, A., Seifert, E. and Jäckle, H.** (1987). The gooseberry-zipper region of Drosophila: five genes encode different spatially restricted transcripts in the embryo. *EMBO J.* **6**, 2793-2801.

**Deutsch, U., Dressler, G. R. and Gruss, P.** (1988). *Pax 1*, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53**, 617-625.

**Dowdy, S. F., Fasching C. L., Araujo, D., Lai, K.-M., Livanos, E., Weissman, B. E. and Stanbridge, E. J.** (1991). Suppression of tumorigenicity in Wilms' tumor by the p15.5-p14 region of chromosome 11. *Science* **254**, 293-295.

**Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O. and Gruss, P.** (1990). *Pax2*, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109**, 787-795.

**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.

**Dunn, L. C., Gluecksohn-Schoenheimer, S. and Bryson, V.** (1940). A new mutation in the mouse affecting spinal column and urogenital system. *J. Hered.* **31**, 343-348.

**Ekblom, P., Lehtonen, E., Saxen, L. and Timpl, R.** (1981). Shift in collagen type as an early response to induction of the metanephric mesenchyme. *J. Cell Biol.* **89**, 276-283.

**Epstein, D. J., Vekemans, M. and Gros, P.** (1991). *Splotch ($Sp^{2H}$)*, a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax3*. *Cell* **67**, 767-774.

**Feinberg, A. P. and Vogelstein, B.** (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.

**Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M.** (1986). Structure of the segmentation gene paired and the Drosophila PRD gene set as part of a gene network. *Cell* **47**, 735-746.

**Gault, E. W., Leung, T. H. W. and Thomas, D. P.** (1974). Clear cell renal carcinoma masquerading as thyroid enlargement. *J. Path.* **113**, 21-25.

**Gehring, W. J., Müller, M., Affolter, M., Percival-Smith, Billeter, M., Qian, Y. Q., Otting, G. and Wüthrich, K.** (1990). The structure of the homeodomain and its functional implications. *Trends Genet.* **6**, 323-329.

**Gerster, T., Balmaceda, G.-G. and Roeder R. G.** (1990). The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J.* **9**, 1635-1643.

**Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H. and Bruns, G. A. P.** (1990). Homozygous deletion in Wilms' tumors of a zinc-finger gene identified by chromosome jumping. *Nature* **343**, 774-778.

**Ghosh, S. and Baltimore, D.** (1990). Activation in vitro of NF-κB by phosphorylation of its inhibitor IκB. *Nature* **344**, 678-682.

**Glaser, T., Jones, C., Douglass, E. C. and Housman, D.** (1989). Constitutional and somatic mutations of chromosome 11p in Wilms' tumor. *Cancer Cells* 7/*Molecular Diagnostics of Human Cancer* CSHL Press.

**Glaser, T., Lane, J. and Housman, D.** (1990). A mouse model of the aniridia-Wilms tumor deletion syndrome. *Science* **250**, 823-827.

**Gluecksohn-Waelsch, S. and Rota, T. R.** (1962). Development in organ tissue culture of kidney rudiments from mutant mouse embryos. *Dev. Biol.* **7**, 432-444.

**Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. and Gruss, P.** (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.* **10**, 1135-1147.

**Grieco, M., Santoro, M., Berlingieri, M. T., Melillo, R. M., Donghi, R., Bongarzone, I., Pierotti, M. A., Porta, G. D., Fusco, A. and Vecchio, G.** (1990). PTC is a novel rearranged form of the rat proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. *Cell* **60**, 557-563.

**Grundy, P., Koufos, A., Morgan, K., Li, F. P., Meadows, A. T. and Cavenee, W. K.** (1988). Familial predisposition to Wilms' tumour does not map to the short arm of chromosome 11. *Nature* **336**, 374-376.

**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.

**Haber, D. A., Buckler, A. J., Glaser, T., Call, K. M., Pelletier, J., Sohn, R. L., Douglass, E. C. and Housman, D. E.** (1990). An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. *Cell* **61**, 1257-1269.

**Hastie, N. D.** (1991). *Pax* in our time. *Current Biol.* **1**, 342-344.

**Hatzopoulos, A. K., Stoykova, A. S., Erselius, J. R., Goulding, M., Neuman, T. and Gruss, P.** (1990). Structure and expression of the mouse Oct2a and Oct2b, two differentially spliced products of the same gene. *Development* **109**, 349-362.

**van Heyningen, V., Bickmore, W. A., Seawright, A., Fletcher, J. M., Maule. J., Fekete, G., Gessler, M., Bruns, G. A. P., Huerre-Jeanpierre, C., Junien, C., Williams, B. R. G. and Hastie, N. D.** (1990). Role for the Wilms tumor gene in genital development. *Proc. Natl. Acad. Sci. USA* **87**, 5383-5386.

**van Heyningen, V. and Hastie, N. D.** (1992). Wilms' tumor: reconciling genetics and biology. *Trends Genet.* **8**, 16-21.

**Helmi, C. and Pruzansky, S.** (1980). Craniofacial and extracranial malformations in the Klippel-Feil syndrome. *Cleft Palate J.* **17**, 65-88.

**Hill, R. E., Favor, J., Hogan, B. L. M., Ton, C. C. T., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D. and van Heyningen, V.** (1991). Mouse *Small eye* results from mutations in a paired-like homeobox-containing gene. *Nature* **354**, 522-525.

**Huff, V., Compton, D. A., Chao, L. -Y., Strong, L. C., Geiser, C. F. and Saunders, G. F.** (1988). Lack of linkage of familial Wilms' tumour to chromosomal band 11p13. *Nature* **336**, 377-378.

**Jostes, B., Walther, C. and Gruss, P.** (1990). The murine paired box gene, Pax 7, is expressed specifically during the development of the nervous and muscular system. *Mech. Dev.* **33**, 27-38.

**Kapiloff, M. S., Farkash, Y., Wegner, M. and Rosenfeld, M. G.** (1991). Variable effects of phosphorylation of Pit-1 dictated by the DNA response elements. *Science* **253**, 786-789.

**Kessel, M. and Gruss, P.** (1990). Murine development control genes. *Science* **249**, 374-379.

**Koufos, A., Grundz, P., Morgan, K., Aleck, K. A., Hadro, T., Lampkin, B. C., Kalbakji, A. and Cavenee, W.** (1989). Familial Wiedemann-Beckwith syndrome and a second Wilms' tumor locus both map to 11p15.5. *Am. J. Hum. Genet.* **44**, 711-719.

**Kozak, M.** (1986). Point mutations define a sequence flanking AUG initiator codons that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.

**Landschulz, W. H., Johnson P. F. and McKnight S. L.** (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.

**Laurie, G. W., Horikoshi, S., Killen, P. D., Segui-Real, B. and Yamada, Y.** (1989). In situ hybridization reveals temporal and spatial changes in cellular expression of mRNA for laminin receptor, laminin, and basement

membrane (type IV) collagen in the developing kidney. *J. Cell Biol.* **109**, 1351-1362.

**Lee, C. Q., Yun, Y., Hoeffler, J. P. and Habener, J. F.** (1990). Cyclic-AMP-responsive transcriptional activation of CREB-327 involves interdependent phosphorylated subdomains. *EMBO J.* **9**, 4455-4465.

**Lopez, A. J. and Hogness, D. S.** (1991). Immunochemical dissection of the Ultrabithorax homeoprotein family in Drosophila melanogaster. *Proc. Natl. Acad. Sci. USA* **88**, 9924-9928.

**Madden, S. L., Cook, D. M., Morris, J. F., Gashler, A., Sukhatme, V. P. and Rauscher III, F. J.** (1991). Transcriptional repression mediated by the WT1 Wilms tumor gene product. *Science* **253**, 1550-1553.

**Mermod, N., O'Neill, Kelly, T. J. and Tjian, R.** (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **58**, 741-753.

**Mierau, G. W., Beckwith, J. B. and Weeks, D. A.** (1987). Ultrastructure and histogenesis of the renal tumors of childhood: an overview. *Ultrast. Pathol.* **11**, 313-333.

**Mumberg, D., Lucibello, F. C., Schuermann, M. and Müller, R.** (1991). Alternative splicing of *fosB* transcripts results in differentially expressed mRNAs encoding functionally antagonistic proteins. *Genes Dev.* **5**, 1212-1223.

**Mundlos, S., Meyer, R., Yamada, Y. and Zabel, B.** (1991). Analysis of cartilage proteoglycan core protein (aggrecan) and link protein gene expression during human skeletal development by in situ hybridization. *Matrix* **11**, 339-346.

**Mundlos, S., Schwahn, B., Reichert, T. and Zabel, B.** (1992). Distribution of osteonectin mRNA and protein during human embryonic and fetal development. *J. Histochem. Cytochem.* **40**, 283-291.

**Nakabeppu, Y. and Nathans, D.** (1991). A naturally occurring truncated form of FosB that inhibits Fos/Jun transcriptional activity. *Cell* **64**, 751-759.

**Nicosia, A., Monaci, P., Tomei, L., De Francesco, R., Nuzzo, M., Stunnenberg, H. and Cortese, R.** (1990). A myosin-like dimerization helix and an extra-large homeodomain are essential elements of the tripartite DNA binding structure of LFB1. *Cell* **61**, 1225-1236.

**Nornes, H. O., Dressler, G. R., Knapik, E. W., Deutsch, U. and Gruss, P.** (1990). Spatially and temporally restricted expression of Pax2 during murine neurogenesis. *Development* **109**, 797-809.

**Opstelten, D.-J. E., Vogels, R., Robert, B., Kalkhoven, E., Zwartkruis, F., de Laaf, L., Destrée O. H., Deschamps, J., Lawson, K. A. and Meijlink, F.** (1991). The mouse homeobox gene, S8, is expressed during embryogenesis predominantly in mesenchyme. *Mech. Dev.* **34**, 29-42.

**Pelletier, J., Bruening, W., Li, F., Haber, D. A., Glaser, T. and Housman, D. E.** (1991a). WT1 mutations contribute to abnormal genital system development and hereditary Wilms' tumour. *Nature* **353**, 431-434.

**Pelletier, J., Bruening, W., Kashtan, C. E., Mauer, S. M., Manivel, J. C., Striegel, J. E., Houghton, D. C., Junien, C., Habib, R., Fouser, L., Fine, R. N., Silverman, B. L., Haber, D. A. and Housman, D.** (1991b). Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* **67**, 437-447.

**Plachov, D., Chowdhury, H., Walther, C., Simon, D., Guenet, J. -L. and Gruss, P.** (1990). Pax 8, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* **110**, 643-651.

**Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., Housman, D., van Heyningen, V. and Hastie, N.** (1990). The candidate Wims' tumor gene is involved in genitourinary development. *Nature* **346**, 194-197.

**Proudfoot, N. J. and Brownlee, G. G.** (1976). The 3′ non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**, 211-214.

**Ptashne, M.** (1988). How eukaryotic transcriptional activators work. *Nature* **335**, 683-689.

**Pulverer, B. J., Kyriakis, J. M., Avruch, J., Nikolakaki, E. and Woodgett, J. R.** (1991). Phosphorylation of *c-jun* mediated by MAP kinases. *Nature* **353**, 670-674.

**Reeve, A. E., Sih, S. A., Raizis, A. M. and Feinberg, A. P.** (1989). Loss of allelic heterozygosity at a second locus on chromosome 11 in sporadic Wilms' tumor cells. *Mol. Cell. Biol.* **9**, 1799-1803.

**Roman, C., Cohn, L. and Calame, K.** (1991). A dominant negative form of transcription activator mTFE3 created by differential splicing. *Science* **254**, 94-97.

**Rosenfeld, M. G.** (1991). POU-domain transcription factors: pou-er-ful developmental regulators. *Genes Dev.* **5**, 897-907.

**Royer-Pokora, B., Ragg, S., Heckl-Östreicher, B., Held, M., Loos, U., Call, K., Glaser, T., Housman, D., Saunders, G., Zabel, B., Williams, B. and Poustka, A.** (1991). Direct pulsed field gel electrophoresis of Wilms' tumors shows that DNA deletions in 11p13 are rare. *Genes, Chrom. Canc.* **3**, 89-100.

**Sanger, F., Nicklen, S. and Coulson, A. R.** (1977). DNA sequencing with chain termination inhibitors, *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

**Saxén, L.** (1987). *Organogenesis of the Kidney.* Cambridge University Press.

**Schwartz, C. E., Haber, D. A., Stanton, V. P., Strong, L. C., Skolnick, M. H. and Housman, D. E.** (1991). Familial predisposition to Wilms tumor does not segregate with the WT1 gene. *Genomics* **10**, 927-930.

**Sheng, M., Thompson, M. A. and Greenberg, M. E.** (1991). CREB: A $Ca^{2+}$-regulated transcription factor phosphorylated by calmodulin-dependent kinases. *Science* **252**, 1427-1430.

**Smeal, T., Binetruy, B., Mercola, D. A., Birrer, M. and Karin, M.** (1991). Oncogenic and transcriptional cooperation with Ha-Ras requires phosphorylation of c-Jun on serines 63 and 73. *Nature* **354**, 494-496.

**Stamminger, T., Fickenscher, H. and Fleckenstein, B.** (1990). Cell type-specific induction of the major immediate early enhancer of human cytomegalovirus by cyclic AMP. *J. Gen. Virol.* **71**, 105-113.

**Struhl, G., Struhl, K. and Macdonalds, P. M.** (1989). The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57**, 1259-1273.

**Stubbs, L., Huxley, C., Hogan, B., Evans, T., Fried, M., Duboule, D. and Lehrach, H.** (1990). The HOX-5 and surfeit gene clusters are linked in the proximal portion of mouse chromosome 2. *Genomics* **6**, 645-650.

**Tanaka, M. and Herr, W.** (1990). Differential transcriptional activation by oct-1 and oct-2: interdependent activation domains induce oct-2 phosphorylation. *Cell* **60**, 375-386.

**Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T.** (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**, 635-636.

**Theill, L. E., Castrillo, J. -L., Wu, D. and Karin M.** (1989). Dissection of functional domains of the pituitary-specific transcription factor GHF-1. *Nature* **342**, 945-948.

**Treisman, J., Harris, E. and Desplan, C.** (1991). The paired box encodes a second DNA-binding domain in the Paired homeo domain protein. *Genes Dev.* **5**, 594-604.

**Ton, C. C. T., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., van Heyningen, V., Hastie, N. D., Meijers-Heijboer, H., Drechsler, M., Royer-Pokora, B., Collins, F., Swaroop, A., Strong, L. C. and Saunders, G. F.** (1991). Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* **67**, 1059-1074.

**Walther, C. and Gruss, P.** (1991). *Pax-6*, a murine paired box gene, is expressed in the developing CNS. *Development* **113**, 1435-1449.

**Walther, C., Guenet, J. -L., Simon, D., Deutsch, U., Jostes, B., Goulding, M. D., Plachov, D., Balling, R. and Gruss, P.** (1991). Pax: a murine multigene family of paired box-containing genes. *Genomics* **11**, 424-434.

**Weissman, B. E., Saxon, P. J., Pasquale, S. R., Jones, G. R., Geiser, A. G. and Stanbridge, E. J.** (1987). Introduction of a normal human chromosome 11 into a Wilms' tumor cell line controls its tumorigenic expression. *Science* **236**, 175-180.

**Yon, J., Palmer, R. W., Sheer, D. and Fried, M.** (1989). Localization of the Surfeit gene cluster containing the ribosomal protein gene S7a to chromosome bands 9q. *Ann. Hum. Genet.* **53**, 149-155.

**Zabel, B. U., Eddy, R. L., Lalley P. A., Scott, J. and Shows, T. B.** (1985). Chromosomal locations of the human and mouse genes for precursors of epidermal growth factors and the β-subunit of nerve growth factor. *Proc. Natl. Acad. Sci. USA* **82**, 469-473.

**Zannini, M., Francis-Lang, H., Plachov, D. and Di Lauro, R.** (1992). PAX-8, a paired domain containing protein, binds to a sequence overlapping with the recognition site of an homeodomain and activates transcription from two thyroid-specific promoters. *Mol. Cell. Biol.* (in press).

*(Accepted 3 June 1992)*