Development 126, 2539-2550 (1999)
Printed in Great Britain © The Company of Biologists Limited 1999
DEV2400

Developmental expression of *Pax1/9* genes in urochordate and hemichordate gills: insight into function and evolution of the pharyngeal epithelium

Michio Ogasawara[1,*], Hiroshi Wada[2], Heiko Peters[3] and Nori Satoh[1]

[1]Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
[2]Seto Marine Biological Laboratory, Graduate School of Science, Kyoto University, Shirahama, Wakayama 649-2211, Japan
[3]GSF Research Center, Institute for Mammalian Genetics, Ingolstaedter Landstrasse 1, 85758 Oberschleissheim, Germany
*Author for correspondence (e-mail: ogasawara@ascidian.zool.kyoto-u.ac.jp)

Accepted 24 February; published on WWW 4 May 1999

SUMMARY

The epithelium of the pharynx contributes to the formation of gills in hemichordates, urochordates, cephalochordates and primitive vertebrates, and is therefore a key structure for understanding developmental mechanisms underlying the establishment of chordate body plans. *Pax1*- and *Pax9*-related genes encode transcription factors which are expressed in the pharyngeal region of cephalochordates as well as in the vertebrate pharyngeal pouch epithelium that forms the thymus and parathyroid glands. To explore the molecular basis underlying the occurrence and modifications of the pharyngeal epithelium during evolution, we isolated cDNA clones for *Pax1*- and *Pax9*-related genes of urochordates (*HrPax1/9* of *Halocynthia roretzi* and *CiPax1/9* of *Ciona intestinalis*) and a hemichordate (*PfPax1/9* of *Ptychodera flava*) from gill cDNA libraries. Each gene is present as a single copy per haploid genome. All of the cDNAs encode typical paired domains and octapeptides but not a homeodomain, as is also true of other *Pax1*- and *Pax9*-related genes. Molecular phylogenetic analysis based on comparison of the paired domain amino-acid sequences suggests that *HrPax1/9*, *CiPax1/9* and *PfPax1/9* belong to the *Pax1/9* subfamily, and that they are descendants of a single precursor of *Pax1/Pax9*. Screening of *HrPax1/9* cDNA clones yielded six different types of transcripts which were generated by alternative splicing. Northern blot, RT-PCR/Southern and in situ hybridization analyses revealed that *HrPax1/9*, *CiPax1/9* and *PfPax1/9* are not expressed during early embryogenesis but are expressed in the epithelia of differentiating gills, suggesting that these genes encode gill-specific transcription factors. The *Pax1/9* genes therefore might provide the first developmental genetic corroboration of hypotheses of organ-level homology that unifies hemichordates, urochordates and cephalochordates.

Key words: *Pax1/9*, Urochordate, Hemichordate, Gill, Chordate body plan

INTRODUCTION

Recent elucidation of the expression and function of developmental genes in different animal groups has made it possible to determine the molecular developmental mechanisms underlying the evolution of various animal body plans. Among numerous evolutionary developmental processes, we have been focusing on the evolution of chordate body plans (Satoh and Jeffery, 1995; Satoh, 1995; Yasuo et al., 1995). The phylum Chordata comprises the subphylum groups urochordates, cephalochordates and vertebrates (cf. Fig. 8). They share several characteristic features, including a notochord, a dorsal hollow neural tube, pharyngeal gill slits and an endostyle (e.g., Brusca and Brusca, 1990; Nielsen, 1995; Gee, 1996; Hall, 1998). In addition, chordates form a superphyletic group of deuterostomes together with two other non-chordate invertebrate groups, echinoderms and hemichordates (cf. Fig. 8). Recent molecular phylogenic studies support the idea that echinoderms, hemichordates and chordates form a monophyletic group (Wada and Satoh, 1994; Turbeville et al., 1994). Therefore, chordate body plans evolved from the ancestor by developing their characteristic features mentioned above. This suggests that the elucidation of molecular developmental mechanisms underlying the formation of these organs will lead to a better understanding of the evolution of chordate body plans. Several studies have focused on developmental mechanisms underlying the formation of the notochord (Yasuo and Satoh, 1993; Harada et al., 1995; Tagawa et al., 1998) and neural tube (e.g., Wada et al., 1998). For example, *Brachyury* is a key developmental gene in the formation of the notochord (reviewed by Di Gregorio and Levine, 1998; Satoh et al., 1999).

The pharyngeal gill slits and endostyle are organs characteristic of filter feeding, which might has arisen when primitive chordates shifted to internal feeding. Therefore, the pharyngeal gill is also an important structure for understanding the evolution of chordate body plans. In a previous study, we characterized two pharyngeal gill (endodermally derived

epithelium of the gill)-specific genes from the ascidian *Halocynthia roretzi*; namely, *HrPhG1* and *HrPhG2*, which encode secretory proteins related to feeding (Tanaka et al., 1996). The expression of these genes is first detected during metamorphosis when the pharyngeal gill is formed, and the gene expression is maintained specifically in the differentiated pharyngeal gill. Therefore, these genes are useful as pharyngeal gill differentiation markers. In the differential screening conditions used in the previous study, however, we could not isolate any genes for transcription factors which are expressed in the pharyngeal gill.

Pax genes constitute a large family of developmental regulatory genes that encode nuclear transcription factors with the DNA-binding motif, the paired domain (reviewed by Strachan and Read, 1994; Dahl et al., 1997). This motif is highly conserved between vertebrate and invertebrate *Pax* genes. In mammalians, nine *Pax* genes (*Pax1* to *Pax9*) have been isolated, which are classified into four paralogous groups according to similarities within structural motifs. *Pax1* and *Pax9* are categorized as the Group I subfamily, which encodes polypeptides with the paired domain and octapeptide but without a homeodomain. However, Group II (*Pax2*, *Pax5* and *Pax8*), Group III (*Pax3* and *Pax7*) and Group IV (*Pax4* and *Pax6*) encode the paired domain and homeodomain. To date, *Pax1/Pax9* genes have been isolated from human, mouse, chick, zebrafish, amphioxus and *Drosophila*. A *Drosophila Pax1/Pax9*-related gene, *Pox meso*, is segmentally expressed in somatic mesoderm at the late phase of germ band elongation (Bopp et al., 1989). In mice, *Pax1* and *Pax9* are expressed in segmented patterns in sclerotomes and later in the intervertebral disks along the body axis, and their expression is required for vertebral column development (Deutsch et al., 1988; Koseki et al., 1993; Wallin et al., 1994; Neubüser et al., 1995). Furthermore, these genes are expressed in the endodermally derived epithelium of the pharyngeal pouches, which form thymus, parathyroid glands and ultimobranchial bodies, and the expression of *Pax9* is required for the formation of these organs (Neubüser et al., 1995; Wallin et al., 1996; Peters et al., 1998). In addition, the cephalochordate amphioxus contains only one *Pax1*/*Pax9* ancestral gene (*AmphiPax1*), which is not expressed in mesodermal tissues such as somites but is expressed in the endodermal cells of the pharynx soon after hatching, and later specifically in the pharyngeal gill (Holland and Holland, 1995). The pharyngeal pouch is thought to be homologous to the pharyngeal gill. The conserved expression of *Pax1/Pax9*-related genes in the epithelium of the pharyngeal pouches suggests that *Pax1/Pax9*-related genes may have a key role in the formation of the pharyngeal gill.

Is *Pax1/Pax9*-related gene expression in the pharyngeal endoderm restricted to cephalochordates and vertebrates, or does it extend to all chordates? This may be answered by determination of urochordate *Pax1/Pax9*-related genes. Furthermore, does it also extend to other deuterostomes? From the time of Bateson (1886), hemichordates have been liked to chordates because they share chordate-specific structures and organizations. The pharyngeal slits are also found in hemichordates, but there is controversy about whether this is a true homology or convergent evolution (e.g., Gee, 1996; Hall, 1998), together with the putative pharyngeal slits in some of the calcichordates (Jefferies, 1986; Peterson, 1995). We reasoned that if the *Pax1/Pax9*-related gene is expressed in the hemichordate gill, this might provide the first molecular evidence for the pharyngeal slits being homologous in hemichordates and chordates. The present study therefore attempted to isolate cDNA clones for *Pax1/Pax9*-related genes from urochordates and hemichordate.

MATERIALS AND METHODS

Biological materials

The ascidian *Halocynthia roretzi* was purchased from fishermen near the Otsuchi Marine Research Center, Ocean Research Institute, University of Tokyo, Iwate, Japan. Another ascidian, *Ciona intestinalis*, was collected near the Marine BioSource Education Center of Tohoku University, Onagawa, Miyagi, Japan. The hemichordate acorn worm, *Ptychodera flava*, was collected at the sand bar, Kaneohe Bay, Oahu, Hawaii, and maintained in the Kewalo Marine Laboratory of the University of Hawaii (Tagawa et al., 1998). After dissection of adults, tissues and organs were quickly frozen in liquid nitrogen. The specimens were kept at −80˚C until use.

H. roretzi eggs were fertilized and raised in filtered seawater at about 12°C. Tadpole larvae hatched at about 40 hours of development. They were allowed to undergo metamorphosis naturally. Juveniles that adhered to plastic dishes were cultured for about 1 month in aquaria with circulating natural seawater. *C. intestinalis* eggs were fertilized and raised at about 20°C. Tadpole larvae hatched at about 16 hours of development. After natural metamorphosis, juveniles that adhered to plastic dishes were cultured for about 1 month. Samples at appropriate developmental stages were frozen in liquid nitrogen for northern blot analysis or fixed for in situ hybridization.

Isolation of RNAs and construction of cDNA libraries

Total RNAs were extracted from frozen samples of *C. intestinalis* pharyngeal gill and *P. flava* pharyngeal gill by the AGPC method (Chomczynski and Sacchi, 1987). Poly(A)$^+$ RNAs were purified by use of Oligotex dT30 beads (Roche Japan). Complementary DNA was synthesized and constructed as described in a previous report (Ogasawara et al., 1996). A pharyngeal gill cDNA library of *C. intestinalis* and gill cDNA library of *P. flava* was constructed using a uni-ZAP-II vector (Stratagene).

Isolation and sequencing of cDNA clones for *Pax1/9* genes

Amino acid sequences of the paired domain of *Pax1/Pax9* gene products are highly conserved among human, mouse, chick, *Xenopus*, zebrafish and *Drosophila* (cf. Fig. 2). The sense oligonucleotide Pax-uniF: GTN AA(CT) CA(AG) (CT)TN GGN GGN GTN TT, which corresponds to the amino acid sequence VNQLGGVF, and the antisense oligonucleotide Pax1/9-R2: GG NAC (AG)TT (AG)TA (TC)TT (AG)TC (AG)CA, which corresponds to the amino acid sequence LLSDCICD, were synthesized. Using these oligonucleotides as primers we amplified target fragments from the *H. roretzi* genomic DNA by means of PCR. Target fragments of *C. intestinalis* and *P. flava* were amplified by means of RT-PCR using pharyngeal gill poly(A)$^+$ RNAs. Probing with candidate 318 bp DNA fragments, which were random labeled with [^{32}P]dCTP (Amersham), full length cDNA clones were isolated from pharyngeal gill cDNA libraries of *H. roretzi* (Tanaka et al., 1996), *C. intestinalis* (Ogasawara and Satoh, 1998) and *P. flava* (the present study) under high stringency conditions (hybridization: 6× SSPE, 0.1% SDS, 1× Denhardt's solution, 50% formamide at 42°C for 16 hours; washing: 2× SSC-0.1% SDS at 55°C for 30 minutes, 1× SSC-0.1% SDS at 55°C for 30 minutes, 0.1× SSC-0.1% SDS at 55°C for 30 minutes). Isolated clones were sequenced using an ABI PRISM 377 DNA Sequencer (Perkin Elmer).

Sequence comparisons and molecular phylogenetic analyses

Sequences were aligned using SeqQpp 1.9 manual aligner for

Macintosh (Gilbert, 1993). Phylogenetic analyses were performed using amino acid sequences of the paired domain. Estimation of molecular phylogeny was carried out by the neighbor-joining method (Saitou and Nei, 1987) using CLUSTAL V program (Higgins et al., 1992). Confidence in the phylogeny was assessed by bootstrap resampling of the data (×1000) (Felsenstein, 1985).

Northern blot analysis

After extraction of total RNA by the AGPC method, poly(A)$^+$ RNA was purified by use of Oligotex dT30 beads. Northern blot hybridization was carried out by the standard procedure (Sambrook et al., 1989) and filters were washed under high stringency conditions. DNA probes for blot hybridizations were labeled with [^{32}P]dCTP using a random primed labeling kit (Boehringer Mannheim).

RT-PCR/Southern blot analysis

Total RNAs were extracted from small pieces of adult tissues that were sonicated in the extraction buffer containing guanidinium thiocyanate. After proteinase K and phenol-chloroform treatment, poly(A)$^+$ RNA was purified with Oligotex dT30 beads. 100 ng of poly(A)$^+$ RNA was reverse-transcribed using an oligo dT primer, and then PCR was performed with specific primers. Amplified fragments were electrophoresed and blotted onto Hybond-N+ nylon membranes (Amersham). The blots were hybridized with random-primed ^{32}P-labeled DNA probes at 42°C for 16 hours and washed under high-stringency conditions.

Genomic Southern analysis

High molecular weight genomic DNA of *H. roretzi* was extracted from a single adult by a standard procedure (Sambrook et al., 1989). After exhaustive digestion with *Bam*HI, *Eco*RI, *Hin*dIII and *Pst*I, and 0.7% agarose gel electrophoresis, the DNA fragments were blotted onto Hybond-N+ nylon membranes (Amersham). The blots were hybridized with random-primed ^{32}P-labeled DNA probes at 42°C for 16 hours and washed under high-stringency conditions.

In situ hybridization

Ascidian juveniles and adult tissues, and branchial region of adult acorn worms were fixed as whole-mount specimens in 4% paraformaldehyde in 0.5 M NaCl, 0.1 M MOPS buffer at 4°C for 12 hours. For ascidian adults, the tunic was stripped off with tungsten needles during fixation in the solution described above (Ogasawara and Satoh, 1998). After dehydration with a graded series of alcohol solutions, specimens were embedded in polyester wax (BDH Chem. Ltd) and sectioned at 6 μm. In situ hybridization of whole-mount and sectioned specimens was carried out basically as described previously (Ogasawara et al., 1996). Probes were synthesized by following the instructions from the supplier of the DIG RNA Labeling kit (Boehringer Mannheim).

RESULTS

Isolation and characterization of cDNA clones for *Halocynthia roretzi Pax1/9* gene

Using degenerate oligonucleotide primers corresponding to the conserved paired box sequence of vertebrate *Pax9*, we amplified target fragments from *H. roretzi* genomic DNA by means of PCR. Sequencing revealed that an amplified fragment contained a *Pax1/Pax9* type paired box. Then, with the random-labeled fragment as a probe, we analyzed the expression of the corresponding gene in adult tissues by northern blotting (data not shown). This northern blotting indicated that the ascidian *Pax1/Pax9*-related gene was predominantly expressed in the pharyngeal gill. Therefore, we screened 3.0×10^5 pfu of an *H. roretzi* pharyngeal gill cDNA library (Tanaka et al., 1996) using the same probe. We found that 53 cDNA clones covered almost the full-length of this gene transcript.

Under our screening conditions, we isolated six different types of cDNAs (type-I to -VI). Nucleotide and predicted amino acid sequences of the type-I cDNA are shown in Fig. 1A. The clone consisted of 3,117 nucleotides excluding the poly(A) tail. This almost full-length sequence contained a single open reading frame of 1,536 bp that predicted a polypeptide of 512 amino acids, and 1,401 bp of 3′ non-coding region, including the potential signal for the polyadenylation site. Like other *Pax1/Pax9* gene products, the polypeptide had the paired type DNA-binding domain and octapeptide (Fig. 2A,B), but not a homeo type DNA-binding domain. The DNA sequence and primary structure of the amino acid sequence therefore suggested that this clone encodes a polypeptide of the Pax1/Pax9 subfamily. We named this gene *HrPax1/9* (*H. roretzi Pax1* and *Pax9*), and also named this clone *HrPax1/9-I* (*HrPax1/9* type-I). The calculated molecular mass of the HrPax1/9-I polypeptide was 55.5 kDa.

Fig. 1C illustrates the structures of *HrPax1/9* type-I to -VI cDNAs. Type-II cDNA contained a different 5′-end sequence (Fig. 1B; exon 2 in Fig. 1C). The predicted N-terminal amino acid sequence was shorter than that of the others (exon 2; Fig. 1C). Because the nucleotide and amino acid sequences of the paired domain and octapeptide of the type-II cDNA were identical to those of the other types, we named this clone *HrPax1/9-II*. The calculated molecular mass of the HrPax1/9-II polypeptide was 51.7 kDa, slightly smaller than that of the type-I polypeptide. Type-III cDNA contained a different 3′-end sequence (exon 8 and noncoding region) without the potential signal for polyadenylation (Fig. 1C). This cDNA therefore encoded a polypeptide with a different C-terminal amino acid sequence. The length of this clone was about 2.3 kb, which was shorter than the others. The calculated molecular mass of the HrPax1/9-III polypeptide was 54.9 kDa. Type-IV cDNA had almost the same structure as type-I cDNA, but 13 nucleotides (exon 7) were missing in the C-terminal region. Therefore, a frame sift occurred in exon 9, resulting in a different C-terminal amino acid sequence. The calculated molecular mass of the HrPax1/9-IV polypeptide was 54.9 kDa. The length of the type-V cDNA was about 3.2 kb, almost the same as those of type-I, -II and -VI, although their structures were different. The type-V clone had an 827 bp insertion (exon 5 and noncoding region) in the C-terminal half region of the type-III sequence. Therefore, the type-V open reading frame was short. The calculated molecular mass of the HrPax1/9-V polypeptide was 40.7 kDa, smaller than the others. The type-VI clone was about 4.0 kb, the longest of the six types. Its open reading frame was the same as that of type-III. The regions of exons 6, 7, 8 and 9 were changed into a 3′ UTR. The molecular mass of the HrPax1/9-VI polypeptide was 55.5 kDa. The sequences near the paired domain and octapeptide were identical among all of the six types (exon 3).

Northern blotting of *HrPax1/9* indicated a major transcript of about 3.2 kb (type-I, -II, -IV and -V) and a minor transcript of about 4.0 kb (type-VI) (Fig. 3A,C). We could not detect a transcript of about 2.7 kb (type-III) under our hybridization conditions. In the western blotting analysis, several signals were detected around 55 kDa (data not shown). To determined whether the transcripts of different size were derived from a

single gene, we carried out genomic Southern blot analysis. The region of the paired box shared by all types was used as a probe. As shown in Fig. 1D, only one single band was detected in the lanes of the *Bam*HI (about 20.0 kb), *Eco*RI (about 1.7 kb), *Hin*dIII (about 7.0 kb) and *Pst*I (about 2.9 kb) digests under high stringency conditions. This result suggests that *HrPax1/9* is present as a single copy per haploid genome of *H. roretzi*. Therefore, we concluded that *HrPax1/9* is present as a single copy and thus different types of transcript were produced by alternative splicing. In further support of this notion, all exons were located in the same locus as determined by genome sequencing (data not shown).

Molecular evolution of *Pax1/9* subfamily genes

Although the overall degree of amino acid identity of *HrPax1/9*-encoded proteins with other *Pax1*/*Pax9* encoded proteins was not high, the amino acid sequence was highly conserved in the paired domain. Fig. 2A shows the comparison of the paired domain amino-acid sequences of Pax1/9 subfamily members. The HrPax1/9 paired-domain amino-acid sequence was highly homologous to those of human Pax1 (90.6%), human Pax9 (90.6%), mouse Pax1 (90.6%), mouse Pax9 (91.4%), chick Pax1 (91.4%), chick Pax9 (90.6%), zebrafish Pax9 (89.1%), amphioxus Pax1 (89.8%), ascidian CiPax1/9 (93.0%; described below), acorn worm PfPax1/9 (91.4%; described below) and *Drosophila* Pox meso (84.4%). The vertebrate Pax1/Pax9s showed more than 90% identity among themselves. The characteristic octapeptide found in every Pax1/9 was also found in HrPax1/9, CiPax1/9 and PfPax1/9.

The evolutionary relationship of *Pax1/9* genes was examined by molecular phylogenetic analysis. A molecular phylogenetic tree was constructed by means of the neighbor-joining method based on the paired domain sequence, because only the paired domain sequences could be aligned and used for calculations. The *Drosophila* Pox meso was used as an outgroup. As shown in Fig. 2C, the tree clearly indicated that HrPax1/9, CiPax1/9 and PfPax1/9 belong to the Pax1/9 subfamily. Furthermore, the tree demonstrated that HrPax1/9, CiPax1/9 and PfPax1/9 are descendants of a single precursor and that the branching of Pax1 and Pax9 occurred in the vertebrate lineage.

HrPax1/9 is expressed in the pharyngeal gill

The temporal and spatial expression patterns of *HrPax1/9* were examined by northern blot analysis and RT-PCR/Southern analysis. Poly(A)$^+$ RNA prepared from the pharyngeal gill, endostyle, body-wall muscle, digestive gland, gonad and intestine were subjected to hybridization. Hybridization signals of 3.2 kb and 4.0 kb were detected (Fig. 3A). The size of the major transcript, about 3.2 kb, coincided with the sizes of the cDNAs of type -I, -II, -IV and -V, whereas the size of the minor

A

```
1     GCTAAGCGTTATTCAAACCGAGAAGCTGGTGAACACACAACCACGCGTCAGCATAAACTAGAGAAGCAGAGTACCCCATTGAAATCAAATCGTTATTAGCATCTTGTGACTGACTTTTTT  120

121   AACTTTGGCATTGAAGTTAAGTCTTCCGGATTAGGAAGGCTATACAATATATTCGGCATCATGATGCAATCACCGACCCAAAATCAGCCGGTTACCTTTAGTTCTCCGCATCCTGGAGCA  240
1                                                                    M  M  Q  S  P  T  Q  N  Q  P  V  T  F  S  S  P  H  P  G  A   20

241   GGAGGAATTGTCTCTATCATGCCAAATTCGACCAACAGTGCCCAAATGTTTCCCCAAAACGGCATCTACACGAATGTGGTCCGATCTACAGAGCAAACGTTCGGCGAAGTCAACCAGCTT  360
21    G  G  I  V  S  I  M  P  N  S  T  N  S  A  Q  (M) F  P  Q  N  G  I  Y  T  N  V  V  R  S  T  E  Q [T  F  G  E  V  N  Q  L]  60

361   GGTGGGATATTTGTCAACGGTCGCCCGCTACCGAACACCATTCGACTGCGTATCGTTGAACTAGCGCAACTCGGCATAAGACCCTGCGACATCAGCAGACAATTGCGTGTTTCTCACGGC  480
61    G  G  I  F  V  N  G  R  P  L  P  N  T  I  R  L  R  I  V  E  L  A  Q  L  G  I  R  P  C  D  I  S  R  Q  L  R  V  S  H  G  100

481   TGTGTTAGTAAAATACTTGCTCGTTATAACGAAACCGGCTCCATTCTTCCTGGAGCCATTGGAGGAAGCAAACCTAGGGTGACAACGCCGTCAGTAGTCAATGCAATCAAGGATTACAAG  600
101   C  V  S  K  I  L  A  R  Y  N  E  T  G  S  I  L  P  G  A  I  G  G  S  K  P  R  V  T  T  P  S  V  V  N  A  I  K  D  Y  K  140

601   ATAAGGGACCCTGGAATATTTGCTTGGGAGATCCGTGACCGCTTATTGTCTGACTGTATTTGTGATAAGTACAATGTACCTTCCGTCAGTTCAATTAGTCGCATTCTTCGCAACAAAATA  720
141   I  R  D  P  G  I  F  A  W  E  I  R  D  R  L  L  S  D  C  I  C  D  K  Y  N  V  P  S  V  S  S  I  S  R  I  L  R  N  K  I  180

721   GGGAATGTGCTTCACCCACACAACCCTCTTAATCCAAACTACGTCCCGGCGACTCAAGAGTGCATGAAGGAGATGAGCTCGGGGTCGGGTGCTCATCCCCATCCTTACTGTACATCGCCA  840
181   G  N  V  L  H  P  H  N  P  L  N  P  N  Y  V  P  A  T  Q  E  C  M  K  E  M  S  S  G  S  G  A  H  P  H  P  Y  C  T  S  P  220

841   GCAGCCATGTACACGCCATACATGATGCAACCAGCTGGATCCAAGCAGCCACCAATGCCTACGGTACCTACAGCACCAGGATGTTTTCGTGCGCCTGCCATACCCTATCCATTGTATTCT  960
221   A  A  M  Y  T  P  Y  M  M  Q  P  A  G  S  K  Q  P  P  M  P  T  V  P  T  A  P  G  C  F  R  A  P  A  I  P  Y  P  L  Y  S  260

961   CATTCAATGAGCGAAATTCTTGGCATTCCAAGATTCGACCAACAGATCCCAGTTACAATTTGCGGCGCGACGATTCCCGGTACCAACAATCATTATCCACCCCACGGCACTATGGTGATT  1080
261   H  S  M  S  E  I  L  G  I  P  R  F  D  Q  Q  I  P  V  T  I  C  G  A  T  I  P  G  T  N  N  H  Y  P  P  H  G  T  M  V  I  300

1081  CCGAACGAAAACATGCCTTCCGCTATGTCTACTAAGATGGAGGACTGGTCCAGAGTTCCCGCCTATGGGGTAACACCGCCAGCCCAAGTTACTTCCGTAAATGTGAATGGAAGTTCGCTA  1200
301   P  N  E  N  M  P  S  A  M  S  T  K  M  E  D  W  S  R  V  P  A  Y  G  V  T  P  P  A  Q  V  T  S  V  N  V  N  G  S  S  L  340

1201  AGCGAGATCGATAAGAATTCAAACGTGACCCATGAGAACATTCCACCTAATACTCACAGTGTGCCACCACAAACACATTCGACGAATTATGCATCGATTTCTGAAAACAATTATCTGACT  1320
341   S  E  I  D  K  N  S  N  V  T  H  E  N  I  P  P  N  T  H  S  V  P  P  Q  T  H  S  T  N  Y  A  S  I  S  E  N  N  Y  L  T  380

1321  ACATCTCCAGTAAGTGTACCGCCTCCAGTATCCGGATATCAAAACAGTCATCCTTACATGTCGCACCAGTATGGAGCTTCCCAAGGCATTAACGGTTGGCATTCGGTGCACTCAGGACAC  1440
381   T  S  P  V  S  V  P  P  P  V  S  G  Y  Q  N  S  H  P  Y  M  S  H  Q  Y  G  A  S  Q  G  I  N  G  W  H  S  V  H  S  G  H  420

1441  GGACCCCAGCGGGCAATATCGCCCCATAACATGATGTCCCAAATGACACAAGGCCCAACGGGCACCATCATACAAATCTCCGACCACAACAAGCAGTTAACCATCAAAGTCCGAATGGCT  1560
421   G  P  Q  R  A  I  S  P  H  N  M  M  S  Q  M  T  Q  G  P  T  G  T  I  I  Q  I  S  D  H  N  K  Q  L  T  I  K  V  R  M  A  460

1561  CTGGCCCTGTGGCTCCCATCAATGTTTATTCCAAGCAACAGCAACAGCATCAGTCGGCGACACCGGTACGCTTAAAATTTGATAACAACGCACCCAATAAAAAACCATCCAATCATTCG  1680
461   L  A  L  W  L  P  S  M  F  I  P  S  N  S  N  S  I  S  R  R  H  P  V  R  L  K  F  D  N  N  A  P  N  K  K  P  S  N  H  S  500

1681  CATACAACGGGTTCTGCGATAGGAACGGCAACAGCATGAAAATCATTCCGTCAGACATCTTTGGAAAGGAAAGAAAGAGCTGTATGCAAATTAGATCAAGGCCTTTAAGGTATGGTCAAG  1800
501   H  T  T  G  S  A  I  G  T  A  T  A  *                                                                                    512

1801  CGAAGGAAGGAGACATATTTTGCTGCCTTATCTATATCATATACATATGCAAAATATGTATTGCTGTATTTTAAACTTTACACAGAGTTATCCATTTAATACATTTACATTATAATTTCC  1920

1921  TTACTTTTGTAGGTTATCATATACTGCCTTATCCTCTTACCATAGTAAATAATTGTCAAATTACTGAAATGCTTGCAGCGTTCACCTATATGCATATATAATAGGCGTAAAGTTTGTTAC  2040

2041  TGTATATCATAAAACTAAAAGGACAGAAGCCATGGAAATTCTTTCTGGATAAAATTGGCAACAATGATTCCTAATGTTGATAATAATCGCTGAAAAGCGAGCCACAAGATACCTCTAGGT  2160

2161  GTTCTAGGGCTTACTCTTTATCGATATATATTTATAAAGTTTGCTTATAGCTACATTGCATTATATTTTCTAGAATTCCAATTTAACTTTTACATTTTCACGTAACGAATGGAAACAGCA  2280

2281  ATAATTCTACGTATCAGTTGTTGGAAACTCCGTAGTCTGAAATACCAAGTGAATGGGCTGGGACTTAACGCTGACTAAGATAGGATCCGATCCGATCCGCTGATAACATAAGGTCTTGGG  2400

2401  TTGCAACCTTAGAAATAAAAAAAGGTTGTGCTTGCACTGCAGCTGAAGGAGTGCATCTTTTATAGTAGAAACTAATTATTTTATTTCAAACGTTTTCTTCAATAAATAAAGGCTTCATAA  2520

2521  AAACGAAAATTTATTGTTATGACAAGTAGAACAAAATACTTGGGATATCATTTTAAAAAGCAATTAAAAAGCGTTTCCCATAATTTTGAACTTTCAACCCAGTCTTGTACATTTTTTTCT  2640

2641  TCGATTAACGAATATTGATGTTGCTATTATTCAACGATTCTAGATTTCAGGGAAGTAAGGTGCCGAATGCTGACCTAAAACATCTGCGTTATGTAGAACCTTTTGTTATTATATTACGTA  2760

2761  TTTTCGGCTGCTGCTCTTCCTGTTAAGAGAATCGAATCATATTACTTATGATGGGTAAATGTTCCGAATATTGAGATGAATAGAATCAGCCACAATATTTTGGTTACAATACTGTTGAAT  2880

2881  GTTTACGTGATTATCTATTTTAACTTCTACCGAATTACAGTATATGAGTGCCATACTTTGCCTTCTCACTTCGCTCATAATTTTTGTTTGAAACGTCGACTAGTTAATGCTAGCAATTTT  3000

3001  GTTCCCTATTTCGACCGTTCATCTCATGTCTGAAATTTCGGTCAACGTGCAAAGATTTTTATTTGATTATATACTGTATCCGTATACTTAATTTTCAATAAACAATTTTGAAAAGTAAA  3120

3121  AAAAAAAAAAAAA  3133
```

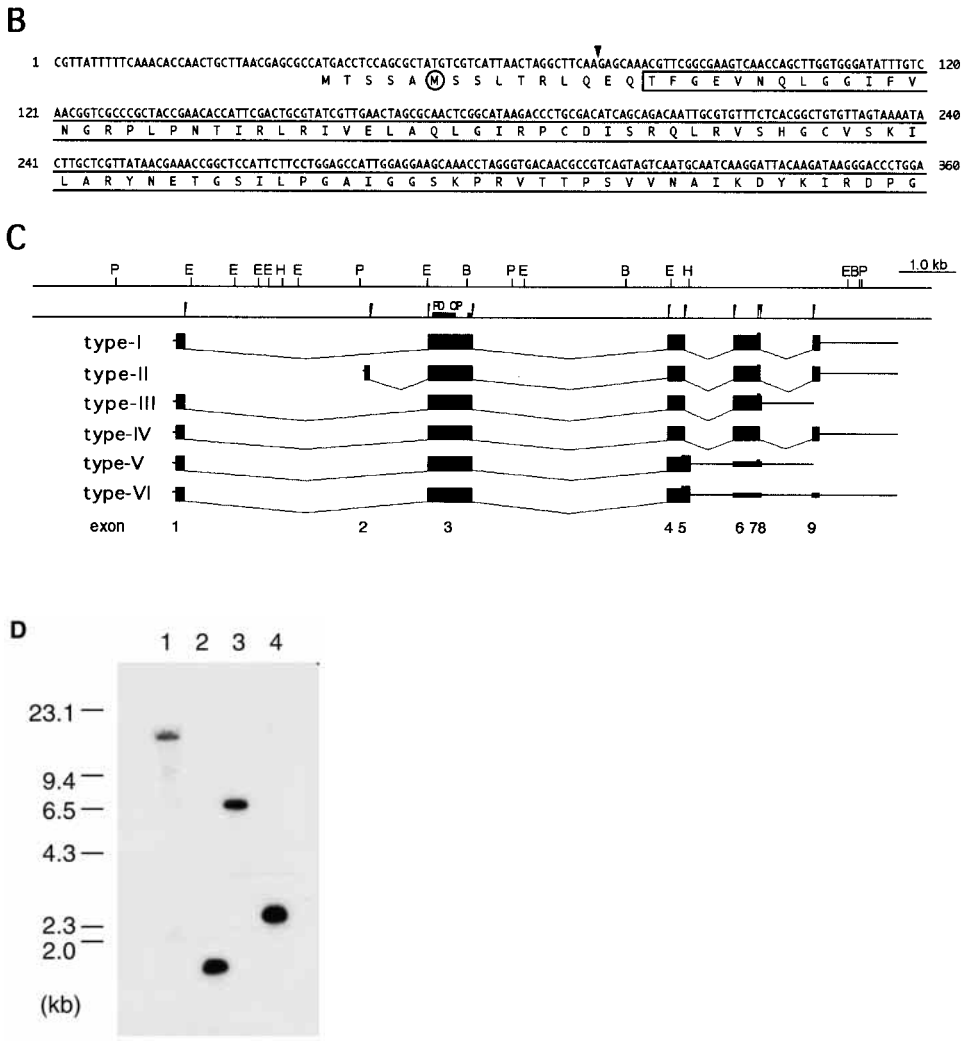
Fig. 1. *HrPax1/9* of *H. roretzi*. (A) Nucleotide and predicted amino acid sequences of cDNA clone for *HrPax1/9* type-I cDNA clone. The single ORF encodes a polypeptide of 512 amino acids, of which the molecular mass was estimated to be 55.5 kDa. The paired domain is boxed and the octapeptide is double underlined. The methionine near the paired domain is circled. The asterisk indicates the termination codon. The potential signal sequence for polyadenylation is underlined. Arrowheads indicate predicted splicing sites. The sequence will appear under the DDBJ/EMBL/GenBank accession number AB021138. (B) The 5′ region of *HrPax1/9* type-II cDNA. The paired domain is boxed. The arrowhead indicates a predicted alternative splicing site. The methionine near the paired

transcript, 4.0 kb, coincided with that of type-VI. No transcript of the size of type-III, about 2.3 kb, was detected under our hybridization conditions. Both the 3.2-kb and 4.0-kb transcripts were restricted to the pharyngeal gill (Fig. 3A, lane 1), and were not detected in the endostyle, muscle, gonad, digestive gland or intestine. Pharyngeal gill-specific expression of *HrPax1/9* was confirmed by means of RT-PCR/Southern blot analysis. As evident from Fig. 3B, expression was detected only in the pharyngeal gill.

The temporal expression pattern of *HrPax1/9* was determined by northern blot analysis using poly(A)$^+$ RNA prepared from specimens at different stages. No hybridization signal was detected in fertilized eggs or in embryos at the 16-cell, 64-cell, gastrula or tailbud stage (data not shown). The hybridization signals of 3.2 kb and 4.0 kb were first detected in swimming larvae (Fig. 3C). The expression of the *HrPax1/9* is maximal in 3-day-old juveniles, decreases in 5-day-old and 7-day-old juveniles, and is then maintained in the adult pharyngeal gill (Fig. 3C). The band intensity of the 3.2 kb and 4.0 kb transcripts did not change during the developmental stages from larva to adult.

To characterize the spatiotemporal expression in detail, in situ hybridization of whole-mount and sectioned specimens was carried out with a DIG-labeled *HrPax1/9* probe. The pharyngeal gills dissected from adults were hybridized as whole-mount specimens with *HrPax1/9* sense and antisense probes. No signal above background level was found in the control specimen hybridized with sense probe (Fig. 3D), whereas the pharyngeal gill hybridized with antisense probe showed an intense signal. The signal was evident in the inner part of the pharyngeal basket, which was derived from the endoderm (Fig. 3E) but not in the outer part of the pharyngeal gill, which originated from the ectoderm. The signal was evident along the wall of the longitudinal vessel and transverse vessel but not in the basal region of the pharyngeal gill.

Fig. 3F,G shows transverse sections of juveniles. The hybridization signal was evident in the inside part of the pharyngeal gill of 3-day-old juveniles (Fig. 3F) and 5-day-old juveniles (Fig. 3G). Fertilized eggs, and embryos at the 8-cell, 64-cell, gastrula, neurula and mid tailbud stages showed no in situ hybridization signals. A weak signal was first detected in the endodermal cells of the head region of swimming larvae, which eventually form the pharyngeal gill (data not shown).



domain is circled. The sequence will appear under the DDBJ/EMBL/GenBank accession number AB021139. (C) Predicted genomic construction and alternative splicing types of *HrPax1/9*. Boxes are ORFs and solid lines are non-coding regions. Locations of the paired domain (PD) and octapeptide (OP) are indicated (exon 3). Arrowheads indicate the predicted splicing sites. There is no splicing site within the region of the paired domain and octapeptide. Restriction sites are indicated for *Bam*HI (B), *Eco*RI (E), *Hin*dII (H), and *Pst*I (P). Type-III to -VI sequences will appear under the DDBJ/EMBL/GenBank accession numbers AB021140 to AB021143, respectively. (D) Genomic Southern blot analysis of *HrPax1/9*. Genomic DNA was isolated from a single adult, and aliquots were digested separately with *Bam*HI (lane 1), *Eco*RI (lane 2), *Hin*dIII (lane 3) and *Pst*I (lane 4). The blots were hybridized with the ^{32}P-labeled DNA probes which contain the paired domain. The filter was washed under high-stringency conditions. 10 μg of digested genomic DNA were loaded per lane. The numbers indicate sizes (in kb) of λ-*Hin*dIII-digested marker.

Isolation and characterization of cDNA for *Ciona intestinalis Pax1/9* gene

We amplified and subcloned a *Pax1/9* fragment that contained a paired box of *C. intestinalis* using an isolation strategy similar to that adopted for *H. roretzi*. RT-PCR using adult tissues also suggested that the *Ciona Pax1/9* gene was

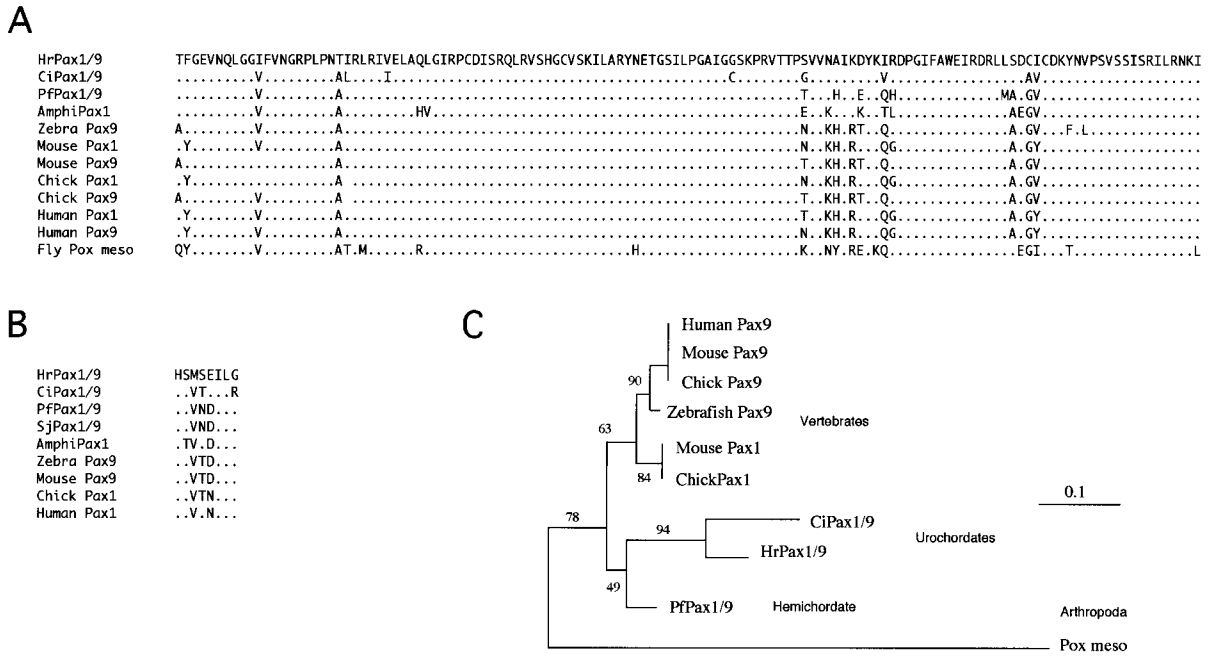


Fig. 2. Relationships of *Pax1/9*-related gene products. Comparison of (A) the amino acid sequence of paired domain, and (B) octapeptide of ascidian HrPax1/9, ascidian CiPax1/9 and acorn worm PfPax1/9 with human Pax1 (Burri et al., 1989), human Pax9 (Stapleton et al., 1993), mouse Pax1 (Deutsch et al., 1988), mouse Pax9 (Neubüser et al., 1995), chick Pax1 (Peters et al., 1995), chick Pax9 (Peters et al., 1995), fish Pax9 (Nornes et al., 1996), amphioxus pax1 (Holland and Holland, 1995), and *Drosophila* Pox meso (Bopp et al., 1989). Identical amino acids are indicated by dots. (C) A neighbor-joining tree based on comparison of the amino acid sequences of the paired domains. The *Drosophila* Pox meso was used as an outgroup. Branch length is proportional to the number of amino acid substitutions; the scale bar indicates 0.1 amino acid substitution in the sequence. The numbers indicate the relative robustness of each node as assessed by boot strap analysis (100 replications). The tree does not include AmphiPax1, because its incorporation disturbed the tree topology severely.

predominantly expressed in the pharyngeal gill. We made a cDNA library of *C. intestinalis* pharyngeal gill. Screening 3.0×10^5 pfu of a *C. intestinalis* adult pharyngeal gill cDNA library yielded 12 cDNA clones that covered almost the full length of this gene transcript. The structure of all these clones was the same, and therefore the corresponding gene was named *CiPax1/9* (*Ciona intestinalis Pax1*/*Pax9*).

The nucleotide and predicted amino acid sequences of the *CiPax1/9* cDNA are shown in Fig. 4. This clone consisted of 1,698 nucleotides, excluding the poly(A) tail. It contained a single open reading frame, predicting a polypeptide with a paired type DNA-binding domain and octapeptide. Like other *Pax1*/*Pax9* genes, this clone lacked a homeo domain. The calculated molecular mass of the *CiPax1/9*-encoded protein was 46.2 kDa. The cDNA sequence and primary structure of the amino acid sequence suggested strongly that *CiPax1/9* belongs to the *Pax1*/*Pax9* subgroup (Fig. 2C).

CiPax1/9 is predominantly expressed in the pharyngeal gill

A hybridization signal of about 1.8 kb was detected in the pharyngeal gill (Fig. 5A). Although two types of *HrPax1/9* hybridization signal were detected in *H. roretzi*, only one *CiPax1/9* signal was detected in *C. intestinalis* (Fig. 5A). In contrast to *H. roretzi*, in which *HrPax1/9* was expressed exclusively in the pharyngeal gill, a weak hybridization signal of *CiPax1/9* was detected in the endostyle, intestine and body wall muscle (Fig. 5A). *CiPax1/9* expression was therefore examined by means of RT-PCR/Southern blot analysis, as in the case of *H. roretzi*. As shown in Fig. 5B, a signal was detected not only in the pharyngeal gill but also in the endostyle, intestine and body wall muscle, although the gonad did not show a signal even by RT-PCR/Southern blot analysis. Therefore, *CiPax1/9* is zygotically expressed not only in the pharyngeal gill but also in other tissues.

The spatiotemporal expression was determined by in situ hybridization of whole-mount and sectioned specimens. As is evident in Fig. 5C,D, an intense signal was detected in the inner part of the pharyngeal basket, which is formed by endoderm. Whole-mount in situ hybridization of a 1-month-old adult showed a signal in the pharyngeal gill and in the endostyle (Fig. 5E). Fig. 5F shows a transverse section of an adult endostyle and pharyngeal gill. Hybridization signals were evident in the inner side of the pharyngeal gill (green arrowheads) and a part of the endostyle (red arrowheads).

Isolation and characterization of cDNA for a hemichordate *Ptychodera flava Pax1/9* gene

A DNA fragment of *P. flava Pax1/9* was obtained by RT-PCR from adult poly(A)$^+$ RNA. We made a cDNA library of *P. flava* adult gill, and screened 3.0×10^5 pfu using the random-labeled fragment as a probe. We found that 9 cDNA clones covered almost the full length of this gene transcript. All of the clones were of the same structure, and the gene was named *PfPax1/9* (*Ptychodera flava Pax1*/*Pax9*).

The nucleotide and predicted amino acid sequences of the *PfPax1/9* cDNA clone are shown in Fig. 6. This clone consisted of 2,892 nucleotides, excluding the poly(A) tail. The clone had

a single open reading frame that encoded a polypeptide with the paired type DNA binding domain and octapeptide (Fig. 2), and lacked a homeo domain. The calculated molecular mass of the *PfPax1/9*-encoded protein was 41.2 kDa. As shown in Fig. 2C, *PfPax1/9* belongs to the *Pax1/Pax9* subgroup.

PfPax1/9 is expressed in the gill

As described above, the ascidian *Pax1/9* genes are either exclusively or predominantly expressed in the pharyngeal gill. The hemichordate *PfPax1/9* expression was first examined by northern blot and RT-PCR/Southern analysis. Poly(A)$^+$ RNA prepared from the proboscis, collar, gill, gonad, hepatic region and caudal region were subjected to hybridization (Fig. 7A). As shown in Fig. 7B, only one hybridization signal, of about 3.0 kb, was detected in the gill (lane 3), and no signal was detected in other regions. RT-PCR/Southern blot analysis confirmed that signal was detected only in the gill (Fig. 7C, lane 3).

To characterize the spatial expression in detail, in situ hybridizations of whole-mount and sectioned specimens were carried out with DIG-labeled *PfPax1/9* probe. A diagram of the structure of the gill is shown in Fig. 7D. The acorn worm gill is primarily composed of branchial septa and tongue bars, which run parallel, perpendicular to the midline of the body. Under our in situ hybridization conditions, there were many spot-like backgrounds with sense probe (Fig. 7E,F), and they were found in certain parts of the branchial septa and tongue bars (Fig. 7F). However, specimens hybridized with antisense probe demonstrated distinct real signal. As is shown in Fig. 7G,H and I, strong signal was evident in the inside region of branchial septa and tongue bars.

DISCUSSION

Urochordate and hemichordate *Pax1/9* genes

Pax1/*Pax9* genes encode polypeptides with a highly conserved paired domain and octapeptide, and lack the homeodomain which is present in all other *Pax* gene subfamilies (reviewed by Strachan and Read, 1994; Dahl et al., 1997). The cDNAs we isolated from two ascidian species *H. roretzi* and *C. intestinalis*, and an acorn worm, *P. flava*, encoded polypeptides with paired domains and octapeptides shared by Pax1/Pax9 family members. Together with the result of molecular phylogenetic analysis, it is evident that *HrPax1/9*, *CiPax1/9* and *PfPax1/9* are *Pax1/Pax9*-related genes.

As for *HrPax1/9*, northern blot analysis detected two transcripts, of about 3.2 kb and 4.0 kb (Fig. 3A,C). In addition, we isolated six types of cDNAs which encoded polypeptides with different N-terminal and C-terminal ends (Fig. 1C). These different types of cDNAs represented transcripts produced by alternative splicing of a single gene (Fig. 1C). In zebrafish, two *Pax9* transcripts, *Pax9a* and *Pax9b*, are generated by alternative splicing (Nornes et al., 1996). The zebrafish Pax9a and Pax9b have divergent C-terminal ends with different *trans*-activating properties (Nornes et al., 1996). These data suggest that the HrPax1/9s with different C-terminal ends have different functions which should be examined in further analyses.

Temporal expression of urochordates and hemichordate *Pax1/9* genes

The *HrPax1/9* transcripts are not detected in early embryogenesis, and are first detected in swimming larvae immediately before pharyngeal gill formation. The amount of transcripts increased during metamorphosis, and was maintained in adults. No maternal transcripts are detected during early embryogenesis. The gene therefore exerts its function in later stages of development and in adults. Similarly, *CiPax1/9* and *PfPax1/9* are expressed only in adults. This

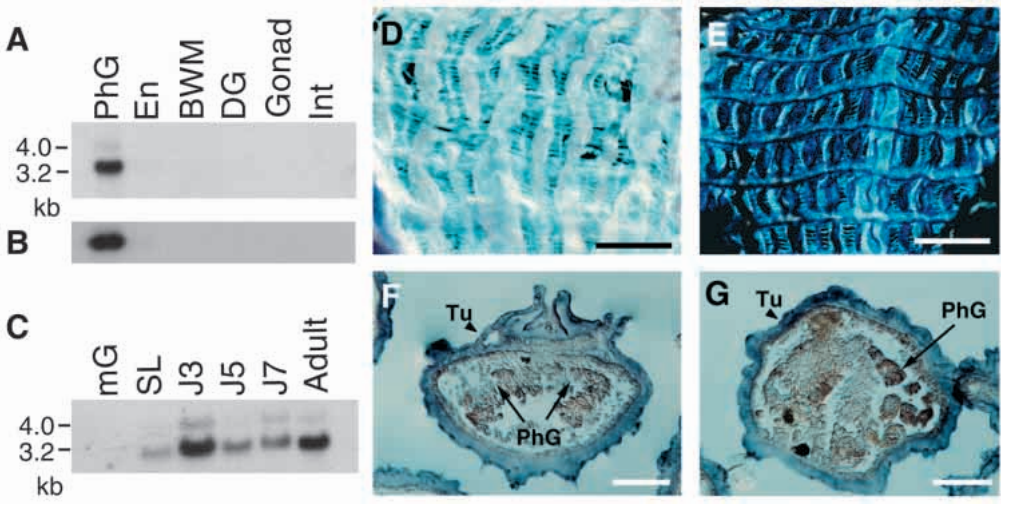


Fig. 3. Distribution of *HrPax1/9* transcript. (A) Northern blots of poly(A)$^+$ RNA prepared from the pharyngeal gill (lane 1), endostyle (lane 2), body-wall muscle (lane 3), digestive gland (lane 4), gonad (lane 5), and intestine (lane 6) of *H. roretzi* were hybridized with the random-primed ^{32}P-labeled DNA probes and the membrane was washed under high-stringency conditions. The *HrPax1/9* transcripts of about 3.2 kb and 4.0 kb in length were detected only in the pharyngeal gill. Each lane contained 10 μg of poly(A)$^+$ RNA. (B) RT-PCR/Southern analysis of *HrPax1/9* transcript in adult tissues and organs. RT-PCR products prepared from the same tissues as A were hybridized and washed under the same conditions. (C) Occurrence of *HrPax1/9* transcripts during *H. roretzi* development. Northern blots of poly(A)$^+$ RNA prepared from gastrulae (lane 1), swimming larvae (lane 2), juveniles at 3 (lane 3), 5 (lane 4), and 7 days (lane 5) after initiation of metamorphosis, and adult endostyle (lane 6). The *HrPax1/9* transcript was not detected in embryos, but was detected in larvae and increased in metamorphosed juveniles. Each lane contained 10 μg of poly(A)$^+$ RNA. (D-G) Localization of *HrPax1/9* transcripts as revealed by in situ hybridization. (D,E) The pharyngeal gills dissected from adults were hybridized as whole-mount specimens with *HrPax1/9* sense (D) and antisense (E) probes. Distinct signals are evident in the endodermal region of the pharyngeal gill when hybridized with antisense probe. Scale bars, 1 mm. (F,G) Cross sections of 3-day-old juvenile (F) and 5-day-old juvenile (G) showing signals in differentiating pharyngeal gill (PhG; arrows). The tunic (Tu) showed high background staining. Scale bars, 100 μm.

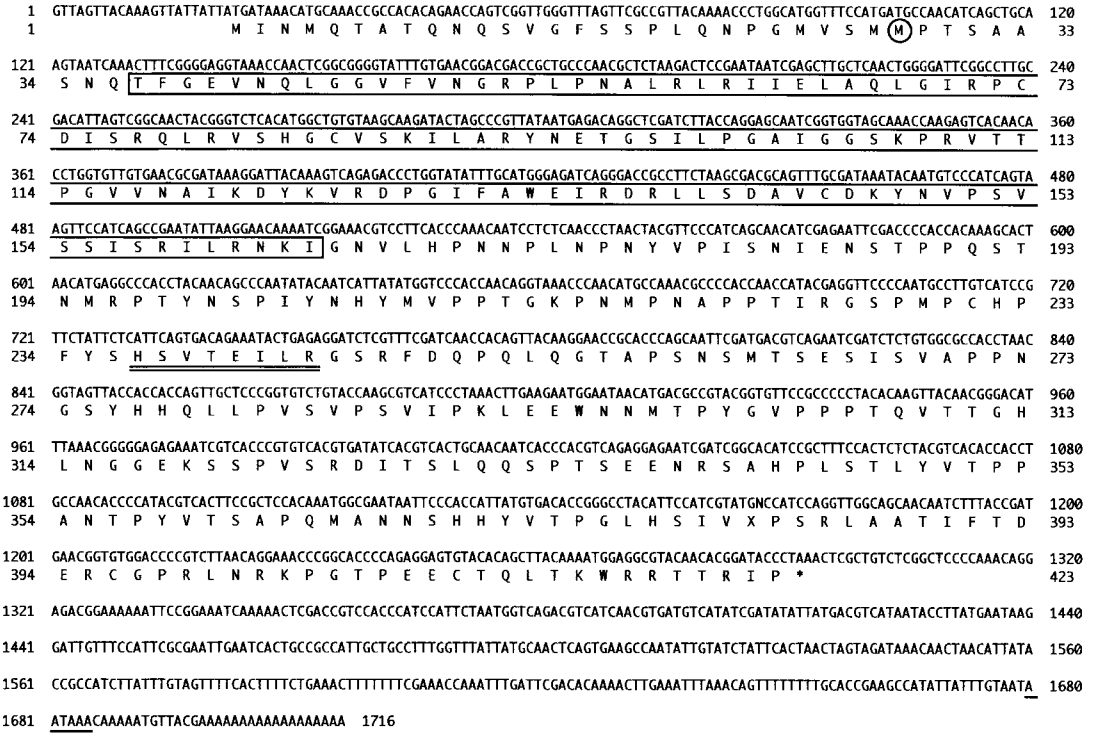


Fig. 4. Nucleotide and predicted amino acid sequences of cDNA clone for *CiPax1/9* of *C. intestinalis*. The single ORF encodes a polypeptide of 423 amino acids, of which the molecular mass was estimated to be 46.2 kDa. The paired domain is boxed and the octapeptide is double underlined. The methionine near the paired domain is circled. An asterisk indicates the termination codon. The potential signal sequence for polyadenylation is underlined. The sequence will appear under the DDBJ/EMBL/GenBank accession number AB020762.

temporal expression pattern is similar to that of mouse *Pax1* and *Pax9* genes; they are not expressed in early embryogenesis, but are expressed during the later stages of development from 8.5 p.c., and maintained in the adult tissues (Deutsch et al., 1988; Neubüser et al., 1995). The amphioxus *AmphiPax1* expression is first detected in the endoderm that forms the pharynx soon after hatching (Holland and Holland, 1995). Furthermore, the sea urchin *Pax1/9* is also not expressed during early embryonic development from the egg to the pluteus stage, but is expressed in adult tissues (Czerny et al., 1997). All of these data indicate that *Pax1/Pax9* subfamily genes are involved in organ formation. In addition, this temporal expression mechanism is highly conserved among deuterostomes.

Spatial expression of *Pax1/Pax9* subfamily genes and evolution of possible functions

Fig. 8 summarizes the phylogenetic relationship of animals and expression patterns of *Pax1/9*-related genes which have been reported to date or characterized in this study. The tree shown

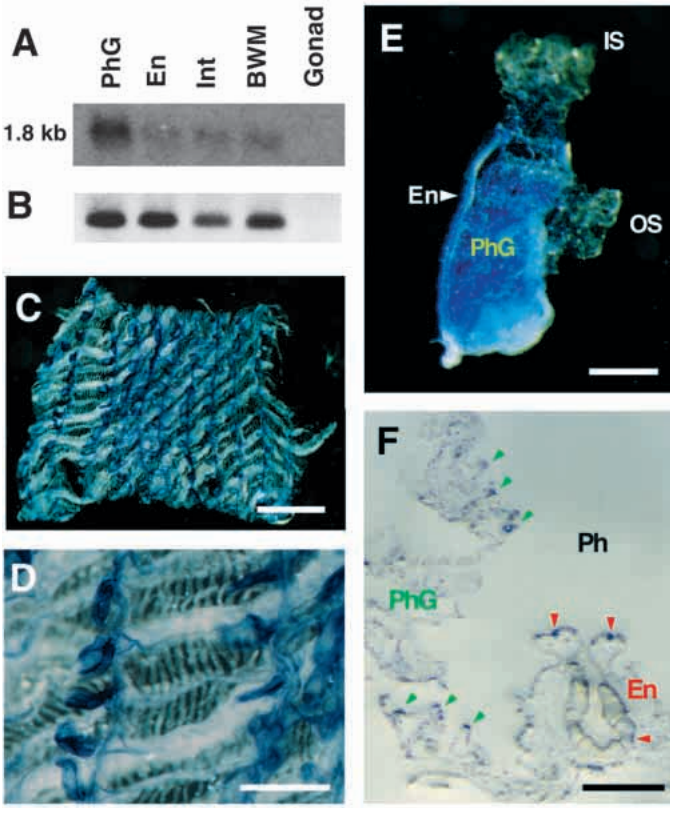


Fig. 5. Distribution of *CiPax1/9* transcript. (A) Northern blot of poly(A)$^+$ RNA prepared from the pharyngeal gill (lane 1), endostyle (lane 2), intestine (lane 3), body wall muscle (lane 4) and gonad (lane 5) of *C. intestinalis* were hybridized with random-primed ^{32}P-labeled DNA probes, and the membrane was washed under high-stringency conditions. The *CiPax1/9* transcripts of about 1.8 kb in length were detected mainly in the pharyngeal gill. Each lane contained 10 μg of poly(A)$^+$ RNA. (B) RT-PCR/Southern analysis of *CiPax1/9* transcripts showing RT-PCR products in the pharyngeal gill (lane 1), endostyle (lane 2), intestine (lane 3) and body-wall muscle (lane 4), but not gonad (lane 5). (C-F) Localization of *CiPax1/9* transcripts, as revealed by in situ hybridization. (C) The pharyngeal gill dissected from an adult was hybridized as a whole-mount specimen with *CiPax1/9* antisense probes. Scale bar, 1 mm. (D) High magnification of the pharyngeal gill showing signals in the endodermal region. Scale bar, 100 μm. (E) A whole-mount specimen of a 1-month-old young adult. Signal is detected in the pharyngeal gill (PhG) and endostyle (En). IS, incurrent siphon; OS, outcurrent siphon. Scale bar, 500 μm. (F) A cross section hybridized with an antisense probe showing signals in the pharynx (Ph) side of the pharyngeal gill (PhG; green arrowheads) and endostyle (En; red arrowheads). Scale bar, 100 μm.

in Fig. 8 is arbitrary. Although the rDNA tree suggests a close relationship between hemichordates and echinoderms (Wada and Satoh, 1994; Turbeville et al., 1994), we prefer to pose hemichordates between echinoderms and chordates, based on the present results as well as expression pattern of other developmental genes (Taguchi et al., personal communication).

Pax genes have been isolated from sponges (Hoshiyama et al., 1998) and cnidarians (Sun et al., 1997). Therefore, the occurrence of *Pax* genes is deep in the evolutionary history of multicellular animals, and thus *Pax1/9* or its ancestral gene first appeared before divergence of the deuterostome and protostome lineages (event 'a'). In Fig. 8, *Pax1/9* expression in the gill is indicated in green, and that in other (mesoderm) tissues is indicated in pink. In contrast to the fact that *Pax* genes of Group II (*Pax2*, *Pax5* and *Pax8*), Group III (*Pax3* and *Pax7*) and Group IV (*Pax4* and *Pax6*) are expressed in the CNS and associated with the development of the CNS, *Pax1/Pax9* genes of the Group I are never expressed in the CNS. *Pax1* and *Pax9* were first characterized in mouse (Deutsch et al., 1988; Koseki et al., 1993; Wallin et al., 1994; Neubüser et al., 1995); the genes are expressed segmentally in the somite, sclerotome and intervertebral disks along the body axis, and the expression is required for the axial skeletal development. In addition, the *Drosophila Pax1/Pax9*-related gene, *Pox meso*, is expressed segmentally in somatic mesoderm (Bopp et al., 1989). Therefore, it has been emphasized that *Pax1/Pax9*-related genes are primarily associated with functions in the mesoderm (pink in Fig. 8; Mansouri et al., 1996).

However, mouse *Pax1* and *Pax9* are also expressed overlappingly in the endodermally derived epithelium of the pharyngeal pouch that forms thymus, parathyroid glands, ultimobranchial bodies, eustachian tube and tonsils (Deutsch et al., 1988; Timmons et al., 1994; Neubüser et al., 1995, 1997). In chick, both *Pax1* and *Pax9* transcripts were first detected in the prospective foregut endoderm in late head-fold stage embryos, prior to the formation of the somites (Müller et al., 1996). During the foregut development, transcripts of chick *Pax1* and *Pax9* were found in the prospective pharyngeal pouch endoderm of the intermediate zone. In addition, the amphioxus *AmphiPax1* transcript is detected in the endodermal cells of pharyngeal gill, but segmental expression along the somite is never detected (Holland and Holland, 1995). The pharyngeal pouch is thought to be homologous to the pharyngeal gill. Furthermore, as was shown in the present study, the ascidian *Pax1/9* genes are expressed in the inner part of the pharyngeal gill which is derived from endoderm, and the hemichordate *Pax1/9* expression is specifically found in the endoderm-derived region of the gill (green in Fig. 8). Therefore, the present study emphasizes an important role of the Group I *Pax* genes during the development of endoderm, in particular, the gill.

Consideration of Fig. 8 suggests at least three scenarios concerning the ancestral role of the *Pax1/9* gene and/or conserved function of the *Pax1/9*-related genes. The first scenario of the *Pax1/9* function is that the ancestral and conserved function is associated with mesodermal

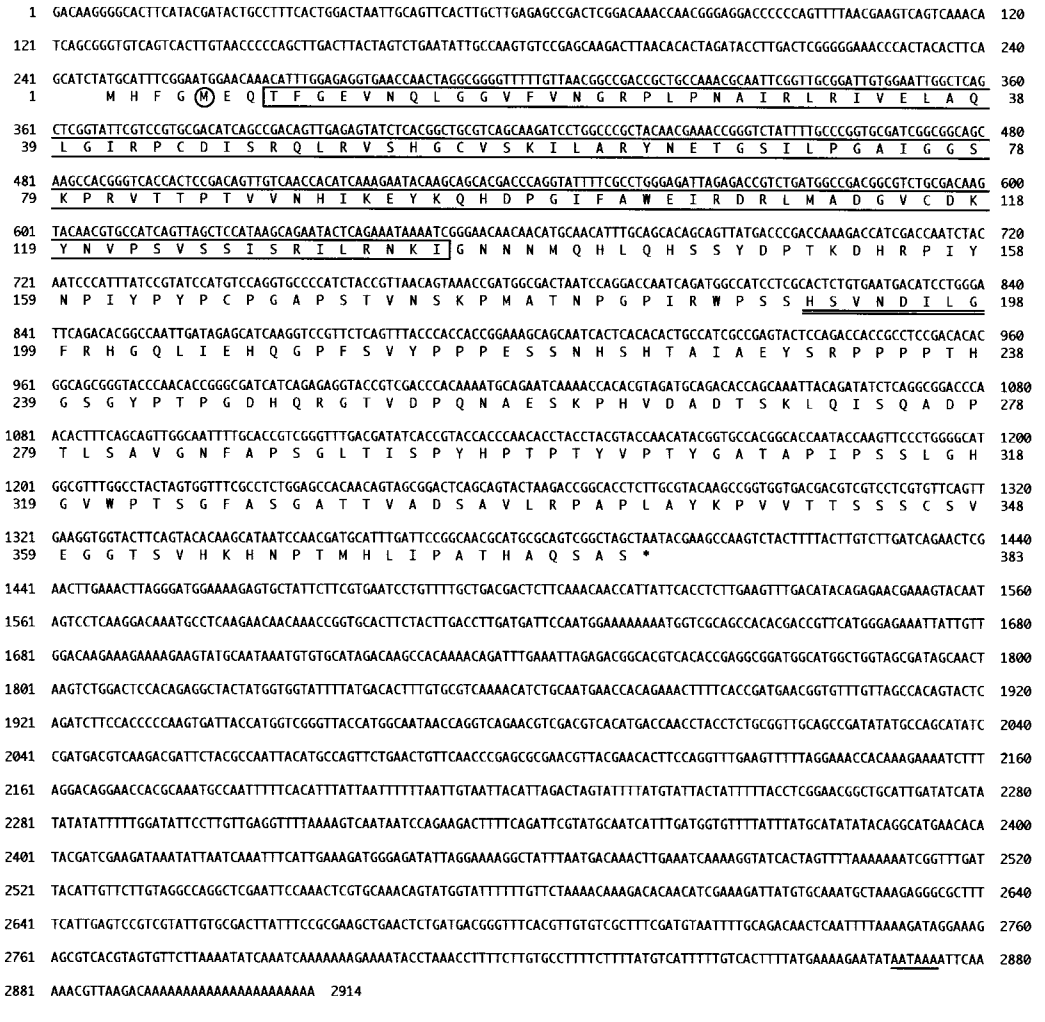


Fig. 6. Nucleotide and predicted amino acid sequences of cDNA clone for *PfPax1/9* of the acorn worm *Ptychodera flava*. The single ORF encodes a polypeptide of 383 amino acids, of which the molecular mass was estimated to be 41.2 kDa. The paired domain is boxed and the octapeptide is double underlined. The methionine near the paired domain is circled. An asterisk indicates the termination codon. The potential signal sequence for polyadenylation is underlined. The sequence will appear under the DDBJ/EMBL/GenBank accession number AB020763.

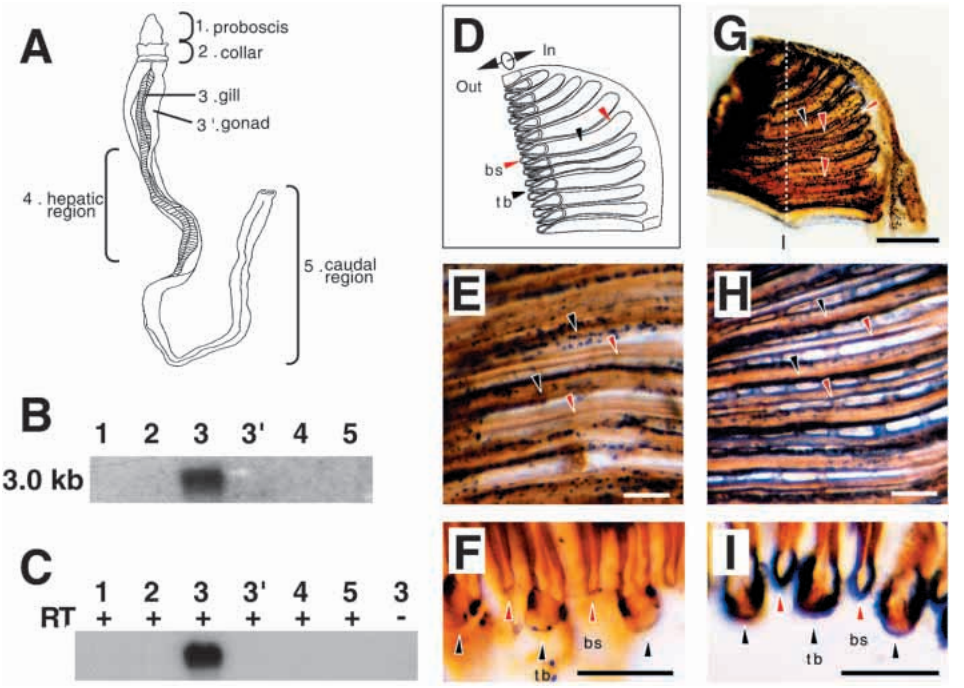


Fig. 7. Distribution of *PfPax1/9* transcripts in the gill of *P. flava*. (A) Diagram of the structure of *P. flava*. Numbers 1, 2, 3, 3′, 4 and 5 indicate the areas sampled in the various lanes in B and C. (B) A Northern blots of poly(A)$^+$ RNA prepared from the proboscis (lane 1), collar (lane 2), gill (lane 3), gonad (lane 3'), hepatic region (lane 4), and caudal region (lane 5) of *P. flava* was hybridized to *PfPax1/9* probe. The 3.0 kb *PfPax1/9* transcript was detected only in the gill. Each lane contained 10 μg of poly(A)$^+$ RNA. (C) RT-PCR/Southern analysis of *PfPax1/9* transcript showing RT-PCR products only in the gill (lane 3). (D) Diagram of the structure of gill, viewed from inside. The gill is composed of branchial septa (bs, read arrowheads) and tongue bars (tb, black arrowheads), which run parallel and perpendicular to the midline of the body. In, inside; Out, outside of the gill. (E-I) Distribution of *PfPax1/9* transcripts revealed by in situ hybridization. (E) The gill hybridized with sense probe, showing many spot-like background staining. Scale bar, 1 mm. (F) A section of (E). Scale bar, 500 μm. (G, H) The gill hybridized with antisense probe, showing true hybridization signal. Scale bar in G, 2 mm. (I) A section of the gill indicated in G. Signal is evident in the inside region of branchial septa (red arrowheads) and of tongue bars (black arrowheads). Scale bars, 500 μm.

development, and thus in cephalochordates, urochordates and hemichordates this function was lost and an additional function for the endodermal development arose. The second scenario is that *Pax1/9* function is associated with endodermal development, and therefore in vertebrates mesodermal function was added during the evolution of this animal group. The third scenario is that *Pax1/9* function is associated with both mesodermal and endodermal development. Because hemichordate, urochordate and cephalochordate *Pax1/9* genes are associated with gill formation, the function of *Pax1/9* genes appears to have been concentrated to this organ during the evolution of protochordates. Thus, the mechanisms for the specific expression might have arisen along with the evolution of the gill structure (event 'b'). During evolution of vertebrates,

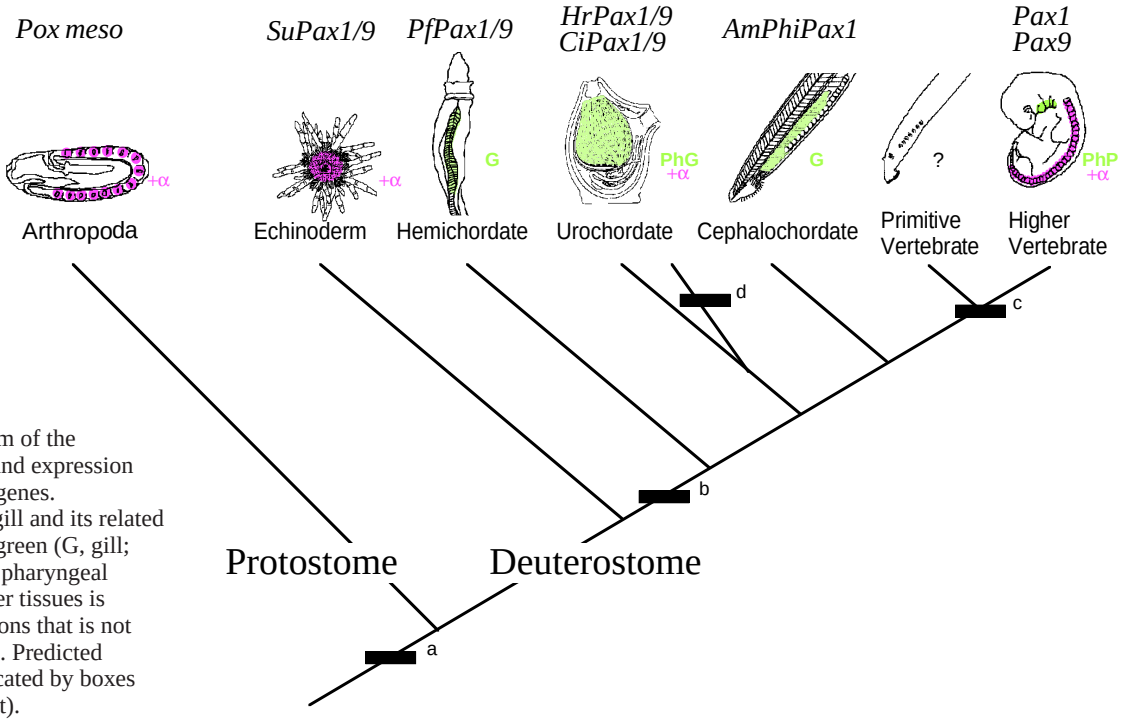


Fig. 8. A schematic diagram of the phylogenetic relationship and expression patterns of *Pax1/9*-related genes. Expression regions of the gill and its related structures are indicated in green (G, gill; PhG, pharyngeal gill; PhP, pharyngeal pouch), and that of the other tissues is indicated in pink. Expressions that is not reported is indicated by '?'. Predicted evolutional events are indicated by boxes 'a' to 'd' (see details in text).

Pax1/9 genes duplicated and evolved to *Pax1* and *Pax9*, and mesodermal expression might have been recovered (event 'c'). Additional expression was recovered in some urochordates (event 'd') during the evolution of the ascidians. Echinoderms have no gill-like structure. Sea urchin *Pax1/9* is expressed in the adult muscle (mesoderm) and intestine (endoderm) (Czerny et al., 1997). If deuterostomes originated from a common ancestor, this expression pattern of sea urchin *Pax1/9* seems unexpected. However, sea urchins show very derived homeobox gene expression patterns (Lowe and Wray, 1997; Arenas-Mena et al., 1998), suggesting that modern echinoderms have a derived bilateral body plan. Therefore, it is possible that the sea urchin *Pax1/9* expression is also derived but not primitive.

In conclusion, the present study investigated a relationship between *Pax1/9* genes and the gill formation during protochordate later embryogenesis. Our results provide evidence for that pharyngeal gill slits are homologous between hemichordates and chordates, and therefore strongly support William Bateson's proposal for a close link between enteropneusts and chordates.

We thank all of the staff members of the Otsuchi Marine Research Center, Ocean Research Institute, University of Tokyo and the Marine BioSource Education Center of Tohoku University, Onagawa, Miyagi for their hospitality. We also thank Drs Kuni Tagawa and Tom Humphreys for collection of acorn worms and Kazuko Hirayama for providing technical assistance. Thanks are also due to two anonymous referees for their suggestions to improve the manuscript.

M. O. is a predoctoral fellow of JSPS with a Monbusho research grant (No. 3252). This research was also supported by a Grant-in-Aid for Specially Promoted Research (No. 07102012) to N. S. from the Monbusho, Japan.

REFERENCES

Arenas-Mena, C., Martinez, P., Cameron, R. A. and Davidson, E. H. (1998). Expression of the *Hox* gene complex in the indirect development of a sea urchin. *Proc. Natl. Acad. Sci. USA* **95**, 13062-13067.

Bateson, W. (1886). The ancestry of the Chordata. *Q. J. Microsc. Sci*. **26**, 535-571.

Bopp, D., Jamet, E., Baumgartner, S., Burri, M. and Noll, M. (1989). Isolation of two tissue-specific *Drosophila* paired box genes, *Pox meso* and *Pox neuro*. *EMBO J*. **8**, 3447-3457.

Brusca, R. C. and Brusca, G. J. (1990). *Invertebrates*. Sunderland, MA: Sinauer.

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.

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

Czerny, T., Bouchard, M., Kozmik, Z. and Busslinger, M. (1997). The characterization of novel *Pax* genes of the sea urchin and *Drosophila* reveal an ancient evolutionary origin of the *Pax2/5/8* subfamily. *Mech. Dev*. **67**, 179-192.

Dahl, E., Koseki, H. and Balling, R. (1997). *Pax* genes and organogenesis. *BioEssays* **19**, 755-765.

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.

Di Gregorio, A. and Levine, M. (1998). Ascidian embryogenesis and the origins of the chordate body plan. *Curr. Opin. Genet. Dev*. **8**, 457-463.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791.

Gee, H. (1996). *Before the Backbone. Views on the origin of the vertebrates.* London: Chapman & Hall.

Gilbert, D. (1993). *SeqApp manual aligner for Macintosh ver. 1.9*, Bloomington: Indiana University.

Hall, B. K. (1998). *Evolutionary Developmental Biology*, Second ed. London: Chapman & Hall.

Harada, Y., Yasuo, H. and Satoh, N. (1995). A sea urchin homologue of the chordate *Brachyury* (*T*) gene is expressed in the secondary mesenchyme founder cells. *Development* **121**, 2747-2754.

Higgins, D. J., Bleasby, A. J. and Fuchs, R. (1992). Clustal V: improved software for multiple sequence alignment. *Comput. Appl. Biosci*. **8**, 189-191.

Holland, N. D. and Holland, L. Z. (1995). An amphioxus Pax gene, *AmphiPax-1*, expressed in embryonic endoderm, but not in mesoderm: implications for the evolution of class I paired box genes. *Mol. Mar. Biol. Biotech*. **4**, 206-214.

Hoshiyama, D., Suga, H., Iwabe, N., Koyanagi, M., Nikoh, N., Kuma, K., Matsuda, F., Honjo, T. and Miyata, T. (1998). Sponge *Pax* cDNA related to *Pax-2/5/8* and ancient gene duplications in the *Pax* family. *J. Mol. Evol*. **47**, 640-648.

Jefferies, R. P. S. (1986). *The Ancestry of the Vertebrates*. London: British Museum (Natural History).

Koseki, H., Wallin, J., Wilting, J., Mizutani, Y., Kispert, A., Ebensperger, C., Herrmann, B. G., Christ, B. and Balling, R. (1993). A role for *Pax-1* as a mediator of notochordal signals during the dorsoventral specification of vertebrae. *Development* **119**, 649-660.

Lowe, C. J. and Wray, G. A. (1997). Radical alterations in the roles of homeobox genes during echinoderm evolution. *Nature* **389**, 718-721.

Mansouri, A., Stoykova, A., Torres, M. and Gruss, P. (1996). Dysgenesis of cephalic neural crest derivatives in *Pax7*$^{-/-}$ mutant mice. *Development* **122**, 831-838.

Müller, T. S., Ebensperger, C., Neubüser, A., Koseki, H., Balling, R., Christ, B. and Wilting, J. (1996). Expression of avian *Pax1* and *Pax9* is intrinsically regulated in the pharyngeal endoderm, but depends on environmental influences in the paraxial mesoderm. *Dev. Biol*. **178**, 403-417.

Neubüser, A., Koseki, H. and Balling, R. (1995). Characterization and developmental expression of *Pax9*, a paired-box-containing gene related to *Pax1*. *Dev. Biol*. **170**, 701-716.

Neubüser, A., Peters, H., Balling, R. and Martin, G. R. (1997). Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation. *Cell* **90**, 247-255.

Nielsen, C. (1995). *Animal Evolution. Interrelationships of the Living Phyla*, Oxford: Oxford Univ. Press.

Nornes, S., Mikkola, I., Krauss, S., Delghandi, M., Perander, M. and Johansen, T. (1996). Zebrafish *Pax9* encodes two proteins with distinct C-terminal transactivating domains of different potency negatively regulated by adjacent N-terminal sequences. *J. Biol. Chem*. **271**, 26914-26923.

Ogasawara, M., Tanaka, K. J., Makabe, K. W. and Satoh, N. (1996). Expression of endostyle-specific genes in the ascidian *Halocynthia roretzi*. *Dev. Genes Evol*. **206**, 227-235.

Ogasawara, M. and Satoh, N. (1998). Isolation and characterization of endostyle-specific genes in the ascidian *Ciona intestinalis*. *Biol. Bull*. **195**, 60-69.

Peters, H., Doll, U. and Niessing, J. (1995). Differential expression of the chicken *Pax-1* and *Pax-9* gene: in situ hybridization and immunohistochemical analysis. *Dev. Dyn*. **203**, 1-16.

Peters, H., Neubüser, A., Kratochwil, K. and Balling, R. (1998). *Pax9*-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev*. **12**, 2735-2747.

Peterson, K. J. (1995). A phylogenetic test of the calcichordate scenario. *Lethaia* **28**, 25-38.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*. **4**, 406-425.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Satoh, N. (1995). Towards a molecular understanding of developmental mechanisms underlying the origin and evolution of chordates. In *Biodiversity and Evolution* (ed. R. Arai, M. Kato and Y. Doi), pp. 267-290. Tokyo: The National Science Museum Foundation.

Satoh, N. and Jeffery, W. R. (1995). Chasing tails in ascidians: developmental insights into the origin and evolution of chordates. *Trends Genet*. **11**, 354-359.

Satoh, N., Hotta, K., Satoh, G., Taguchi, S., Yasuo, H., Tagawa, K.,

Takahashi, H. and Harada, Y. (1999) Developmental mechanisms underlying the origin and evolution of chordates. In *The Biology of Diversity* (ed. M. Kato), Tokyo: Springer, in press.

Stapleton, P., Weith, A., Urbanek, P., Kozmik, Z. and Busslinger, M. (1993). Chromosomal localization of seven *Pax* genes and cloning of a novel family member, *Pax9*. *Nature Genet*. **3**, 292-304.

Strachan, T. and Read, A. P. (1994). PAX genes. *Curr. Opin. Genet. Dev*. **4**, 427-438.

Sun, H., Rodin, A., Zhou, Y., Dickinson, D. P., Harper, D. E., Hewett-Emmett, D. and Li, W. H. (1997). Evolution of paired domains: isolation and sequencing of jellyfish and hydra *Pax* genes related to *Pax-5* and *Pax-6*. *Proc. Natl. Acad. Sci. USA* **94**, 5156-5161.

Tagawa, K., Humphreys, T. and Satoh, N. (1998). Novel pattern of *Brachyury* gene expression in hemichordate embryos. *Mech. Dev*. **75**, 139-143.

Tanaka, K. J., Ogasawara, M., Makabe, K. W. and Satoh, N. (1996). Expression of pharyngeal gill-specific genes in the ascidian *Halocynthia roretzi*. *Develop. Genes Evol*. **206**, 218-226.

Timmons, P. M., Wallin, J., Rigby, P. W. and Balling, R. (1994). Expression and function of *Pax 1* during development of the pectoral girdle. *Development* **120**, 2773-2785.

Turbeville, J. M., Schulz, J. R. and Raff, R. A. (1994). Deuterostome phylogeny and the sister group of the chordates: Evidence from molecules and morphology. *Mol. Biol. Evol*. **11**, 648-655.

Wada, H. and Satoh, N. (1994). Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA* **91**, 1801-1804.

Wada, H., Saiga, H., Satoh, N. and Holland, P. W. H. (1998). Tripartite organization of the ancestral chordate brain and the antiquity of placodes: insights from ascidian *Pax-2*/*5*/*8*, *Hox* and *Otx* genes. *Development* **125**, 1113-1122.

Wallin, J., Wilting, J., Koseki, H., Fritsch, R., Christ, B. and Balling, R. (1994). The role of *Pax-1* in axial skeleton development. *Development* **120**, 1109-1121.

Wallin, J., Eibel, H., Neubüser, A., Wilting, J., Koseki, H. and Balling, R. (1996). *Pax1* is expressed during development of the thymus epithelium and is required for normal T-cell maturation. *Development* **122**, 23-30.

Yasuo, H. and Satoh, N. (1993). Function of vertebrate *T* gene. *Nature* **364**, 582-583.

Yasuo, H., Harada, Y. and Satoh, N. (1995). The role of *T* genes in the organization of the notochord during chordate evolution. *Seminars Dev. Biol*. **6**, 417-425.