

# The *Trichoplax PaxB* Gene: A Putative Proto-*PaxA/B/C* Gene Predating the Origin of Nerve and Sensory Cells

Thorsten Hadrys,\*†‡<sup>1</sup> Rob DeSalle,† Sven Sagasser,\* Nina Fischer,\* and Bernd Schierwater\*§

\*ITZ, Ecology & Evolution, Hannover, Germany; †Department of Invertebrate Zoology & Molecular Labs, American Museum of Natural History; ‡Department of Physiology and Neuroscience, NYU Medical Center; and §Yale University, MCDB

*Pax* genes play key regulatory roles in embryonic and sensory organ development in metazoans but their evolution and ancestral functions remain widely unresolved. We have isolated a *Pax* gene from Placozoa, beside Porifera the only metazoan phylum that completely lacks nerve and sensory cells or organs. These simplest known metazoans also lack any kind of symmetry, organs, extracellular matrix, basal lamina, muscle cells, and main body axis. The isolated *Pax* gene from *Trichoplax adhaerens* harbors a paired domain, an octapeptide, and a full-length homeodomain. It displays structural features not only of *PaxB* and *Pax2/5/8*-like genes but also of *PaxC* and *Pax6* genes. Conserved splice sites between Placozoa, Cnidaria, and triploblasts, mark the ancient origin of intron structures. Phylogenetic analyses demonstrate that the *Trichoplax PaxB* gene, *TriPaxB*, is basal not only to all other known *PaxB* genes but also to *PaxA* and *PaxC* genes and their relatives in triploblasts (namely *Pax2/5/8*, *Pax4/6*, and *Poxneuro*). *TriPaxB* is expressed in distinct cell patches near the outer edge of the animal body, where undifferentiated and possibly multipotent cells are found. This expression pattern indicates a developmental role in cell-type specification and/or differentiation, probably in specifying-determining fiber cells, which are regarded as proto-neural/muscle cells in *Trichoplax*. While *PaxB*, *Pax2/5/8*, and *Pax6* genes have been linked to nerve cell and sensory system/organ development in virtually all animals investigated so far, our study suggests that *Pax* genes predate the origin of nerve and sensory cells.

## Introduction

Transcription factors of the *Pax* gene family serve crucial functions in several developmental processes, particularly with respect to the development and differentiation of the central nervous system and sensory organs, both in vertebrates and invertebrates (Walther et al. 1991; Halder, Callaerts, and Gehring 1995; Rinkwitz-Brandt, Arnold, and Bober 1996; Torres, Gomez-Pardo, and Gruss 1996; Callaerts et al. 1999; Czerny et al. 1999; Holland et al. 1999; Kavalier et al. 1999; Kozmik et al. 1999; Groger et al. 2000; Kozmik et al. 2003). A structural characteristic of *Pax* genes is a paired-type DNA-binding domain, which was first identified in the *Drosophila* pair-rule gene *paired* (Frigerio et al. 1986). In addition, most *Pax* genes contain a complete or partial homeodomain, and some *Pax* genes also possess an octapeptide close to the C-terminal of the paired domain. Regions between domains are less well conserved. Based on their structural properties *Pax* genes are grouped into five subfamilies in triploblasts: namely *Pax1-9/Poxmeso*, *Pax2-5-8/sparkling*, *Pax3-7/paired/gooseberry*, *Pax6-4/eyeless*, and *Poxneuro* (Breitling and Gerber 2000; Miller et al. 2000). In diploblasts, *Pax* genes from Cnidaria and Porifera belong to four classes, *PaxA-D* (Balczarek, Lai, and Kumar 1997; Sun et al. 1997; Catmull et al. 1998; Hoshiyama et al. 1998; Groger et al. 2000; Miller et al. 2000; Kozmik et al. 2003). Orthological relationships between diploblast and triploblast *Pax* genes are not finally resolved yet, but it has been proposed that cnidarian *PaxB* genes are related to triploblast *Pax2/5/8* genes, *PaxC* genes to *Pax4-6*, and *PaxD* genes to *Pax1-9* and *Pax3-7* genes (Sun et al. 1997; Miller et al. 2000).

*Pax2/5/8*-related genes seem to be primarily associated with the development of mechanosensory systems in both invertebrates and vertebrates. In higher vertebrates *Pax2*, *Pax5*, and *Pax8* genes are present in multiple paralogs (likely as a result of chromosomal or whole-genome duplications) and are expressed in the developing inner ear and central nervous system in mammals (Rinkwitz-Brandt, Arnold, and Bober 1996; Torres, Gomez-Pardo, and Gruss 1996). In *Drosophila* only one *Pax2/5/8* gene (*D-Pax2*; *sparkling*) is present, which has a crucial function for the development of mechanosensory bristles (Fu et al. 1998; Kavalier et al. 1999), ommatidial cone, and pigment cells (Fu and Noll 1997). A *Pax2/5/8* gene identified in ascidians is expressed in the atrial primordium (Wada et al. 1998), a structure that comprises sensory cells similar to those of the vertebrate inner ear (Bone 1978), and in gastropods, a *Pax2/5/8* gene is expressed in the statocyst (O'Brien and Degan 2003). Interestingly the cnidarian *Pax2/5/8* counterpart, *PaxB*, is implicated in nerve cell differentiation in a hydrozoan (Groger et al. 2000) and in sensory organ (statocyst and eye) development in the cubozoan *Tripedalia cystophora* (Kozmik et al. 2003). The latter mirrors the combined expression (and function) of *Pax6* (eye) and *Pax2/5/8* (statocyst) genes in triploblastic animals. The *Tripedalia Pax* gene, *TcPaxB*, not only unites functional but also structural features of *Pax2/5/8* and *Pax6*-like genes. The paired domain is similar to *Pax2/5/8* genes, whereas the homeodomain displays features of *Pax6*-like genes. Kozmik et al. (2003) demonstrated that the *PaxB* protein is a functional hybrid of *Pax2/5/8* and *Pax6*.

Different hypotheses on the origin of metazoan *Pax* genes have been proposed. One hypothesis suggests that a *PaxA*-like paired domain was fused to a homeodomain and founded the *Pax* gene family (Galliot and Miller 2000; Miller et al. 2000). Breitling and Gerber (2000) postulated that *Pax*-like genes evolved by fusion of a DNA-binding domain of an ancestral transposase (Proto-*Pax* transposase) to a homeodomain shortly after the emergence

<sup>1</sup> Present address: IBU, Neurogenetics, Carl von Ossietzky University, Oldenburg, Germany.

Key words: *PaxB*, *Pax* gene evolution, Proto-*Pax*, *Trichoplax*, Placozoa.

E-mail: thorsten.hadrys@uni-oldenburg.de.

*Mol. Biol. Evol.* 22(7):1569–1578. 2005

doi:10.1093/molbev/msi150

Advance Access publication April 27, 2005

of metazoan animals about 1 billion years ago. The authors further propose a single homeodomain fusion event followed by an early duplication of *Pax* genes before the divergence of Porifera. In order to unravel the early evolution of *Pax* genes we need data from all putative basal metazoan groups. While *Pax* genes have been isolated from sponges and cnidarians, no data have been available from the last and possibly most crucial diploblast phylum, the Placozoa.

Here we report the isolation and characterization of a single *Pax* gene from the morphologically most simple organized metazoan animal, the placozoan *Trichoplax adhaerens*, which lacks any kind of nervous system and/or sensory organs. It is important to note that Placozoa are not secondarily reduced cnidarians (Ender and Schierwater 2003), and thus lack of nerve cells most likely is a plesiomorphy. The *Trichoplax Pax* gene, *TriPaxB*, is expressed in distinct cell patches in a ring-shaped pattern near the lower-upper epithelium boundary.

Our structural and phylogenetic analyses show that the *Trichoplax Pax* gene is basal to *PaxA*-, *B*- and *C*-type genes and harbors structural features of both *Pax2/5/8* and *Pax6* genes. These findings suggest that *TriPaxB* gave rise to at least four of the five *Pax* gene families in higher metazoan animals and provide support for Millers' hypothesis on the origin of *Pax* genes (Miller et al. 2000). The *TriPaxB* gene meets expectations for a Proto-*Pax* gene or the early descendant of a Proto-*Pax* gene in metazoan animals.

## Materials and Methods

### Polymerase Chain Reaction Amplification of Paired-Box Sequences

*Trichoplax* genomic DNA was isolated as described previously (Ender and Schierwater 2003). Messenger RNA (mRNA) from growing and reproducing *Trichoplax* individuals was isolated using the Invitrogen (San Diego, Calif.) "Micro-Fast Track" Kit according to the manufacturer's protocol. Different sets of degenerate primers were used to amplify a 344-bp fragment of the paired domain. Two sets of degenerate primers are described in Hoshiyama et al. (1998). A third set was designed based upon conserved paired domain sequences from other diploblastic *Pax* genes by using the CODEHOP program (<http://blocks.fhcrc.org/blocks/codehop.html>). Complementary DNA (cDNA) preparations from growing and reproducing *Trichoplax* individuals served as template DNA. Double-stranded cDNA was synthesized from mRNA using the "Creator Smart" System (Clontech, Palo Alto, Calif.) according to the manufacturer's protocol. Polymerase chain reaction (PCR) fragments were obtained only with the primer pairs S1-AS3 and S2-AS3; the sequences are as follows (see also Hoshiyama et al. 1998): forward: S1 5'-CAGGATCCCARYTIGGNGGNGTNTT- (corresponding to the "QLGGVF" motif); reverse: AS3 5'-GTGAATTCATYTCCCANGCRAADAT-3' (corresponding to IFAWEI). Other primers designed from highly conserved amino acid sequences within the paired domain did not result in the amplification of *Pax*-specific PCR products. These primers were forward: S2 5'-GAGGATCCTTYGTNAAAYGGNMGNCC-3' (corresponding to FVNGRP); reverse: AS1 5'-GTGAATTCYKRTCNK-

DATYTCCCA-3' (corresponding to WEIRD[RK]); see Hoshiyama et al. (1998), and two primers designed using CODEHOP: forward: S3 5'-CAAGATCTGTGCCGGTACTAYGARACNGG-3' (corresponding to KILSRYYETG) and reverse: 5'-CTCCAGCAGGCAGTCCCKDATYTCCCA-3' (corresponding to WEIRDCLLQ). PCR conditions were 30 s 95°C, 30 s 50°C, and 60 s 68°C, and 40 cycles were performed. PCR fragments were subcloned in pGEM-T vector Promega, Madison, Wisc.). Plasmid minipreparations were sequenced in both directions using ABI (Foster City, Calif.) BigDye terminator chemistry on an ABI-310 capillary sequencer.

### Rapid Amplification of cDNA Ends and Genome Walk PCR

Starting from the paired-box cDNA fragment, the coding sequence of *Trichoplax PaxB* was amplified using the "SMART RACE" system (Clontech). The following primers were designed from the sequence of the isolated paired-box cDNA fragment (forward: 3' Walk 1: ATCAACTACCGTTGGTGTGCCACCT; 3' Walk 2: CGATATGACGACGTATTGCTTCACGC; reverse: 5' Walk 1: CTTGCTTCCTCCAATAATACCTGGGC; 5' Walk 2: CTTCCATTTTCAAACACACCACCCAG). The 3' and 5' rapid amplification of cDNA ends (RACE)-PCR reactions were performed according to the manufacturer's manual (Clontech). PCR conditions were 95°C 15 s, 68°C 3 min, 35 cycles. The obtained RACE products were subcloned (pGEM-T) and sequenced.

To characterize the corresponding *Trichoplax PaxB* gene structure a "Genome Walk" (Clontech) was carried out. PCR reactions were performed using long-template *Taq*-polymerase as described in the manufacturer's manual. PCR fragments were subcloned (pGEM-T) and characterized by sequencing.

### Expression Analyses

Whole-mount in situ hybridization experiments were performed using a modified protocol developed for Cnidaria and Placozoa, respectively (Grogger et al. 2000; Jakob, et al. 2004). Animals were fixed in Lavdowsky's fixative as described in Jakob et al. (2004). *TriPaxB*-, *Trox2*-, and *actin*-RNA-Probes were synthesized from subcloned cDNA fragments (pGEM-T easy; Promega) using digoxigenin (DIG) and fluorescein isothiocyanate-uridine triphosphate (FITC-UTP) labeling (Roche, Mannheim, Germany) according to the manufacturer's manual. Reverse transcriptase (RT)-PCR was done as described previously (Hadrys et al. 2004).

### Phylogenetic Analyses

Distance and Maximum Parsimony analyses were carried out in order to infer phylogenetic relationships between *Pax* genes. All known paired domain sequences from diploblasts were included in the analysis. For rooted tree analyses a *Pseudomonas* transposase sequence served as an out-group (Breitling and Gerber 2000).

Bayes analysis was done with MrBayes (Huelsenbeck and Ronquist 2001). The parsmodel was applied and the

M P H S **▼** G H V T I N Q L G G V F E N G R P L R E A I R R H I V  
 ATGCCGATTTCAGGACCGTTACAATAATCAGCTTGGTGGCGTGTTTGAAAATGGAAGGCCACTGCGTGAAGCAATACGTCGCATATCGTTC  
**Q L A Q S G V R P C D I S R Q L R V S H G C V S K I L C R Y Y Q**  
 AATTAGCTCAAGTGGTGTTCCTCCGTGTGATATACGCCAATTACGAGTATCACATGGATGTGTGATAGATATATATGTCGATATATCA  
**T G S V S P G I I G G S K P K V A T P T V V D K I A E Y K R N**  
 AACGGTTCGTGTAGCCAGGTATTATTGGAGGAAGCAAGCTAAGGTGGCAACACCAAGCGTGTGATAAAATTCGTGAATATAAAGGAAT  
**N S T I F A W E I R E K L L G D K I C D A S N V P S V S S I N**  
 AATTCAACGATATTTCGCTGGGAAATTCGTGAAAATTCGTAGCGGATAAAATTCGCGACGCTAGCAATGTACCTAGTGTAGTTCATCAATCAATC  
**R I V R S K V L S A Y K Y N D G N H S R K Q S A L S T G V E A S**  
 GGATGTTCAGAAAGTTCAGCGCCTACAAGTATAACGACGGCAATCATAGCAGAAAACAATCAGCTCTCTACTGGAGTAGAAGCATC  
**C H Y R I N D L L G** I P Q P Q T S S V G H Y K T A N S S I Q A  
 ATGTCATTATAGAATTAACGACCTATTAGGTATCCACAGCCAGACATCATCTGTTGGCCATTACAAGACTGTAATTCATCCATACAAGCT  
 N D A G L V N D Y T N K S N L I Y S S K M D H Q A A D L L K L Q  
 AACGGTTCGTGTAGCCAGGTATTATTGGAGGAAGCAAGCTAAGGTGGCAACACCAAGCGTGTGATAAAATTCGTGAATATAAAGGAAT  
**K L R R N R T M F T D E Q I K K L E D I F K S T Q Y P D V Y T R**  
 AATTACGCTCAAGTAAACAATGTTTACCGATGAGCAGATTAAGAAGTTAGAAGATATTTTAAATCAACACAGTATCTGTATTTATACCCG  
**E E L A S K I G L S E A R V Q V W F S N R R A K W R K E G K H**  
 AGAAGAATTGGCTTCAAAAATGGATTATCTGAAGCTCGTGTTCAGGTATGGTTTAGCAATCGTCGCGCAAAAATGGCGTAAAGAGGGAAAACAT  
**R S P S K L S N D G S T T L F N S K S T N S S V S E Q S D V C**  
 CGAAGCCCAAGCAACTATCTAATGATGGATCTACTACGTTATTTAATTCAAAAGCACAATTCAGCGTGTAGTGAACAATCTGATGTCTGCA  
 M M L V P  
 TGATGCTAGTGCCAT

FIG. 1.—Nucleotide and amino acid sequence of *TriPaxB*. The isolated coding region comprises 955 nucleotides (318 amino acids). The *TriPaxB* protein includes a paired domain (underlayed in gray), an octapeptide (boxed) and a “full” homeodomain (underlayed in light gray). The positions of two introns are indicated by arrowheads (black). An intron located directly upstream the paired domain was found in all *Pax* genes investigated so far. An intron located between the octapeptide and the homeodomain was also found in other *Pax2/5/8* genes, for example, in *Drosophila D-Pax2* (*sparkling*; here the intron is much longer, however).

following parameters were used: Markov chain Monte Carlo (MCMC) with 100,000 four chains and sampling frequency of 10. The trees generated from the MCMC simulation were imported into PAUP, and a Bayesian tree was visualized using the 50% majority rule option in PAUP (Swofford 2003).

For likelihood ratio tests the method of Shimodaira and Hasegawa (1999, 2001) as implemented in the PROML program in PHYLIP (Felsenstein 2003, <http://evolution.genetics.washington.edu/phylip.html>) was used to calculate likelihood ratios of the best neighbor-joining (NJ) tree and the parsimony tree relative to trees that were generated by RETREE (PHYLIP package) that removed and regrafted the *TriPaxB* gene at all nodes in the parsimony tree.

In this way the basal position of the *TriPaxB* gene could be tested in comparison to its position within the *PaxA–C* clade and within the *PaxB* clade (see fig. 3). The likelihoods of over 20 trees generated by RETREE were included in these tests using a likelihood that took into account site-specific rate differences using a gamma correction.

Accession numbers for sequences included in the analyses are *Acropora millepora PaxA* (*AmPaxA*): AF053458; *A. millepora PaxC* (*AmPaxC*): AF053459; *A. millepora PaxD* (*AmPaxD*): AF241311; *Chrysaora quinquecirrha PaxA1* (*CqPaxA1*): U96195; *Cladonema californicum PaxB* (*CcPaxB*): AF260128; *Chrysaora quinquecirrha PaxB* (*CqPaxB*): U96197; *Drosophila melanogaster eyeless* (*ey*): X79493; *Drosophila melanogaster paired* (*prd*): M14548; *Ephydatia fluviatilis Pax2/5/8* (*EfPax258*): AB007462; *Halocynthia roretzi Pax-37* (*HrPax-37*): D84254; *Hydra littoralis PaxA* (*HIPaxA*): U96194; *Hydra littoralis PaxB* (*HIPaxB*): U96194; *Mus Musculus Pax2*: A60086; *Mus musculus Pax3*: NM\_008781; *Mus Musculus Pax5*: M97013; *Mus musculus Pax6*: BC011272; *Paracene-*

*trotus lividus Pax258* (*suPax258*): AF016884; *Podocoryne carnea* (*PcPaxB*): AJ249563; *Pseudomonas syringae transposase*: AF169828; *T. cystophora PaxB* (*TcPaxB*): AY280703.

**Results**

Isolation and Structural Features of the *T. adhaerens PaxB* Gene

Using different sets of primers, we obtained PCR fragments of the expected size (344 bp) with the primer combination S1/AS3 only. Because primers were designed according to the most conserved regions of the paired-box motif, they are expected to amplify paired-box sequences of all *Pax* gene subfamilies. From a total of 20 clones sequenced all of which were 100% identical in sequence. By means of 5' and 3' RACE 955 nucleotides of the coding sequence, including the paired and homeodomain (318 amino acids) were isolated (fig. 1). The 3' RACE reactions also revealed the presence of two weaker, slightly larger PCR products, indicating the presence at least two alternative transcripts (data not shown).

The *Trichoplax Pax* gene contains a paired domain, an octapeptide, and a “full-length” homeodomain (figs. 1 and 2). The paired domain displays structural features of *PaxB* and *Pax2/5/8* genes, and it harbors several amino acid positions that are regarded as diagnostic for this class of proteins (Kozmik et al. 2003; fig. 1). The full-length homeodomain, however, is more similar to *PaxC/Pax6*-like genes (fig. 2B).

Two exon-intron junctions were mapped via genome walk PCR. The first intron is located directly upstream and adjacent to the paired box (fig. 2A). The location of this first intron is conserved in all *Pax* genes investigated so far. The second intron is located upstream of the homeodomain and



<b>A</b>		<b>Paired domain</b>
TriPaxB		GHVTINQLGGVFENGRPLREAIRRHIVQLAQSGVRPCDISRQLRVSHGCVSKILCRYYQT
TcPaxB		S.GGV.....V.....P.QV..R..E..HQ.....G...E.
EfPax258		.QGGV....L.V.....P.S...K..E.S.N.....G...E.
DmSpar		.GGV.....V.....PDVV.QR..E..HN.....S...E.
MmPax2		R.GGV.....V.....PDVV.QR..E..HQ.....G...E.
AmPaxA		.PGGV.....V.....PDYM.HR..E..HC...SEI...L.....G...E.
AmPaxC		S.GG.....P.V.....PDY..HR.....AC...EI..R.L.....G.F.E.
DmEye		..SGV.....VG....PDST.QK..E..H..A.....I.Q..N.....G...E.
MmPax6		S.SGV.....V.....PDST.QK..E..H..A.....I.Q..N.....G...E.
TriPaxB		GSVSPGIIGGSKPKVATPTVVDKIAEYKRNNTIFAWEIREKLLGDKICDASNVPSVSSI
TcPaxB		..IK.....G..S.....A.P.M.....DR..Q.SV.SQE.....
EfPax258		..IK..V.....SK..L..ED..QE.PS.....DR..Q.GV..KV.....
DmSpar		..FKA.V.....P...A.N...E.P.M.....DR..AEA..SQD.....
MmPax2		..IK..V.....K.....Q.P.M.....AQ..REG..NDT.....
AmPaxA		..R..A.....P..S..LA..ED.PC.....NN..S.GV..K.....
AmPaxC		..IR..S.....P..N..VQ..QQ.P.....DR.VEEGV..RE.T.....
DmEye		..IR.RA.....R...AE..S..SQ...ECPS.....DR..QENV.TND.I.....
MmPax6		..IR.RA.....R...E..S...Q...ECPS.....DR..SEGV.TND.I.....
TriPaxB		NRIVRSKV
TcPaxB		....NRI
EfPax258		....TRA
DmSpar		....N.A
MmPax2		..I.T..
AmPaxA		..L.NAA
AmPaxC		..L.N.A
DmEye		..VL.NLA
MmPax6		..VL.NLA
<b>B</b>		<b>Homeo domain</b>
TriPaxB		LRRNRTMFTDEQIKKLEDIFKSTQYPDVYTRRELASKIGLSEARVQVWFSNRRAKWRKEG
TcPaxB		N.K..YN..P..TDL..QL.EK.P...AT...I.K.TN.....M..QD
EfPax258		EGD.TPTLL.A.VHE..RSLGDCA...TA.VQD..CRL..T.GQI.S.LKA.QPSPAPW.
DmSpar		K.QRMSTYSGD.LYTNISWG.WCIKD.HKLLA..GNLTASTGNCPATYYEASNGFSTTPI
MmPax2		KHLRADT..QQ.LEA.DRV.ERPS...FQAS.HIKSEQGN.YSLPALTPGLDEVKSSLS
AmPaxC		I.....T.SP..LEM..KE.EKSH...A.....DM.....RHQ
DmEye		.Q...S..ND..DS..KE.ER.H...FA..R..G...P...I.....R.E
MmPax6		.Q...S..Q...EA..KE.ER.H...FA..R..A..D.P...I.....R.E

FIG. 2.—Alignment of paired domain (A) and homeodomain (B) sequences of several *PaxA*, *PaxB*, *PaxC*, *Pax2/5/8*, and *Pax6* genes. Amino acids underlined in gray are *Pax6* specific. *Trichoplax PaxB* and *Tripedalia PaxB* harbor paired domains that are *Pax2/5/8* related and homeodomains that are *PaxC/Pax6* related. Am: *Acropora millepora*; Dm: *Drosophila melanogaster*; Ef: *Ephydatia fluviatilis*; eye: *eyeless*; Hl: *Hydra littoralis*; Mm: *Mus musculus*; Pc: *Podocoryne carnea*; spar: *sparkling*; Tc: *Tripedalia cystophora*; Tri: *Trichoplax adhaerens*.

comprises 350 bp. The accession number of the coding sequence is DQ22561.

### Phylogenetic Analyses

In phylogenetic analyses the *TriPaxB* paired domain clusters basal to the *PaxB/Pax2/5/8* subfamily (fig. 3). Furthermore, *TriPaxB* appears to be basal also relative to all but one Pax family. *TriPaxB* always comes out basal to *PaxA*, *PaxB*, and *PaxC* genes, independent of the algorithm and also independent of whether paired domain sequences from triploblastic animals were included or not. The topology shown in figure 3A does not change when randomly chosen paired domain sequences from triploblasts are added to the analysis (fig. 3C).

To test the robustness of the basal position of the *TriPaxB* gene relative to *PaxA*, *PaxB*, and *PaxC* genes, we used the likelihood ratio test as developed by Shimodaira and Hasegawa (1999, 2001). The results of tests using the

PROML program in PHYLIP (Felsenstein 2003) indicated that the tree with *TriPaxB* placed basal to all other *Pax* genes (except for *PaxD*) was the best tree according to likelihood scores (table 1). Furthermore, any tree tested where the *TriPaxB* gene was placed in the *PaxA* clade was highly statistically significantly indicated as worse than the *TriPaxB* basal tree. Placement of the *TriPaxB* gene into the clade in figure 3A that holds most of the other *PaxB* genes, however, indicates that while these trees have worse likelihood scores than the *TriPaxB* basal tree, the trees are not statistically significantly worse. In table 1, tree 1 is the “*TriPaxB* basal” tree. Trees 2–5 and 12 and 13 are trees where *TriPaxB* was grafted onto a *PaxA* or *PaxC* branch in the tree in figure 3A. All other trees except for tree 15 are cases where *TriPaxB* was grafted into places in the *PaxB* clade in figure 3A. Tree 15 retained *TriPaxB* as basal but as sister to the single *PaxD* gene.

A second approach we took was to examine the support for the NJ tree and the parsimony tree using Bayesian

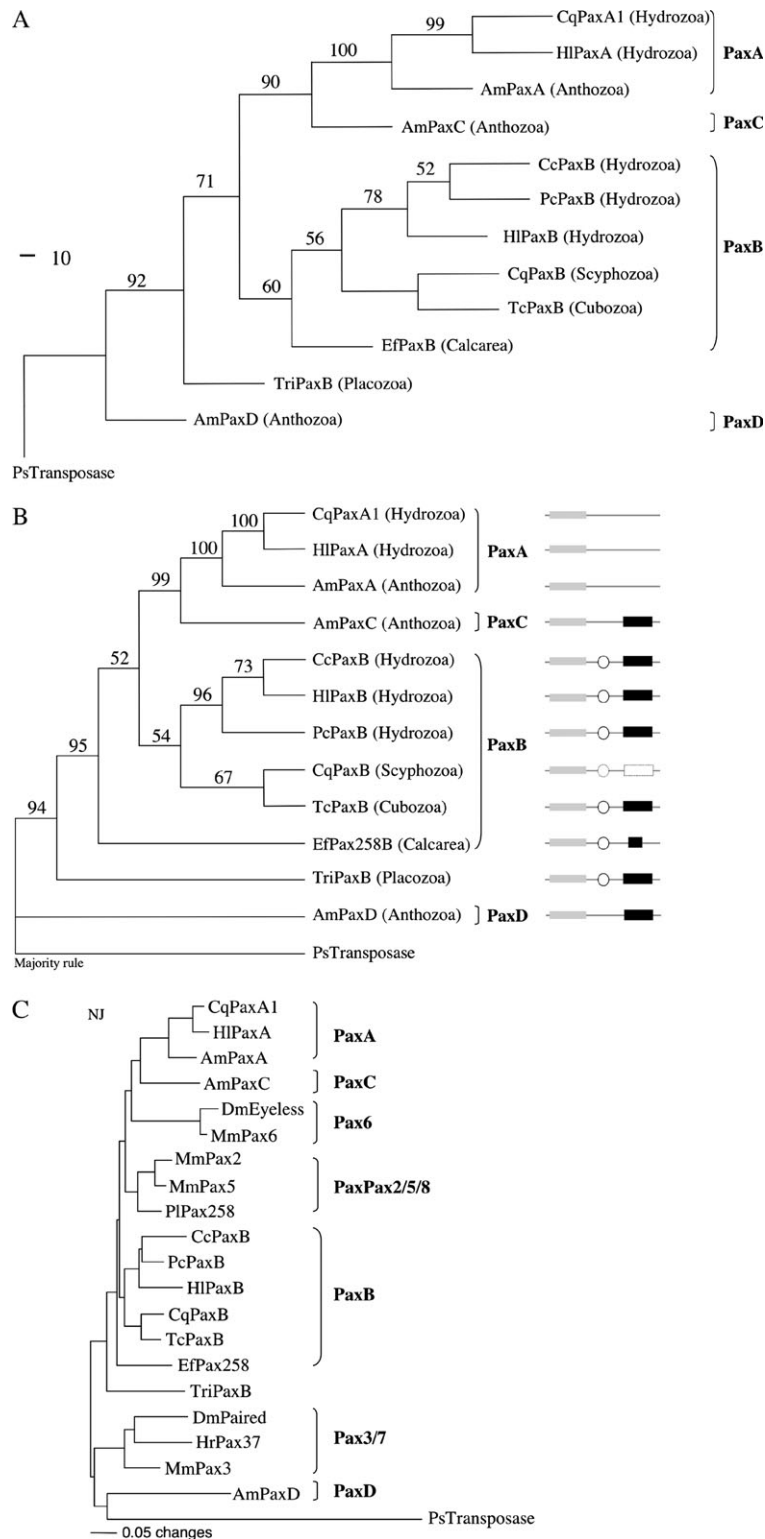


FIG. 3.—(A) Rooted neighbor-joining tree of paired domain sequences of all known diploblast *PaxA*, *PaxB*, *PaxC*, and *PaxD* genes. Bootstrap values result from 1,000 stepwise addition replicates. The shown topology is identical to the parsimony tree. A *pseudomonas* transposase sequence serves as outgroup. (B) Bayesian analysis for all known diploblast *Pax* genes. The Bayes proportions are shown on the branches. At 95% Bayesian proportion, a basal position of the *TriPaxB* gene compared to *PaxA–C* and *PaxB* clades is supported. In this figure the structural features of the corresponding *Pax* genes are also illustrated. Paired domains = light gray boxes, homeodomains = black boxes, octapeptides = white circles. For *CqPaxB* so far only sequences of the paired domain are available. *EfPax258* contains a partial homeodomain. (C) Neighbor-joining tree of all known diploblast together with several triploblast *Pax* gene paired domains. Note that the inclusion of triploblast sequences results in a loss of bootstrap support. Am, *Acropora millepora*; Cc, *Cladonema californicum*; Cq, *Chrysaora quinquecirrha*; Dm, *Drosophila melanogaster*; Ef, *Ephydatia fluviatilis*; Hl, *Hydra littoralis*; Hr, *Halocynthia roretzi*; Mm, *Mus musculus*; Pc, *Podocoryne carnea*; Pl, *Paracentrotus lividus*; Tc, *Tripedalia cystophora*; Tri, *Trichoplax adhaerens*.

**Table 1**  
**Results of Likelihood Ratio Tests (as Developed by Shimodaira and Hasegawa, 1999, 2001) Using the PROML Program in PHYLIP (Felsenstein, 2003)**

Tree	Log L	Differentiated Log L	P Value	Significantly Worse
1	-1,773.0 ← Best			
2	-1,839.3	-66.3	0.000	Yes
3	-1,838.6	-65.6	0.000	Yes
4	-1,822.9	-49.9	0.002	Yes
5	-1,800.1	-27.2	0.041	Yes
6	-1,788.3	-15.3	0.284	No
7	-1,784.2	-11.2	0.429	No
8	-1,785.5	-12.5	0.358	No
9	-1,790.5	-17.6	0.174	No
10	-1,788.2	-15.3	0.268	No
11	-1,781.2	-8.2	0.573	No
12	-1,822.9	-49.9	0.002	Yes
13	-1,800.3	-27.3	0.039	Yes
14	-1,781.2	-8.2	0.569	No
15	-1,773.0	-0.0	0.889	No
16	-1,781.0	-8.1	0.579	No
17	-1,780.9	-7.9	0.588	No
18	-1,783.8	-10.8	0.425	No
19	-1,785.5	-12.5	0.358	No
20	-1,788.8	-15.8	0.214	No

statistics. The Bayesian analysis suggests that the *TriPaxB* gene is not supported as a member of either the *PaxA–C* or *PaxB* clades and supports at 95% Bayesian proportion, the basal position of the *TriPaxB* gene. The Bayes proportions are shown on the branches of the tree in figure 3B.

## Expression of *TriPaxB*

Semiquantitative RT-PCR experiments revealed that *TriPaxB* is expressed in adult, i.e., growing, and vegetatively reproducing animals. Here, *TriPaxB* expression is significantly higher than that of the regulatory *Antp* superclass gene, *EMX* but lower than expression of the *HSP70* gene (fig. 5). Whole-mount in situ hybridization studies revealed expression in distinct cell patches along a ring region close to the outer edge of the animal body (fig. 6A, E, and F). Control hybridization with an *actin* antisense probe shows homogeneous expression throughout the entire body, as expected for a housekeeping gene (Fig. 6B). Control experiments with sense probes did not reveal any specific hybridization signals (data not shown, but see Jakob et al. 2004).

Interestingly, expression signals found in smaller animals were weaker than those found in larger animals (compare fig. 6A and E; data not shown). Analysis of tissue sections revealed that *TriPaxB*-expressing cells are not epithelial cells but cells inside the animal (fig. 6C and D). Possibly these cells are undifferentiated fiber cells. Interestingly, the *Hox/ParaHox* gene, *Trox2*, shows a similar spatial expression pattern in this region of cell differentiation (on average the *Trox2* signal, however, is stronger and more evenly spaced; fig. 6G and Jakob et al. 2004).

## Discussion

The structure of the putative ancestral *Pax* gene has been controversially discussed (cf. Sun et al. 1997; Catmull et al. 1998; Hoshiyama et al. 1998; Breitling and Gerber 2000). The data obtained from *T. adhaerens* strongly suggest

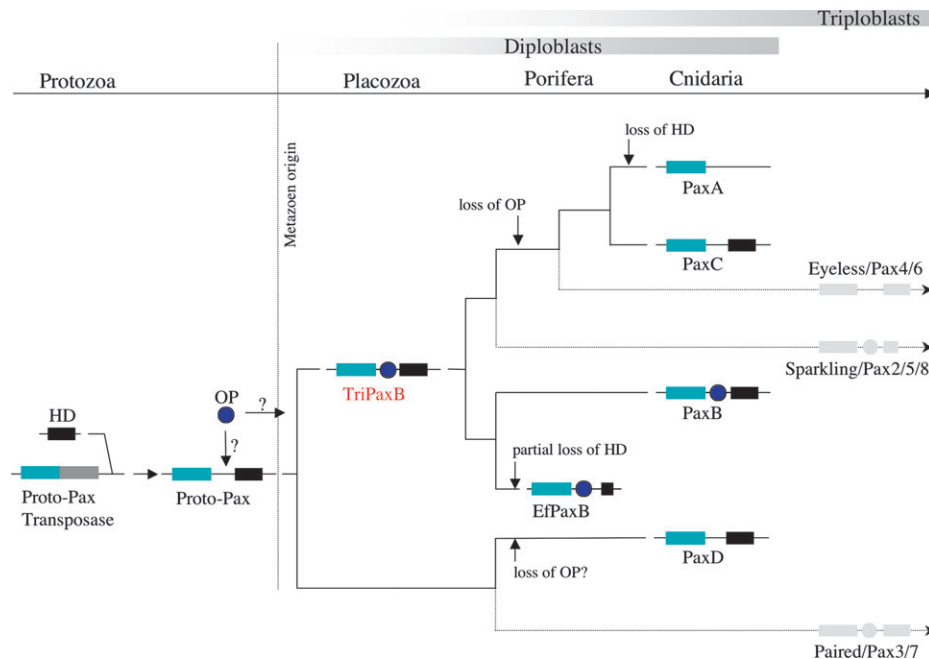


FIG. 4.—*Pax* gene evolution model. A Proto-*Pax* gene (B or D like) derived from a gene fusion event between a Proto-*Pax* transposase and a homeodomain (HD) in protozoans, as first proposed by Breitling and Gerber (2000). The octapeptide (OP) capturing occurred either after or before the first gene duplication event. In the latter case the octapeptide got lost in the D lineage. Two more rounds of gene duplication (B > C and C > A) followed by partial losses of homeodomain and/or octapeptide sequences did lead to the current *Pax* gene assembly in cnidarians. In sponges a partial loss of the homeodomain took place. Proposed relationships to triploblast *Pax258*, *Pax6*, and *Pax37* genes are indicated by dotted lines. For abbreviations see figure 3. Paired domains = green boxes, homeodomains = black boxes, octapeptides = blue circles.



FIG. 5.—Semi-quantitative RT-PCR analysis of *HSP70*, *TriPaxB*, and *EMX* using a cDNA preparation from reproducing stages of *Trichoplax adhaerens*. *TriPaxB* expression is significantly higher than expression of the regulatory *Antp* superclass gene, *EMX*, but lower than expression of the *HSP70* gene.

that a putative “Proto-Pax” gene harbored at least two of the three Pax gene motifs, that is, a paired domain and a homeodomain, and possibly also the third motif, the octapeptide (fig. 4). The transition from an “Ur-Pax” gene (a homeodomain fused to a paired domain; Breitling and Gerber 2000; Miller et al. 2000) to the *Trichoplax PaxABC* gene, requires only a single step, the incorporation of the octapeptide. From this “fully loaded” ancestral *PaxABC* gene other Pax genes may be derived by subsequent deletion events, that is, loss of the (1) homeodomain, (2) partial homeodomain, and/or (3) octapeptide (fig. 4). The supported model is consistent with the paired domain gene trees in figure 3, current knowledge on the phylogenetic position of Placozoa, and the comparison of Pax gene functions.

The Pax evolution model in figure 4 incorporates the assumption that a Proto-Pax gene derived from a gene fusion event between a paired domain (e.g., from a Proto-Pax transposase) (Breitling and Gerber 2000) and a homeodomain in protozoans. In addition to the original arguments, this scenario seems plausible also because no Pax genes have been found in protists and the best-supported Pax paired domain phylogeny is obtained with transposase as out-group. The diploblast Pax gene tree (fig. 3A and B) is in accordance with a proposed basal position of Placozoa (for overview and references see Syed and Schierwater 2002). If Porifera were basal, however, the Pax evolution model would require a slight modification. Here, the last common ancestor of Placozoa and Porifera would be assumed to have harbored a *Trichoplax*-type *PaxB* gene, whose structure remained unchanged in Placozoa but experienced a partial loss of the homeodomain in the lineage leading to Porifera. Because it is known that Placozoa are not secondarily reduced Cnidaria (Ender and Schierwater 2003) and because their bauplan cannot easily be derived from a sponge bauplan (Syed and Schierwater 2002), one has to argue that Pax genes of the A class, and—particularly interesting—also of the B and C class, predated the invention of nerve and sensory cells.

With respect to the evolution of PaxD-like genes additional research is needed to decide whether PaxD branched off even earlier (as shown in our evolution model)

or if the PaxD paired domain originated in the next common ancestor of Cnidaria and triploblasts. In the first scenario a PaxD-type gene either got lost or escaped surveys in Placozoa and Porifera. In the second scenario insufficient taxon sampling or nonoptimal out-group choice may have hindered phylogenetic resolution in the analysis. Although it seems unlikely that a PaxD-type gene escaped our PCR screen, we cannot rule out, however, that Placozoa possess more than one Pax gene. To decide between the alternatives more data will be needed, which will likely come from ongoing whole-genome sequencing efforts in Placozoa, Porifera, and Cnidaria.

Functional information from triploblast Pax genes may also add to our understanding of early duplication events in diploblastic animals. It was previously assumed that PaxB is a precursor of Pax2/5/8, whereas PaxC could be a precursor of Pax6 genes in triploblasts. Plaza et al. (2003) recently demonstrated that DNA-binding characteristics of cnidarian PaxB and PaxC proteins display no simple relationship to Pax2/5/8 and Pax6 genes. The authors showed that *A. millepora* PaxB and PaxC proteins can both bind to *eyeless* (*ey*) targets in vivo and in vitro, which casts doubt on the postulated direct relationship between cnidarian Pax genes and the bilaterian Pax6 and Pax2/5/8 classes. Given that our analysis suggests that Cnidarian PaxB and PaxC (and also PaxA) genes are derived from a gene with similar organization to the placozoan PaxB-like gene, one could speculate that the ancestral PaxABC gene unites functional features which were retained in cnidarian PaxB and PaxC as well as in triploblast Pax2/5/8 and Pax6 genes. This hypothesis is indeed supported by Kozmik et al. (2003) who showed that *T. cystophora* PaxB (the only Pax gene in cubomedusa found so far) contains a Pax2/5/8-type paired domain and octapeptide but a Pax6-type homeodomain. The *Tripedalia PaxB* gene is expressed in larval stages, in the retina, lens, and statocyst. According to functional properties, that is, binding specificity, the ability to rescue *spa* (a *Drosophila* Pax2 mutant) and to induce ectopic eyes in *Drosophila*, the authors suggest that the ancestor of the cubozoan PaxB-like protein was the primordial Pax protein in eye evolution and that Pax6-like genes evolved in triploblasts after separation from Cnidaria. *Trichoplax PaxB* meets these expectations for an ancestral Proto-PaxABC gene. Most interestingly, *Trichoplax* does not possess any kind of sensory organs or nerve cells. Expression of *TriPaxB* in small irregular cell patches along the outer edge of the animal possibly relates to undifferentiated cells and is spatially overlapping with the *Trox2* expression domain (the only *Hox/ParaHox* gene found in Placozoa; see fig. 6G and Jakob et al. 2004). Quite noteworthy, *TriPaxB* is not expressed in differentiated fiber cells, which represent putative proto-neural/muscular cells and are located between the upper and lower epithelium throughout the center region of the body (fig. 6C and D). *TriPaxB* could, however, function in cell determination of fiber cells from undifferentiated and multipotent precursor cells (cf. Jakob et al. 2004). We propose that *TriPaxB* and *Trox2* both demarcate a particular zone of cell proliferation and differentiation.

We further propose that PaxB was co-opted in the last common ancestor of cnidarians and triploblasts for sensory organ and nerve cell development and that two



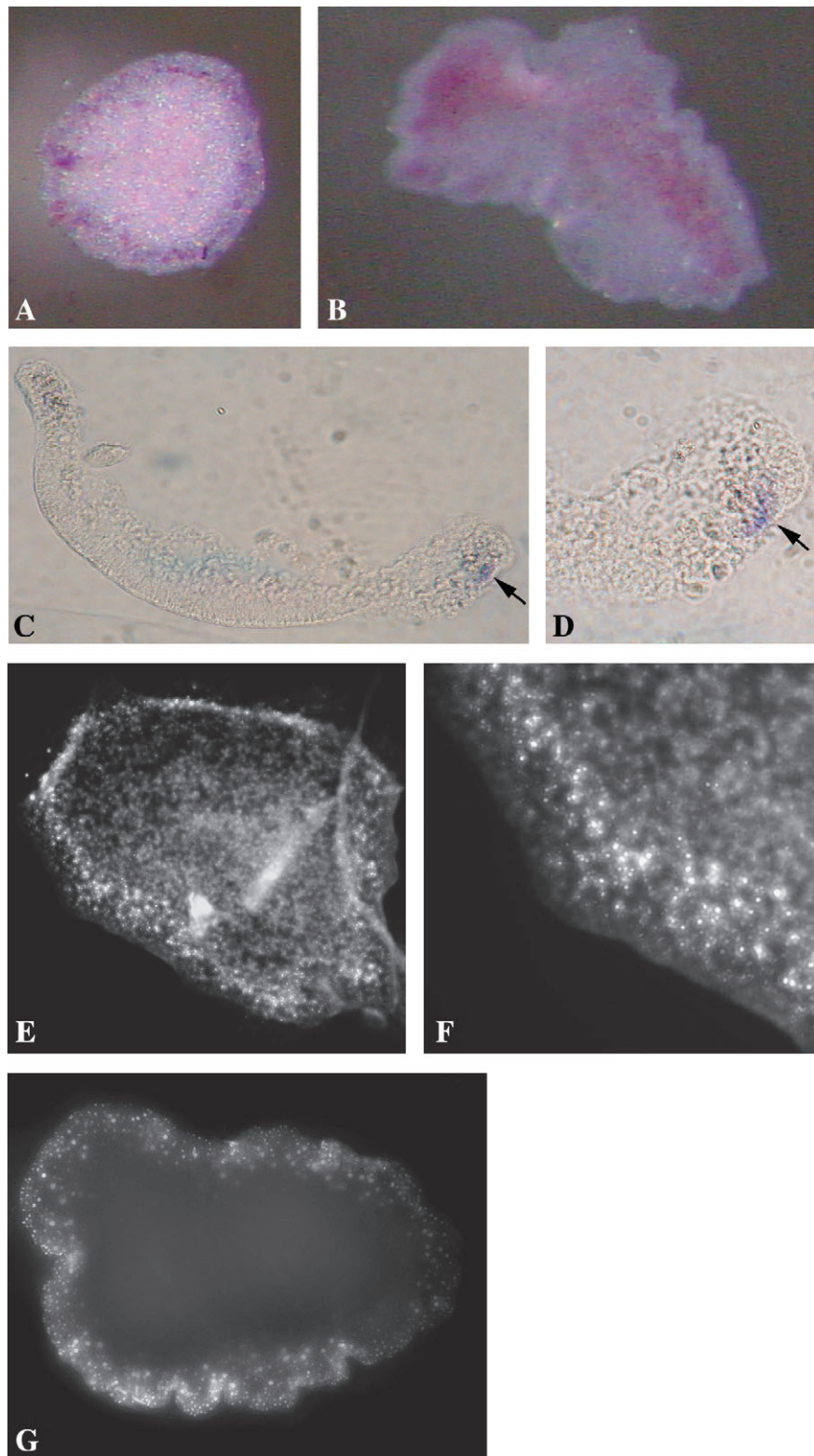


FIG. 6.—Whole-mount in situ expression analyses of *TriPaxB*, *Trox2*, and *actin*. The *TriPaxB* gene is expressed in growing and vegetatively reproducing stages of *Trichoplax adhaerens* (cf. fig. 5) in cell patches at some distance to the outer edge of the animal body (A, C–F). *Actin* (B) and *Trox2* (G) expression is shown for comparison. A–D represent DIG-labeled whole-mount in situ hybridizations, whereas E–G were labeled with FITC-UTP. Note that *TriPaxB* is not expressed in differentiated fiber cells that are located in between the upper and the lower epithelium in the center of the animal body (C and D). *Trox2* (G) is also expressed in the proposed zone of cell proliferation and differentiation close the outer edge of the animal body, but on average stronger and more homogeneously (Jakob et al. 2004).



rounds of gene duplication (B → C and C → A) followed by partial losses of homeodomain and/or octapeptide sequences led to the current Pax gene assembly in cnidarians (fig. 4). In the triploblast lineage additional duplication-deletion events have taken place and among others resulted in the functional split of protein function of Pax2/5/8 and Pax6 genes (fig. 4). Because *TriPaxB* is basal to all other known PaxB genes (and also to PaxA and PaxC genes), it is basal also to Pax2/5/8 and Pax6 genes (Sun et al. 1997; fig. 4). Our data suggest that a PaxB similar gene (harboring functional features of both Pax2/5/8 and Pax6 genes) was the original gene involved in sensory organ development and evolution. A functional split into Pax2/5/8 (mechanosensory) and Pax6 (eye/light sense) likely occurred in the last common ancestor of diploblasts and triploblasts. While the placozoan *TriPaxB* gene most likely predates the origin of nerve and sensory cells, its ancestral developmental function needs to be investigated in more detail.

### Acknowledgments

We acknowledge very helpful comments from two anonymous reviewers and support from the Deutsche Forschungsgemeinschaft (DFG Schi 277/10) and Human Frontier Science Program (HFSP RGP0221/2001-M).

### Literature Cited

- Balczarek, K. A., Z. C. Lai, and S. Kumar. 1997. Evolution of functional diversification of the paired box (Pax) DNA-binding domains. *Mol. Biol. Evol.* **14**:829–842.
- Bone, Q., and K. P. Ryan. 1978. Cupular sense organs in ciona (Tunicata: Ascidiacea). *J. Zool. Lond.* **186**:417–429.
- Breitling, R., and J. K. Gerber. 2000. Origin of the paired domain. *Dev. Genes Evol.* **210**:644–650.
- Callaerts, P., A. M. Munoz-Marmol, S. Glardon, E. Castillo, H. Sun, W. H. Li, W. J. Gehring, and E. Salo. 1999. Isolation and expression of a Pax-6 gene in the regenerating and intact Planarian *Dugesia(G)tigrina*. *Proc. Natl. Acad. Sci. USA* **96**:558–563.
- Catmull, J., D. C. Hayward, N. E. McIntyre, J. S. Reece-Hoyes, R. Mastro, P. Callaerts, E. E. Ball, and D. J. Miller. 1998. Pax-6 origins—implications from the structure of two coral pax genes. *Dev. Genes Evol.* **208**:352–356.
- Czerny, T., G. Halder, U. Kloter, A. Souabni, W. J. Gehring, and M. Busslinger. 1999. Twin of eyeless, a second Pax-6 gene of *Drosophila*, acts upstream of eyeless in the control of eye development. *Mol. Cell* **3**:297–307.
- Ender, A., and B. Schierwater. 2003. Placozoa are not derived cnidarians: evidence from molecular morphology. *Mol. Biol. Evol.* **20**:130–134.
- Felsenstein, J. 2003. *Inferring Phylogenies*. Sinauer Associates, Inc., Sunderland, Mass.
- Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene paired and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**:735–746.
- Fu, W., H. Duan, E. Frei, and M. Noll. 1998. Shaven and sparkling are mutations in separate enhancers of the *Drosophila* Pax2 homolog. *Development* **125**:2943–2950.
- Fu, W., and M. Noll. 1997. The Pax2 homolog sparkling is required for development of cone and pigment cells in the *Drosophila* eye. *Genes Dev.* **11**:2066–2078.
- Galliot, B., and D. Miller. 2000. Origin of anterior patterning. How old is our head? *Trends Genet.* **16**:1–5.
- Groger, H., P. Callaerts, W. J. Gehring, and V. Schmid. 2000. Characterization and expression analysis of an ancestor-type Pax gene in the hydrozoan jellyfish *Podocoryne carnea*. *Mech. Dev.* **94**:157–169.
- Hadrys, T., V. Prince, M. Hunter, R. Baker, and S. Rinkwitz. 2004. Comparative genomic analysis of vertebrate Hox3 and Hox4 genes. *J. Exp. Zool. B Mol. Dev. Evol.* **302**:147–164.
- Halder, G., P. Callaerts, and W. J. Gehring. 1995. Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science* **267**:1788–1792.
- Holland, L. Z., M. Schubert, Z. Kozmik, and N. D. Holland. 1999. *AmphiPax3/7*, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. *Evol. Dev.* **1**:153–165.
- Hoshiyama, D., H. Suga, N. Iwabe, M. Koyanagi, N. Nikoh, K. Kuma, F. Matsuda, T. Honjo, and T. Miyata. 1998. Sponge Pax cDNA related to Pax-2/5/8 and ancient gene duplications in the Pax family. *J. Mol. Evol.* **47**:640–648.
- Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**:754–755.
- Jakob, W., S. Sagasser, S. Dellaporta, P. Holland, K. Kuhn, and B. Schierwater. 2004. The *Trox-2* Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. *Dev. Genes Evol.* **214**:170–175.
- Kavaler, J., W. Fu, H. Duan, M. Noll, and J. W. Posakony. 1999. An essential role for the *Drosophila* Pax2 homolog in the differentiation of adult sensory organs. *Development* **126**:2261–2272.
- Kozmik, Z., M. Daube, E. Frei, B. Norman, L. Kos, L. J. Dishaw, M. Noll, and J. Piatigorsky. 2003. Role of Pax genes in eye evolution: a cnidarian PaxB gene uniting Pax2 and Pax6 functions. *Dev. Cell* **5**:773–785.
- Kozmik, Z., N. D. Holland, A. Kalousova, J. Paces, M. Schubert, and L. Z. Holland. 1999. Characterization of an amphioxus paired box gene, *AmphiPax2/5/8*: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development* **126**:1295–1304.
- Miller, D. J., D. C. Hayward, J. S. Reece-Hoyes, I. Scholten, J. Catmull, W. J. Gehring, P. Callaerts, J. E. Larsen, and E. E. Ball. 2000. Pax gene diversity in the basal cnidarian *Acropora millepora* (Cnidaria, Anthozoa): implications for the evolution of the Pax gene family. *Proc. Natl. Acad. Sci. USA* **97**:4475–4480.
- O'Brien, E. K., and B. M. Degnan. 2003. Expression of Pax258 in the gastropod statocyst: insights into the antiquity of metazoan geosensory organs. *Evol. Dev.* **5**:572–578.
- Plaza, S., D. M. De Jong, W. J. Gehring, and D. J. Miller. 2003. DNA-binding characteristics of cnidarian Pax-C and Pax-B proteins in vivo and in vitro: no simple relationship with the Pax-6 and Pax-2/5/8 classes. *J. Exp. Zool. B Mol. Dev. Evol.* **299**:26–35.
- Rinkwitz-Brandt, S., H. H. Arnold, and E. Bober. 1996. Regionalized expression of Nkx5-1, Nkx5-2, Pax2 and sek genes during mouse inner ear development. *Hear Res.* **99**:129–138.

- Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**:1114–1116.
- . 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**:1246–1247.
- Sun, H., A. Rodin, Y. Zhou, D. P. Dickinson, D. E. Harper, D. Hewett-Emmett, and W. H. Li. 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.
- Swofford, D. L. 2003. PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Version 4. Sinauer Associates Inc., Sunderland, Mass.
- Syed, T., and B. Schierwater 2002. *Trichoplax adhaerens*: discovered as a missing link, forgotten as a hydrozoan, re-discovered as a key to metazoan evolution. *Vie Milieu* **52**: 177–187.
- Torres, M., E. Gomez-Pardo, and P. Gruss. 1996. Pax2 contributes to inner ear patterning and optic nerve trajectory. *Development* **122**:3381–3391.
- Wada, H., H. Saiga, N. Satoh, and P. W. Holland. 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.
- Walther, C., J. L. Guenet, D. Simon, U. Deutsch, B. Jostes, M. D. Goulding, D. Plachov, R. Balling, and P. Gruss. 1991. Pax: a murine multigene family of paired box-containing genes. *Genomics* **11**:424–434.

Claudia Kappen, Associate Editor

Accepted April 3, 2005