Developmental Biology 343 (2010) 106–123

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/developmentalbiology

Evolution of Developmental Control Mechanisms

Origin of Pax and Six gene families in sponges: Single *PaxB* and *Six1/2* orthologs in *Chalinula loosanoffi*

April Hill [a,*,1], Werner Boll [b,1], Carolin Ries [b], Lisa Warner [a], Marisa Osswalt [c,2], Malcolm Hill [a], Markus Noll [b,*]

[a] Department of Biology, University of Richmond, Richmond, VA 23113, USA
[b] Institute of Molecular Life Sciences, University of Zürich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[c] Biology Department, Boston College, Boston, MA 02467, USA

ARTICLE INFO

Article history:
Received for publication 14 October 2009
Revised 11 February 2010
Accepted 16 March 2010
Available online 25 March 2010

Keywords:
Pax genes
Six genes
Pax–Six gene regulatory network
Evolution
Porifera
Haliclona
Chalinula loosanoffi

ABSTRACT

Pax genes play an important role in networks of transcription factors that determine organogenesis, notably the development of sensory organs. Other members of this regulatory network include transcription factors encoded by the Six gene family. Sponges lack organs and a nervous system, possibly because they have not evolved a Pax/Six network. Here we show that the demosponge *Chalinula loosanoffi* encodes only one Pax and one Six gene, representatives of the *PaxB* and *Six1/2* subfamilies. Analysis of their temporal transcription patterns during development shows no correlation of their mRNA levels while their spatial patterns show some overlap of expression in adult tissue, although cellular resolution was not achieved. These results do not suggest that these genes form a major network in this basal phylum, although its existence in a minor fraction of cells is not excluded. We further show that sponge PaxB can substitute for some of the Pax2, but not of the Pax6 functions in *Drosophila*. Finally, we have analyzed the phylogeny of Pax and Six genes and have derived a model of the evolution of the Pax gene subfamilies in metazoans. It illustrates a diversification of Pax genes into subfamilies mostly in triploblasts before the protostome–deuterostome split, whereas few subfamilies were lost in various phyla after the Cambrian explosion.

Introduction

Pax genes encode a small family of transcription factors, including the DNA-binding paired domain and, in some of its subfamilies, also a *prd*-type homeodomain (Noll, 1993; Hanson and van Heyningen, 1995; Chi and Epstein, 2002). Paired domains and Pax genes were discovered in a test of the gene network hypothesis (Bopp et al., 1986, 1989). It was based on our concept of the evolution of gene networks and their integrated functions by the deployment of network-specific domains, regardless of whether these are of a protein-coding or *cis*-regulatory nature (Frigerio et al., 1986). Thus, the evolution of Pax genes and their networks regulating integrated functions (Noll, 1993) was an eminent question ever since their discovery and probable restriction to metazoans (Burri et al., 1989). An early analysis, based on the limited number of Pax genes isolated from *Drosophila* (Bopp et al., 1986, 1989), mouse (Gruss and Walther, 1992), and humans (Burri et al., 1989), predicted the existence of four subfamilies, *Pax2/5/8*, *Pax3/7*, *Pax1*, and *Pax4/6*, before the protostome–deuterostome split, while considering *Drosophila Pox neuro* (*Poxn*) to be a member of the *Pax2/5/8* subfamily (Noll, 1993). This pedigree was extended to include more primitive metazoans, primarily through the isolation and analysis of Pax genes from cnidarians (Sun et al., 1997; Gröger et al., 2000; Miller et al., 2000; Sun et al., 2001; Kozmik et al., 2003; Hoshiyama et al., 2007; Matus et al., 2007), a placozoan (Hadrys et al., 2005), and sponges (Hoshiyama et al., 1998; Larroux et al., 2006). The chief conclusion from these studies was that two of the subfamilies, *Pax2/5/8* and *Pax3/7*, also exist in anthozoan cnidarians as *PaxA/B/C* and *PaxD* genes (Miller et al., 2000; Matus et al., 2007), whereas in medusozoan cnidarians (Sun et al., 1997; Gröger et al., 2000; Kozmik et al., 2003) and sponges (Hoshiyama et al., 1998; Larroux et al., 2006) only a single subfamily, represented by *PaxB* and *PaxA*, was found. It was also proposed that the *PaxA* gene was most closely related to *Poxn* and that the ancestral *PaxC* might have been the precursor of the *Pax6* subfamily (Miller et al., 2000). It remained unclear, however, whether additional paired-domain subfamilies existed in the basal metazoan phylum of sponges (Hoshiyama et al., 1998; Larroux et al., 2006).

Additional interest in the evolution of Pax genes and their networks was raised by the discovery of a conserved role of *Pax6* in eye development of bilaterians (Quiring et al., 1994; for reviews, see

* Corresponding authors. A. Hill is to be contacted at Department of Biology, University of Richmond, Richmond, VA 23113, USA. Fax: +1 804 289 8233. M. Noll, Institute of Molecular Life Sciences, University of Zürich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. Fax: +41 44 635 6829.
E-mail addresses: ahill2@richmond.edu (A. Hill), markus.noll@imls.uzh.ch (M. Noll).
[1] Equal contribution.
[2] Present address: Infinity Pharmaceuticals, Cambridge, MA 02139, USA.

0012-1606/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2010.03.010

Hanson and van Heyningen, 1995; Gehring and Ikeo, 1999; Kozmik, 2005), in agreement with one of the predictions of the gene network hypothesis that gene networks are conserved (Noll, 1993). Ectopic expression in leg, wing, and antennal discs of *Drosophila* Eyeless (Ey) or its mouse homolog Pax6 is able to induce small ectopic eyes on the corresponding adult structures (Halder et al., 1995). This property is even conserved in PaxB of *Tripedalia cystophora*, a jellyfish with complex eyes, and in Pax2 of *Drosophila*, although the ectopic eyes are smaller and induced only on legs (Kozmik et al., 2003). However, ectopic eyes were as big as those induced by Ey when the binding specificity of PaxB was changed to that of Pax6 (Kozmik et al., 2003) by the substitution of only three amino acids in the paired domain (Czerny and Busslinger, 1995). Since PaxB also regulates transcription of the eye crystallin in the jellyfish (Kozmik et al., 2003), these results demonstrate a close relationship of paired domains between the PaxB/2/5/8 and Pax6 subfamilies and suggest that a PaxB-like protein was the primordial Pax protein in eye evolution (Kozmik et al., 2003). In addition, jellyfish PaxB retained Pax2 functions, as shown by the extensive, although incomplete, ability to substitute for the eye-specific function of Pax2 in *Drosophila* spa^{pol} mutants (Kozmik et al., 2003).

Normal or ectopic eye development in *Drosophila* is induced by a network of transcription factors (Chen et al., 1997; Pignoni et al., 1997; Shen and Mardon, 1997; Halder et al., 1998; Treisman, 1999) that is activated by a combination of signaling pathways (Kumar and Moses, 2001) and appears to be conserved in bilaterians (Treisman, 1999; Kawakami et al., 2000). A prominent member of this Pax–Six–Eya–Dac gene network, in addition to the *Pax6* homologs *ey* and *twin of eyeless* (*toy*) (Czerny et al., 1999; Kronhamn et al., 2002), is the *sine oculis* (*so*) gene of *Drosophila* (Cheyette et al., 1994) or its *Six1/2* orthologs in other bilaterians (Seo et al., 1999). Six genes encode a family of transcription factors containing a conserved Six domain, required for interaction with the transcription factor Eya (Pignoni et al., 1997), and an adjacent Six-type homeodomain. Six genes can be grouped into three subfamilies, *Six1/2*, *Six3/6*, and *Six4/5*, that are characterized by diagnostic tetrapeptides close to the N-terminus of their homeodomains (Seo et al., 1999). The presence of Six gene clusters in mammals that include one member of each subfamily led to the proposal of an ancestral Six gene cluster in the progenitor of protostomes and deuterostomes (Boucher et al., 2000). Indeed, such a cluster of Six genes probably already existed before the cnidarians diverged from bilaterians, as Six genes of all three subfamilies were isolated from *Cladonema radiatum*, a jellyfish with lens eyes (Stierwald et al., 2004). Even in the more basal phylum of sponges Six genes exist, although only an incomplete sequence of a Six domain, isolated from the demosponge *Haliclona* sp., has been reported (Bebenek et al., 2004). However, it was neither clear to which subfamily this Six gene belonged nor whether additional subfamilies exist in *Haliclona*. A single Six gene belonging to the *Six1/2* subfamily was also found in the draft sequence of the genome of the demosponge *Amphimedon queenslandica* (Larroux et al., 2008). Like Pax genes, Six genes appear to be restricted to metazoans (Kawakami et al., 2000). No Pax or Six genes were found by BLAST searches in the genomes of the two choanoflagellates, *Monosiga brevicollis* (draft genomic sequence of DOE Joint Genome Institute) and *Proterospongia* sp. (draft genomic sequence of The Broad Institute), a member of the closest extant relatives of metazoans (King et al., 2008), nor are they present in the genome of the filose amoeboid symbiont, *Capsaspora owczarzaki* (draft genomic sequence of The Broad Institute), another unicellular opisthokont closely related to metazoans (Ruiz-Trillo et al., 2004, 2007). Members of both families participate in homologous Pax–Six–Eya–Dac gene networks that are conserved from insects to mammals and not restricted to eye development, but deployed in many other programs of organogenesis (Kawakami et al., 2000), such as in myogenesis (Heanue et al., 1999), nephrogenesis (Xu et al., 2003) or inner ear development (Ozaki et al., 2004).

Given metazoan monophyly, one might surmise that sponges possess a core set of genes that supports all characteristics of being an animal. Recent work strongly suggests that sponges utilize animal-specific genetic pathways to create the sponge body (Adell et al., 2003; Perovic et al., 2003; Hill et al., 2004; Funayama et al., 2005; Larroux et al., 2006). Analysis of the draft sequence of the first sponge genome of *A. queenslandica* (previously named *Reniera* sp.) and of cDNA sequences of the demosponge *Oscarella carmela* illustrates that sponges possess at least one member of many of the transcription factor and signal transduction families known to be crucial in development of complex metazoans (Larroux et al., 2006, 2008; Nichols et al., 2006). Here we isolate a *PaxB* and a *Six1/2* gene from the marine demosponge *Chalinula loosanoffi* (also known as *Haliclona loosanoffi*) and show that they are the only members of their gene families in this species. Thus, these genes are probably the representatives of the primordial Pax and Six genes. Analysis of their temporal expression patterns during development and tissue aggregation does not suggest that they are part of a Pax–Six gene network in a major fraction of cells in which they are expressed. However, such a network is not excluded to exist in a small portion of developing or adult sponge cells. *Chalinula* PaxB can substitute to a large extent for Pax2 in the developing *Drosophila* eye but is unable to induce even small ectopic eyes in *Drosophila* legs, wings, or antennae and hence has no Pax6 function. Finally, we derive a model of Pax gene evolution by an extensive phylogenetic analysis of Pax genes, including Pax genes from several genome projects that we specifically annotated to obtain a better coverage of lower phyla.

Materials and methods

Collection of sponges

Chalinula loosanoffi (this species was recently reclassified from *H. loosanoffi* (Hartman, 1958) to *C. loosanoffi*) was collected from either Long Island Sound at Anchor Beach, Milford, CT, or the Chesapeake Bay at Virginia Institute of Marine Science, Gloucester Point, VA. When necessary, sponges were reared in recirculating seawater tanks, but tissue was usually processed immediately and stored at −80 °C.

Culturing of larvae and primmorphs

Larvae were collected from individual sponges placed in beakers of sterile, filtered seawater. Reproductive sponges released larvae into the water from where newly released larvae were collected with a pipet. Larvae were cultured in 24-well plates in sterile, filtered seawater, which was replaced daily, and were staged as described (Hill et al., 2004). All tissues were stored in RNA stabilization solution RNA*later* (Ambion) overnight and stored at −80 °C for subsequent isolation of RNA.

For culturing of sponge cell aggregates, adult *Chalinula* was dissociated into single cells in Ca^{2+},Mg^{2+}-free seawater and filtered through nylon mesh cups as described (Leith, 1979). Aliquots of dissociated cells were stored in RNA*later* (Ambion). Sponge cell aggregates (primmorphs) were obtained from single cells as described (Custodio et al., 1998; Müller et al., 1999). Primmorphs of approximately 5 mm in diameter were stored in RNA*later* for subsequent RNA isolation. Smaller primmorphs were cultured at 16–22 °C in suspension by constant circulation in natural seawater, supplemented with 0.2% RPMI-1640 medium, 0.060 mM silicate, as described (Krasko et al., 2002). After transfer of the primmorphs to a Petri dish in the same medium supplemented with 0.030 mM Fe^{3+}, circulation was omitted to allow the primmorphs to attach to the substrate, proliferate, and spread across the dish (Krasko et al., 2002). Spreading and differentiating aggregates were cultured for 5 days before they were harvested for RNA isolation.

Isolation of genomic DNA, isolation of RNA, and RT–PCR analysis

Chalinula genomic DNA was isolated by a modified CTAB method (Doyle and Doyle, 1987; Cullings, 1992). After grinding 100 mg of sponge tissue in 0.35 ml of CTAB buffer (50 mM Tris–HCl, pH 8.0, 0.7 M NaCl, 10 mM EDTA, 1% hexadecyl-trimethylammonium bromide, and 0.1% β-mercaptoethanol), 0.35 ml of CTAB buffer was added, and proteins were digested with 0.020 ml of Proteinase K (20 mg/ml) at 65 °C for 1 h. The solution was extracted with 0.60 ml of chloroform–isoamyl alcohol (24:1) and the DNA precipitated with an equal volume of isopropanol at −20 °C overnight. After centrifugation, the pellet was washed with 70% ethanol, allowed to air dry, and the DNA dissolved in TE, pH 8.

Chalinula RNA was isolated from free-swimming larvae or reaggregated adult tissue by use of the Nucleospin RNA II kit (Clontech) and treated with DNAse I to degrade any contaminating genomic DNA. For RT–PCR, 0.50 μg of RNA was reverse-transcribed with the ThermoScript Reverse Transcriptase kit (Invitrogen) and oligo(dT) primers. To detect *PaxB* transcripts, the cDNA was amplified by PCR with Platinum *Taq* DNA polymerase (Invitrogen) and the primers 5′-CCAACTAGGCGGACTCTTCG-3′ (forward) and 5′-CTTTGTCAGGGAGGTCAAGC-3′ (reverse), spanning the first to the fourth coding exon of *PaxB* (positions 1182 to 3069 of the genomic, or positions 149 to 1323 of the cDNA sequence). To detect *Six1/2* mRNA, the cDNA was amplified by PCR with the primers 5′-CGAAAGCGTGCTGAAAGCAAAGG-3′ and 5′-CACGGGGGGATGGGTATGGATTC-3′ that span positions 224 to 543 of the cDNA sequence. Since this sequence did not include an intron, a "cDNA synthesis" reaction without reverse transcriptase was amplified by PCR as a control to exclude artifacts caused by amplification of contaminating genomic DNA. RT–PCR control reactions were performed with *Chalinula*-specific actin gene primers as described (Hill et al., 2004).

In situ hybridization to tissues

Sponge tissues were fixed overnight in sterile seawater, 4% paraformaldehyde, 0.02% glutaraldehyde, transferred into ascending concentrations of methanol, and stored in 100% methanol at −80 °C. Fixed tissues were rehydrated through a methanol and PTw (PBS, 0.1% Tween-20) series and subsequently fixed again in PBS, 4% paraformaldehyde. Tissues were washed in detergent solution (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% SDS, and 0.5% Tween-20) for 30 min, followed by six washes in PTw, processed to a 1:1 solution of PTw and hybridization buffer (5× SSC, pH 5, 50% formamide, 0.050 mg/ml heparin, 0.25% Tween-20, 1% SDS, and 0.100 mg/ml sonicated and denatured salmon sperm DNA), prehybridized in 0.500 ml of hybridization buffer at 60 °C for at least 3 hours, and hybridized with 10 ng of riboprobe at 60 °C overnight. After hybridization, tissues were washed 7 times in hybridization buffer at 60 °C and gradually processed to room temperature through washes in TBST (25 mM Tris–HCl, pH 7.5, 140 mM NaCl, 2.5 mM KCl, 0.1% Tween-20) and hybridization buffer (1:1). After a few washes in TBST, tissues were incubated in 0.500 ml of TBST, 1% BSA to block nonspecific binding of antibodies, and stained histochemically with alkaline phosphatase according to standard protocols. For the preparation of sections, the tissue was embedded in paraffin and cut with a Leica Microsystems RM2245 rotary microtome.

All probes were labeled with the Dig RNA labeling kit (Roche). For *PaxB*, either the entire cDNA or a 371-bp region spanning the paired domain, amplified by the primers 5′-GGCGTGAACCAACTAGGCGGACT-3′ and 5′-CGACTGCGAACGATGCGATTTAT-3′, was used to generate a riboprobe. Both probes produced the same *in situ* staining pattern. For *Six1/2*, a 320-bp region spanning two-thirds of the Six domain and half of the adjacent homeodomain, amplified by the primers 5′-CGAAAGCGTGCTGAAAGCGAAGG-3′ and 5′-CACGGGGGGATGGGTATGGATTC-3′, was used to generate a riboprobe. All DNA fragments were cloned in the TOPO TA Cloning Dual Promoter Vector (Invitrogen), from which both sense and antisense riboprobes were produced that were partially hydrolized to an average length of 100–200 nucleotides suitable for hybridization.

Isolation and sequencing of PaxB gene

Paired box DNA was isolated from genomic DNA by PCR with degenerate primers of conserved paired box sequences, encoding the peptides HGCVSKI and M(I)FA(T)WEIR and biased toward the corresponding Cnidaria and Porifera Pax gene sequences: 5′-CAYGGITGYGTIWSYAARAT-3′ and 5′-CKDATYTCCCAIGYRAAIAT-3′. The resulting DNA fragment of about 450 bp was cloned and verified to encode part of the paired domain by DNA sequencing. This fragment provided the basis for an inverse PCR strategy (Ochman et al., 1988) to walk along the genomic DNA and obtain the complete *PaxB* gene locus. All PCR products were separated by agarose gel electrophoresis; fragments of expected sizes were cut out of the gel, eluted, and cloned by use of the TOPO TA Cloning kit (Invitrogen); and clones were sequenced with the Big Dye Terminator kit (Applied Biosystems) on a 3730 DNA Analyzer (Applied Biosystems).

To determine the exon/intron structure of the *PaxB* gene, *PaxB* cDNA was amplified by PCR from 3′RACE-ready *PaxB* cDNA, synthesized according to the manufacturer's protocol by use of a 3′RACE kit (Clontech) from 1 μg of total RNA isolated from adult *Chalinula* tissue. PCR of the 3′RACE was performed with 5′-CCAACTAGGCGGACTCTTCG-3′ as forward primer (positions 1182–1201 of genomic *PaxB* DNA) and the UPM reverse primer (Clontech). An aliquot of the 3′RACE reaction was used for nested PCR with 5′-GGACTCTTCGTGACCGGTCG-3′ as forward primer (positions 1192–1211 of genomic *PaxB* DNA) and the NUP reverse primer (Clontech) under the same PCR conditions. The remaining 5′ portion of *PaxB* cDNA was isolated by PCR by trying various forward primers, the sequence of which had been derived from genomic *PaxB* DNA sequence. It generated a *PaxB* cDNA spanning the genomic *PaxB* sequence from positions 819 to 5339.

Isolation and sequencing of Six1/2 gene

For Six gene isolation from genomic and cDNA, PCRs were performed with the degenerate primers used previously for the isolation of *Six1/2* and *Six4/5* genes of the jellyfish *C. radiatum* (Stierwald et al., 2004). Single and identical genomic and cDNA fragments of 269 bp, spanning the region encoding amino acids 107 to 196 of Cl-Six1/2, were obtained and extended by 120 bp and 24 bp at their 5′ and 3′ ends, respectively, by use of nested primers whose sequences were obtained from a partial sequence of a *Six1/2* gene isolated from a *Haliclona* sp. (Bebenek et al., 2004). Finally, the cDNA was extended by 5′ and 3′RACE with the GeneRacer RLM-RACE kit (Invitrogen). The *Six1/2* cDNA generated consisted of several fragments covering a total length of 1713 bp. Multiple sets of primers derived from the cDNA sequence were used for amplification by PCR of the corresponding genomic DNA, revealing a single intron of 418 bp interrupting the coding region after the second nucleotide of the codon for Ser207. PCR products were purified and cloned as described above. Clones were sequenced with the SequiTherm EXCEL II kit (Epicenter) on a LiCor System or sequenced with the Big Dye Terminator kit (Applied Biosystems) on a 3730 DNA Analyzer (Applied Biosystems).

DNA constructs and generation of transgenic flies

Chalinula PaxB was expressed in *Drosophila* by the Gal4/UAS system (Brand and Perrimon, 1993). Target constructs were generated by cloning the complete *Cl-PaxB* cDNA into the pP{UAST} vector. The Pax6-specific amino acids I, Q, and N were introduced into the paired domain of Cl-PaxB at positions 42, 44, and 47 of the paired domain by PCR-based mutagenesis of the *PaxB* cDNA by use of the following primers: 5′-GCTATTTAGGTGACACTATAG-3′ (sp6pri), 5′-

CTACTTATGTCGCAAGGACGTA-3′ (splink), 5′-GCGACATAAGTA-GAATCCTACAAGTGTCTAACGGTTGCGTG-3′ (IQNmut), and 5′-GTAA-TACGACTCACTATAGGG-3′ (T7pri).

P-element constructs were coinjected with the transposase donor plasmid pUChsΔ2-3 (designed by D. Rio; FlyBase FBmc0000938) into the posterior pole of stage 2 *y w* embryos. P-element mediated germ line

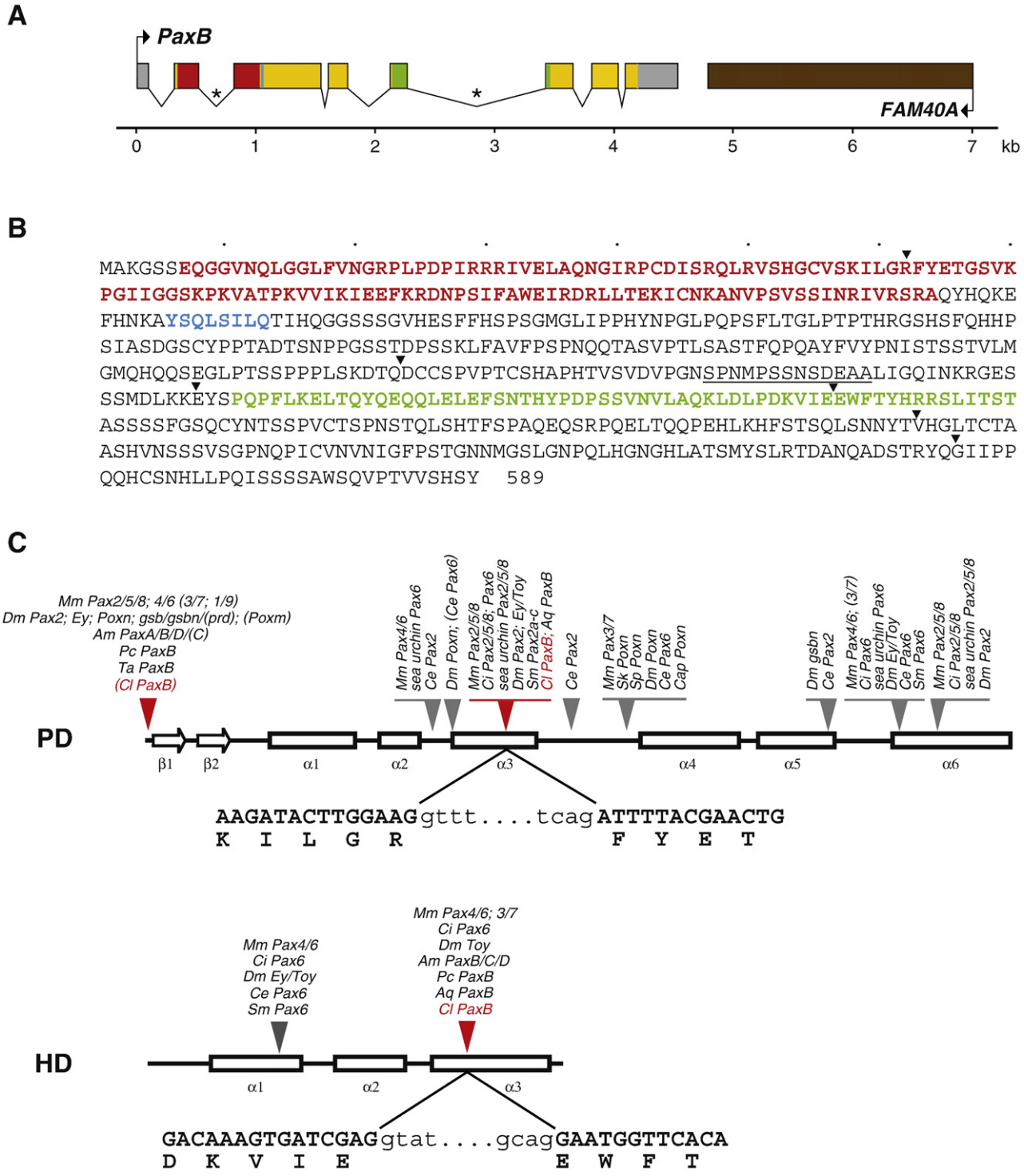
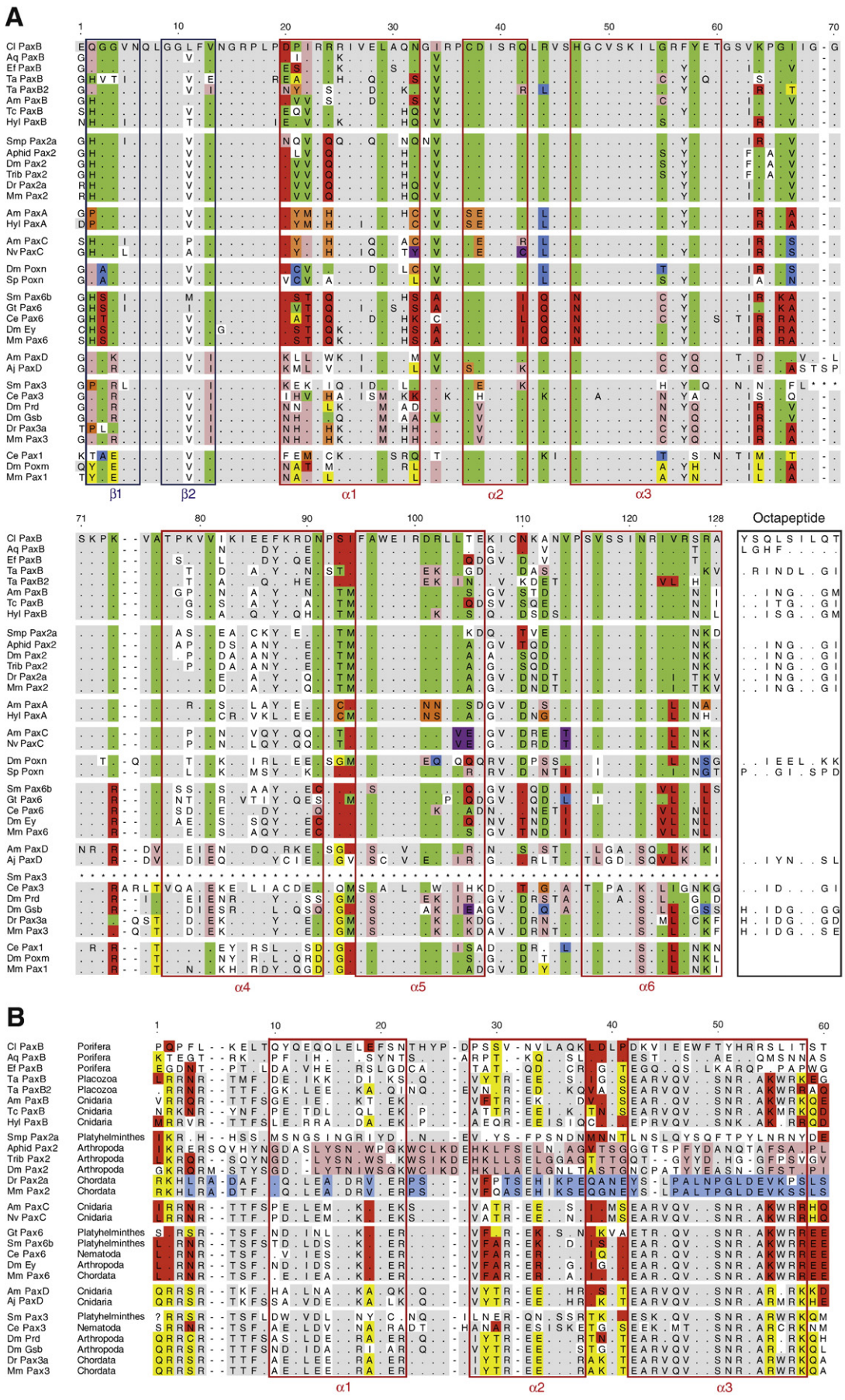


Fig. 1. Structure of the *C. loosanoffi PaxB* gene and its product. (A) Map of the *Cl-PaxB* gene. The eight exons of *PaxB* are mapped with respect to a scale (in kb) shown below whose origin is at the 5′end of the tentative first exon, which is not coding. Regions encoding the paired domain (red), octapeptide (blue), the homeodomain (green), and the remaining part of the protein (yellow) are colored, while untranslated 113 bp 5′-leader and 344 bp 3′-trailer regions are shown in gray. Introns whose positions have been conserved in other Pax genes are marked with asterisks. The brown box downstream of the *PaxB* gene represents a conserved ORF, FAM40A, which is transcribed in a direction opposite to that of *PaxB* (arrows). The GenBank accession number for the genomic DNA and cDNA of the *Cl-PaxB* gene is GQ985310. (B) Sequence of the ORF of the 589-amino-acid Cl-PaxB protein. The paired domain is shown in red, the octapeptide in blue, the homeodomain in green, while the peptide conserved in different sponges is underlined. Positions of introns are marked by filled triangles. (C) Conserved intron positions in the paired domain and homeodomain of Cl-PaxB. Paired domain (PD) and homeodomain (HD) are depicted with their characteristic features in secondary structure, the β-hairpin (β1 and β2) and the six α-helices in the PD and the three α-helices in the HD. Intron positions in these domains are indicated for different Pax genes by gray arrowheads, for *Cl-PaxB* by red arrowheads. For the two conserved introns, the sequence flanking the splice donor and acceptor sites are shown below. Pax genes whose intron positions are close to those conserved strictly in other Pax genes are in parentheses. For abbreviations of species and Pax genes, see Fig. S2.


A
Octapeptide
β1
β2
α1
α2
α3
α4
α5
α6
B
α1
α2
α3

transformation led to several stable transgenic lines, which were established in a *y w* background. To express Cl-PaxB, several of the transgenic lines were crossed to flies carrying *Gal4* drivers, *y w*; *spa-Gal4* (Jiao et al., 2001) or *y w*; *Sp/CyO*; *dpp-Gal4-40C.6/TM6B* (Staehling-Hampton et al., 1994).

Phylogenetic analysis of Pax and Six protein families

Metazoan Pax and Six protein sequences were obtained from the published literature or BLAST searches of the NCBI GenBank, except for sequences from *A. queenslandica* and *Saccoglossus kowalevskii*, obtained from the respective Trace Archive draft genomes at NCBI, and some sequences from *Trichoplax adhaerens* (*Ta PaxB2*) that were retrieved by BLAST search of the *T. adhaerens Grell-BS-1999* v1.0 genome database at the DOE Joint Genome Institute. Pax gene sequence information of *Schmidtea mediterranea* and *Capitella capitata* was retrieved by BLAST search from the SmedGD v1.3.14 (Robb et al., 2008) and the Capca1 database of the DOE Joint Genome Institute, respectively. Paired domain sequences of selected Pax proteins and Six domain and Six-type homeodomain sequences of selected Six proteins were aligned with the multiple sequence alignment program ClustalX version 2.0 with default parameters (Larkin et al., 2007), and alignments were manually refined where necessary.

Maximum likelihood (ML) phylogenetic inferences were derived by use of the PhyML program version 3.0 (Guindon and Gascuel, 2003). For the analyses of Pax and Six proteins, we evaluated the best amino acid substitution model by ProtTest v2.2 (Abascal et al., 2005) and applied the LG model of sequence evolution accordingly (Le and Gascuel, 2008). Two classes of sites were assumed, one class being invariable and the other free to change and follow a gamma shape distribution, which was calculated by use of a discrete approximation with four categories of sites. As starting tree, either the BioNJ tree or five random trees were chosen, and the best resulting tree was regarded as output tree. Tree topologies were estimated with the NNI and SPR methods (Hordijk and Gascuel, 2005). Statistical branch support for the ML analysis was assessed with the approximate likelihood ratio test (aLRT; Anisimova and Gascuel, 2006). Phylogenetic trees are displayed with the Tree Explorer which is included in the Mega4 software (Tamura et al., 2007). Only branches with aLRT values larger than 0.50 are resolved and trees are rooted at midpoint.

Results

C. loosanoffi contains a single Pax gene

To isolate Pax gene DNA from *C. loosanoffi*, genomic DNA was amplified by PCR with degenerate primers specific for highly conserved paired box sequences, which generated a single DNA fragment encoding a partial paired domain of a *PaxB*-type gene. A second amplification by PCR produced additional DNA fragments encoding partial paired domains of other Pax gene subfamilies: *Pax6* [100% identity to zebrafish (NP_571379)] and *Pax3/7* [(87% identity to *Drosophila paired* (AAB59221)]. However, only the PCR product of the first amplification detected a hybridization signal on Southern blots of genomic *Chalinula* DNA, apparent as a single band (Fig. S1), which implies that the sequences amplified by the second PCR did not originate from *Chalinula* but rather from contaminating tissues, probably of other marine metazoans and, more importantly, that the *C. loosanoffi* genome includes only a single Pax gene. This notion of a single Pax gene in sponges is supported by BLAST searches in the NCBI trace archive of the *A. queenslandica* (*Reniera*) genome that yielded only one paired domain, that of *PaxB*, on 23 independent sequences.

Structural features of Chalinula PaxB

Extending the genomic DNA from the *PaxB*-type gene fragment by repeated use of inverse PCR (Ochman et al., 1988) and combining it with the corresponding cDNA, we isolated and sequenced most of the *PaxB* gene as overlapping fragments comprising 5349 bp, at least 4521 nucleotides of which are transcribed. The gene consists of at least eight exons, the last seven of which comprise the entire open reading frame, encoding a protein of 589 amino acids that includes a paired domain, homeodomain, and octapeptide (Figs. 1A and B). The N-terminal paired domain is encoded by exons 2 and 3 that are separated by an intron whose location is conserved among many members of the *Pax2/5/8* and *Pax6* subfamilies in other metazoans (Fig. 1C). A single intron also splits the coding region of the homeodomain at a position conserved in most homeodomains of those Pax subfamilies that include a homeodomain (Fig. 1C).

Immediately downstream of the *C. loosanoffi PaxB* gene, a large open reading frame was detected (Fig. 1A) that encodes a protein of 758 amino acids well conserved in other metazoans, including mammals, nematodes, and insects (NCBI Blast: 42% identity with mouse FAM40A (Q8C079) and 38% identity with CG11526-PA of *Drosophila melanogaster*). The polypeptide contains several putative transmembrane domains, and its N-terminal portion resembles the acidic N1221 domain (PFAM PF07923) in yeast. The function of this putative protein is presently unknown.

The paired domain of *Chalinula* PaxB is most closely related to that of PaxB from other sponges, *A. queenslandica* (91% identity; 96% similarity) and *Ephydatia fluviatilis* (83%; 95%), and very similar to that of cnidarian PaxB or members of the Pax2/5/8 subfamily (>73%; >88%) (Fig. 2A). The octapeptide (Burri et al., 1989), a motif essential for the recruitment of the Groucho corepressor by interaction with the SP domain and possibly the Q/TLE N-terminal domain of Groucho (Eberhard et al., 2000) and conserved in Pax2/5/8 and cnidarian PaxB proteins even as nonapeptide YSINGILGI (Kozmik et al., 2003), is weakly conserved, sharing four identical amino acids with these Pax subfamilies (Fig. 2A). The only other octapeptide identified in a sponge, namely in PaxB of *A. queenslandica*, shares the last 5 amino acids with that of *Chalinula*, but only 2 with octapeptides of cnidarian PaxB or Pax2/5/8 proteins (Fig. 2A). That the octapeptide we identified in *PaxB* of sponges (Fig. 2A) may be functional is suggested by our search in the *A. queenslandica* genome for a Groucho protein, the sequence of which is strikingly conserved. Each of its two highly conserved portions, consisting of a 115-amino acid Q/TLE N-terminal domain and a 300-amino acid C-terminal moiety of seven WD-40 repeats, display about 70% identity and 80% similarity with respect to *Drosophila* Groucho.

The homeodomain of Cl-PaxB is closest to the homeodomain of the Pax6 subfamily (about 38% identity) and to the conserved N-terminal half of the homeodomain of the Pax2/5/8 subfamily in chordates and echinoderms (Fig. 2B). Strikingly, in arthropods, this partial

Fig. 2. Alignments of paired domains, octapeptides, and homeodomains of Pax proteins. (A) Paired domains and octapeptides labeled in the left margin by abbreviations of species and subfamilies of Pax genes. Extent of β-hairpins and α-helices of paired domains are boxed. Dots indicate identities, gray color similarities of amino acids with the Cl-PaxB sequence shown at the top. Gaps in the alignment are marked by dashes and not included in the numbering of amino acids. Paired domain sites diagnostic for different subfamilies are colored as defined in Fig. 5. (B) Homeodomains of Pax proteins labeled as in (A) but including phyla. The three α-helices are boxed. Dots indicate identities in amino acids with the Cl-PaxB sequence shown at the top, while amino acids similar to the most frequent amino acid at a given position are in gray. Gaps in the alignment are marked by dashes and not included in numbering of amino acids. Homeodomain sites diagnostic for Pax3/7 (yellow) and Pax6 (red) protein subfamilies are colored. The homologies extending beyond the partial homeodomains of Pax2 in arthropods and chordates are labeled in pink and blue colors, respectively. Note that the homeodomain of the Pax2-like protein of *S. mansoni*, Smp Pax2a, also extends only over the N-terminal half with about 50% similarity to cnidarian PaxB homeodomains. Fig. S2 shows an expanded and more representative list of 134 complete paired domains and two PAI subdomains with associated octapeptides and homeodomains and includes sequence accession information.

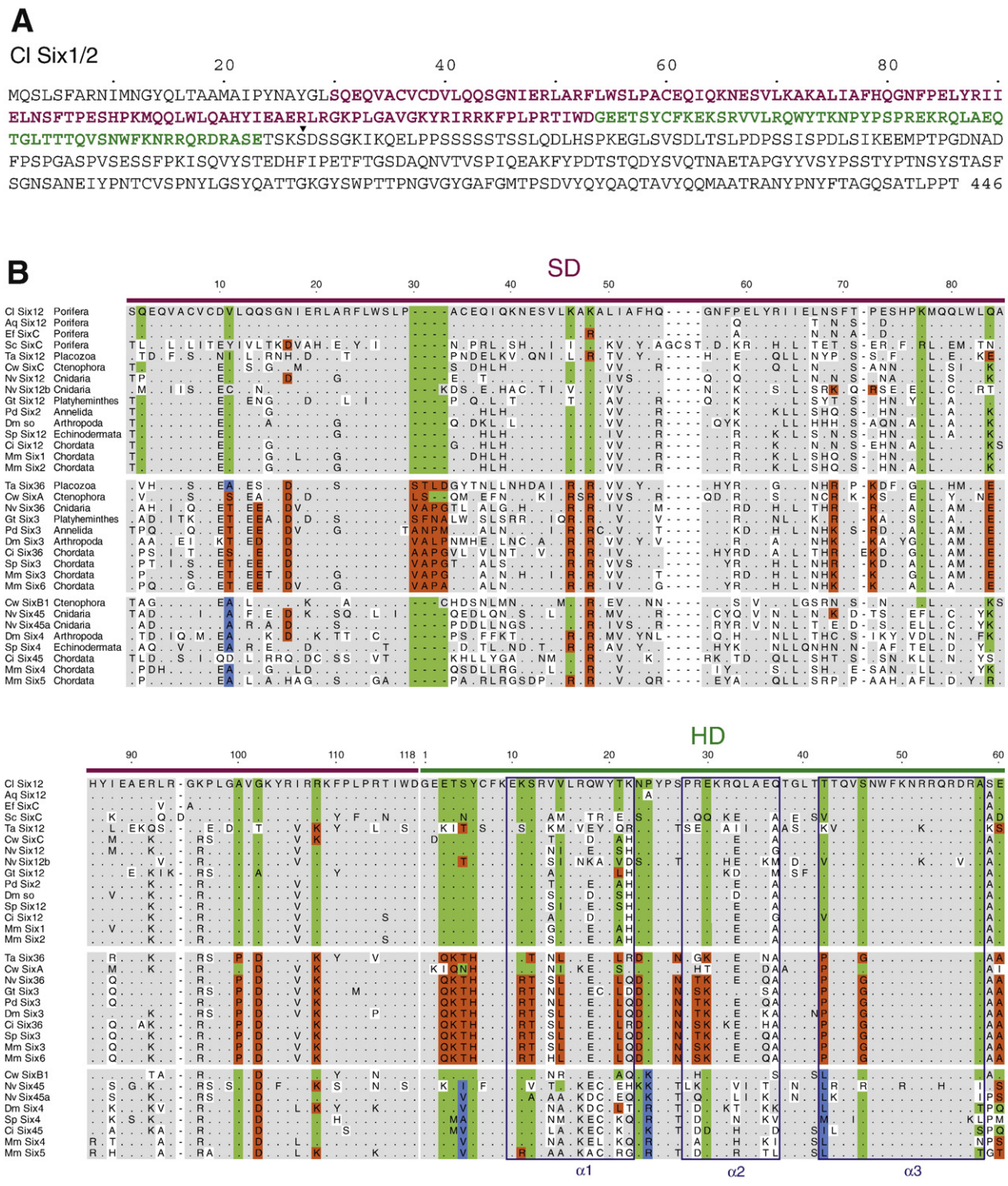


Fig. 3. Six1/2 protein of *C. loosanoffi* and alignment of its Six domain (SD) and adjacent Six-type homeodomain (HD) with other SDs and Six-type HDs. (A) Sequence of the ORF of the 446-amino-acid Cl-Six1/2 protein. The SD (purple) and abutting Six-type HD (green) are colored, and the position of the single intron, which is conserved in human Six genes (Gallardo et al., 1999; Boucher et al., 2000), is marked by a filled triangle. The GenBank accession number for the genomic DNA and cDNA of the *Cl-Six1/2* gene is GQ985311. (B) Alignment of SDs and Six-type HDs. SDs and Six-type HDs are labeled in the left margins by abbreviations of species, subfamilies of Six genes, and, in the upper part, by phyla of species. The extents of SDs and HDs are marked by purple and green lines and those of α-helices in HDs are boxed. Dots indicate identities and gray color similarities of amino acids with the Cl-Six1/2 sequence shown at the top. Gaps in the alignment are marked by dashes and not included in numbering of amino acids. SD and Six-type HD sites diagnostic for different Six subfamilies are colored in green (Six1/2), orange (Six3/6), and blue (Six4/5). An expanded and more representative list of 62 SDs, used for construction of the phylogenetic tree in Fig. 6, is shown in Fig. S3, which includes NCBI accession numbers and full names of species.

homeodomain is conserved only over the 12 N-terminal amino acids with regard to the Pax6-type homeodomain, but extends among arthropods over the length of a typical homeodomain (about 67% similarity between *Drosophila* and *Tribolium* or *Aphid*; marked pink in Fig. 2B). This extension of homology of the partial homeodomain over the length of a true homeodomain is also observed among vertebrate

Pax2/58 proteins (marked blue in Fig. 2B). It remains to be seen whether a similar extension of homology of the partial homeodomain we found in Pax2a of *Schistosoma mansoni* (Fig. 2B) also exists for other Platyhelminthes. Like the homeodomains of other sponge PaxB proteins but in contrast to those well conserved among PaxB proteins in other lower phyla (Cnidaria and Placozoa), the homeodomain of Cl-PaxB displays virtually no sequence conservation in the 5 N-terminal and 7 C-terminal amino acids (Fig. 2B). The distal part of the third α-helix is degenerated, and the amino acids I/V47 and N51, in 'standard' homeodomains essential for the recognition of the (A/T)AAT core binding site (Laughon, 1991; Pomerantz and Sharp, 1994), are not conserved either. Also, other amino acids of the Pax6-type homeodomain proposed to be important for its interaction with DNA as well as the Pax6-type paired domain (Bruun et al., 2005) show only a low degree of conservation. Curiously, in all three sponge genes, the two turns between the three alpha-helices are best conserved in relation to homeodomains of the Pax6 subfamily (Fig. 2B). In addition to the paired domain and homeodomain, we found a 'sponge-specific' triskaidekapeptide SPNMPSSNSDEAA (Fig. 1B), which is the only additional 'domain' that is conserved in PaxB proteins of demosponges (85% identity) but not of other lower metazoans, or in Pax2/5/8 proteins of higher metazoans.

A single Six gene in C. loosanoffi

A Six gene was isolated from *C. loosanoffi* as genomic and cDNA by PCR with degenerate primers (Stierwald et al., 2004), and subsequent 5′- and 3′-extension by RACE of the cDNA, which was used to clone the corresponding genomic DNA (cf., Materials and methods). Since PCR generated only a single DNA fragment, despite the use of degenerate primers able to amplify Six genes of all subfamilies, only one subfamily of Six genes exists in *C. loosanoffi*. This subfamily was most likely represented by a single gene, as evident from Southern blot analysis (Fig. S1). This conclusion is consistent with BLAST searches for the Six domain and adjacent homeodomain in the genome of *A. queenslandica* that detected only a single Six gene, *Six1/2*, on 10 independent sequences in the NCBI trace archive. The gene belongs to the *Six1/2* subfamily because it includes in the N-terminal portion of its homeodomain the tetrapeptide ETSY (Seo et al., 1999; Kawakami et al., 2000), as well as in its Six domain and homeodomain other amino acids that are diagnostic for the So/Six1/2 subfamily of Six proteins (Fig. 3). As is characteristic of Six-type homeodomains (Kawakami et al., 2000), the Cl-Six1/2 homeodomain deviates at two positions conserved in all other homeodomains, R5 and Q12, that are changed to serines (Fig. 3B). The *Cl-Six1/2* gene encodes a highly conserved Six domain (77% identity or 92% similarity with that of mouse, *Mm* Six2; Fig. 3B) and adjacent Six-type homeodomain (88% identity with that of *Mm* Six2; Fig. 3B) on a presumptive first exon, separated by a 418-bp intron from its other exon encoding the C-terminal moiety of the Cl-Six1/2 protein (Fig. 3A). The presence of a single intron at this position has been highly conserved in human *SIX2* and *SIX6* genes (Gallardo et al., 1999; Boucher et al., 2000). The combined domains are best conserved among Six1/2 proteins of demosponges (97% similarity with those of *E. fluviatilis* or *A. queenslandica*). Surprisingly, these domains are considerably less conserved in SixC of the calcareous sponge *Sycon calcaravis* (73% similarity; Hoshiyama et al., 2007) than in Six1/2 of the mouse or So of *Drosophila* (>88% similarity).

Phylogenetic analysis of Cl-PaxB and Cl-Six1/2

Phylogenetic analysis of Pax genes, based on their paired domains, shows the *Cl-PaxB* gene to fall within a well-supported sponge-specific clade that includes Pax genes from marine and fresh water demosponges (Fig. 4). The poriferan Pax genes cluster within the *PaxB/2/5/8* subfamily, but relationships within this clade are poorly resolved. This subfamily forms a clade with the paired domains encoded by the *Poxn/PaxA/PaxC* and *Pax6* subfamilies (Fig. 4). More informative is an approach that considers within paired domains the 41 sites of subfamily-specific amino acids (Fig. 5A). According to this criterion, the sponge *PaxB* sequences resemble the *PaxB* subfamily at least at 36 sites, while they match the *Pax6* subfamily at most at 22 positions.

Phylogenetic analysis of Six genes, based on the amino acid sequences of their Six domains and Six-type homeodomains, unambiguously grouped the *Cl-Six1/2* gene with the Six genes of the marine demosponge *A. queenslandica*, *Aq-Six1/2*, and of the fresh water demosponge *E. fluviatilis*, *Ef-SixC* (Hoshiyama et al., 2007). The calcareous sponge *S. calcaravis* is not included in the same branch, but all poriferan Six genes fall within the *Six1/2* subfamily in the ML analysis which produced a phylogenetic tree that is consistent within the three subfamilies (Fig. 6).

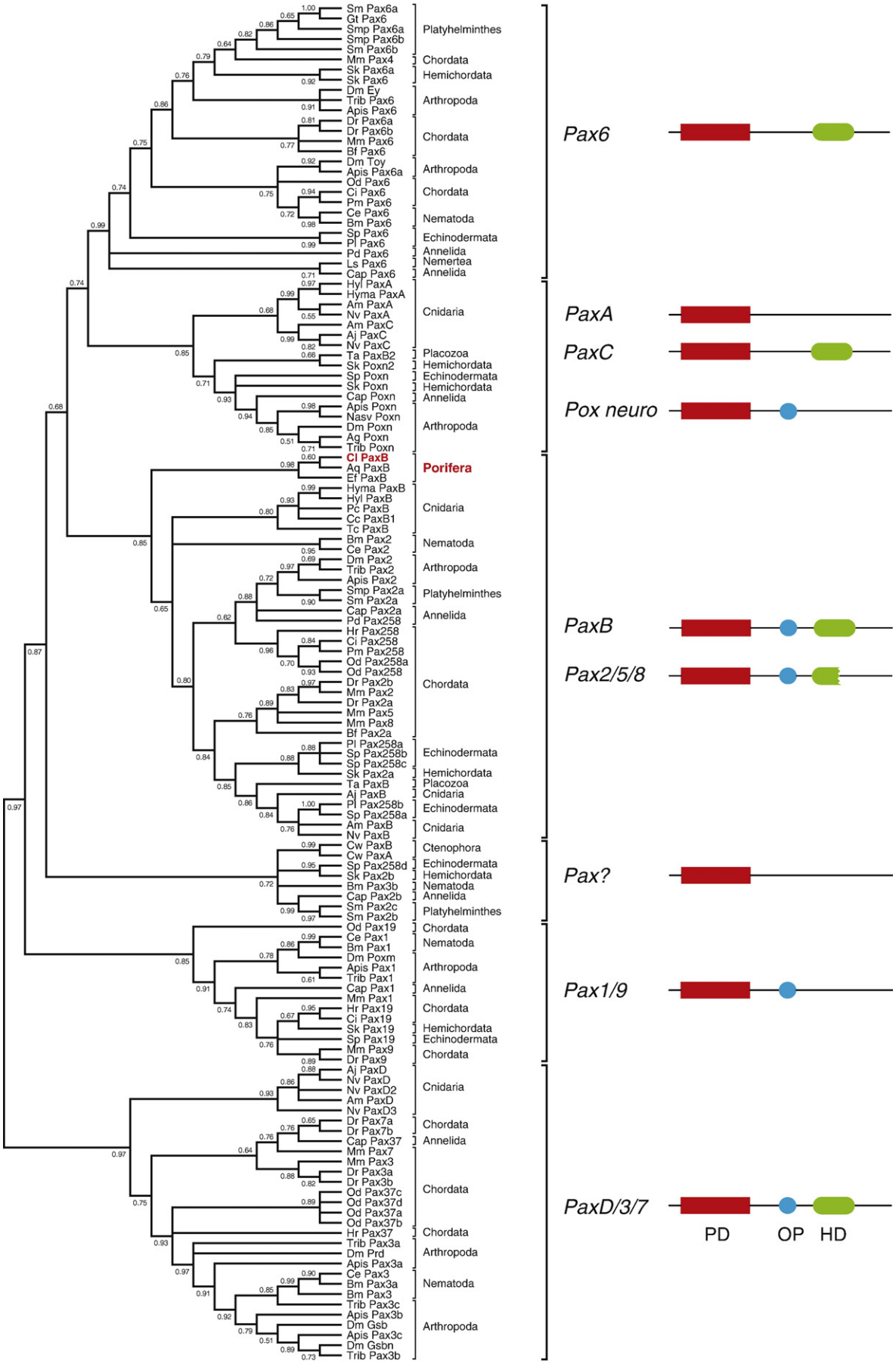
Interestingly, *T. adhaerens*, the only known species of the phylum Placozoa, has two Six genes that clearly belong to two subfamilies, *Six1/2* and *Six3/6* (Figs. 3B and 6). This suggests that the *Six3/6* subfamily was lost in sponges or, what seems more probable, Porifera are basal to Placozoa and the *Six3/6* subfamily originated after the separation of Placozoa from Porifera.

PaxB and Six1/2 expression analysis

Transcript levels of *Cl-PaxB* are very low during all postrelease larval stages, from free-swimming larvae to larvae attached to the substrate, which may imply expression in only a small fraction of all larval cells, whereas overall transcript levels are substantially increased in adult tissue (R in Fig. 7). In contrast, transcript levels of *Cl-Six1/2* are consistently much higher than those of *PaxB* throughout all larval stages and in adults (Fig. 7). Real-time RT–PCR indicates that *Six1/2* transcript levels also increased in adult tissue compared to larval stages when normalized to actin transcript levels (data not shown). Similarly, in dissociated and reaggregated adult cells, a striking difference is observed between transcript levels of *PaxB* and *Six1/2* (Fig. 7). *PaxB* transcript levels are highest during primmorph cellular attachment and spreading, whereas levels in disaggregated tissue and reaggregating primmorphs are low. Conversely, *Six1/2* is most highly expressed in disaggregated cells and reaggregated proliferating primmorphs, while considerably lower levels are apparent during primmorph cellular attachment and spreading (Fig. 7).

Analysis of transcript patterns by *in situ* hybridization showed that *PaxB* transcripts are abundant in reproductive *Chalinula* adults (Fig. 8A) but not detectable in embryos and prerelease larvae that are part of that tissue (inset of Fig. 8A). *Six1/2* is also expressed in adult tissue undergoing reproduction. However, in contrast to *PaxB*, transcripts of *Six1/2* are observed in larvae, mainly in their outer epithelial layer but not in their inner cell mass, as evident from sections of reproductive adult tissue (Fig. 8B). Although RT–PCR analysis demonstrated very low levels of *PaxB* transcripts but very high levels of *Six1/2* transcripts during postrelease larval development (Fig. 7), we were only able to assess spatial expression in earlier stages, i.e., in embryos and prerelease larvae (insets of Figs. 8A and B). After release from the mother sponge, free-swimming larvae of this species are especially fragile, and no fixation methods were able to preserve intact larvae suitable for *in situ* hybridization analysis.

It was possible, however, to compare *PaxB* and *Six1/2* transcript patterns in adult tissue since both genes are transcribed to reach relatively high levels in adult sponges (R in Fig. 7). Preservation of tissue and cell structure in sections of adults is particularly challenging because of their high spicule content (red arrows in Figs. 8C and D); nevertheless, global expression patterns of *PaxB* and *Six1/2* were analyzed in adult tissue. The leuconoid morphology of the adult *Chalinula* sponge includes four major 'regions' (Figs. 8B–D). The first is the atrial opening of the oscular chimney, the osculum's major conduit


Sm Pax6a
Gt Pax6
Smp Pax6a
Smp Pax6b
Sm Pax6b
Platyhelminthes
Mm Pax4
Chordata
Sk Pax6a
Sk Pax6
Hemichordata
Dm Ey
Trib Pax6
Apis Pax6
Arthropoda
Dr Pax6a
Dr Pax6b
Mm Pax6
Bf Pax6
Chordata
Dm Toy
Apis Pax6a
Arthropoda
Od Pax6
Ci Pax6
Pm Pax6
Chordata
Ce Pax6
Bm Pax6
Nematoda
Sp Pax6
Pl Pax6
Echinodermata
Pd Pax6
Annelida
Ls Pax6
Nemertea
Cap Pax6
Annelida
Hyl PaxA
Hyma PaxA
Am PaxA
Nv PaxA
Cnidaria
Am PaxC
Aj PaxC
Nv PaxC
Ta PaxB2
Placozoa
Sk Poxn2
Hemichordata
Sp Poxn
Echinodermata
Sk Poxn
Hemichordata
Cap Poxn
Annelida
Apis Poxn
Nasv Poxn
Dm Poxn
Ag Poxn
Trib Poxn
Arthropoda
Cl PaxB
Porifera
Aq PaxB
Ef PaxB
Hyma PaxB
Hyl PaxB
Pc PaxB
Cc PaxB1
Tc PaxB
Cnidaria
Bm Pax2
Ce Pax2
Nematoda
Dm Pax2
Trib Pax2
Apis Pax2
Arthropoda
Smp Pax2a
Sm Pax2a
Platyhelminthes
Cap Pax2a
Pd Pax258
Annelida
Hr Pax258
Ci Pax258
Pm Pax258
Od Pax258a
Od Pax258
Dr Pax2b
Mm Pax2
Dr Pax2a
Mm Pax5
Mm Pax8
Bf Pax2a
Chordata
Pl Pax258a
Sp Pax258b
Sp Pax258c
Echinodermata
Sk Pax2a
Hemichordata
Ta PaxB
Placozoa
Aj PaxB
Cnidaria
Pl Pax258b
Sp Pax258a
Echinodermata
Am PaxB
Nv PaxB
Cnidaria
Cw PaxB
Cw PaxA
Ctenophora
Sp Pax258d
Echinodermata
Sk Pax2b
Hemichordata
Bm Pax3b
Nematoda
Cap Pax2b
Annelida
Sm Pax2c
Sm Pax2b
Platyhelminthes
Od Pax19
Chordata
Ce Pax1
Bm Pax1
Nematoda
Dm Poxm
Apis Pax1
Trib Pax1
Arthropoda
Cap Pax1
Annelida
Mm Pax1
Hr Pax19
Ci Pax19
Chordata
Sk Pax19
Hemichordata
Sp Pax19
Echinodermata
Mm Pax9
Dr Pax9
Chordata
Aj PaxD
Nv PaxD
Nv PaxD2
Am PaxD
Nv PaxD3
Cnidaria
Dr Pax7a
Dr Pax7b
Chordata
Cap Pax37
Annelida
Mm Pax7
Mm Pax3
Dr Pax3a
Dr Pax3b
Od Pax37c
Od Pax37d
Od Pax37a
Od Pax37b
Chordata
Hr Pax37
Chordata
Trib Pax3a
Dm Prd
Apis Pax3a
Arthropoda
Ce Pax3
Bm Pax3a
Bm Pax3
Nematoda
Trib Pax3c
Apis Pax3b
Dm Gsb
Apis Pax3c
Dm Gsbn
Trib Pax3b
Arthropoda
Pax6
PaxA
PaxC
Pox neuro
PaxB
Pax2/5/8
Pax?
Pax1/9
PaxD/3/7
PD
OP
HD

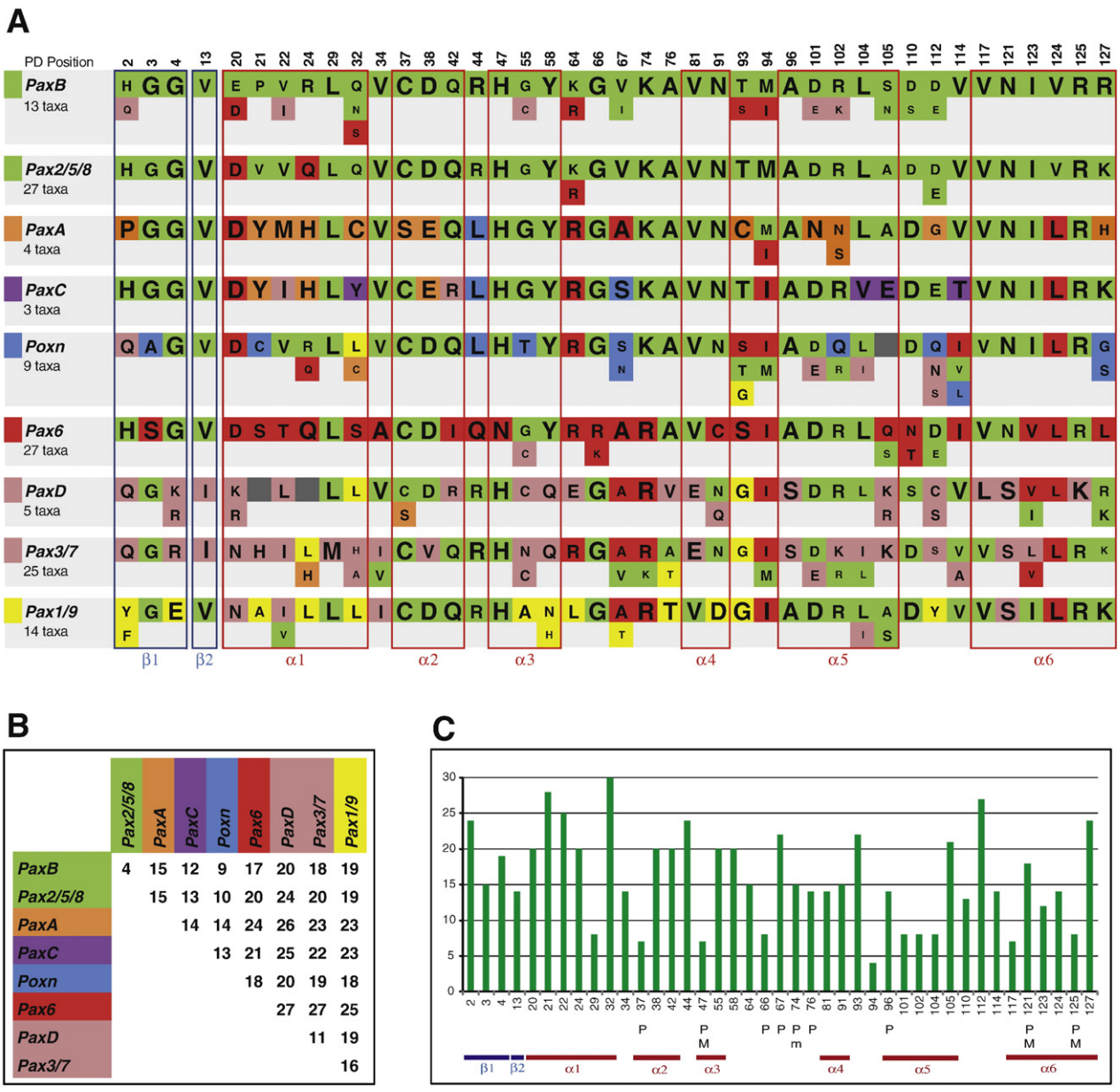


Fig. 5. Diagnostic sites of paired domains. (A) Amino acids at 41 diagnostic positions, numbered at the top, of paired domain subfamilies are shown. At each of these positions, the amino acid of at least one subfamily differs characteristically from that in other subfamilies. Amino acids with a single incidence or a frequency below 20% at a certain position within a subfamily are not listed. One or several amino acids diagnostic for a subfamily are marked by the color of the subfamily indicated in the left margin. If more than one subfamily shares the same amino acid at a position, the amino acid is considered to be diagnostic for the subfamily that displays the highest incidence of this amino acid within the subfamily, weighed by the number of paired domains of that subfamily among the 126 paired domains analyzed (Fig. S2). The amino acids are shown in four different letter sizes, with decreasing size indicating frequencies ≥90%, ≥66%, ≥33%, and <33%. (B) Table listing numbers of differences at diagnostic positions between paired domain subfamilies. Only differences equal to or larger than a factor of 5 in frequency of the amino acid are considered to be significantly different. (C) Frequencies of amino acid changes at diagnostic sites of paired domains. The number of amino acid changes taken from (A) is plotted as a function of their position within the paired domain. Below the histogram, contacts of amino acids with the DNA backbone (P) and/or base in the major (M) or minor (m) groove are indicated (Xu et al., 1999).

through which the water is extruded. A pinacodermal layer of cells separates the atrial opening of the oscular chimney from the sponge body, the largest portion of which consists of the choanosome or mesohyl. This internal region comprises the choanocyte chambers consisting of choanocytes, their associated canals, and a variety of other cell types (e.g., archeocytes and schlerocytes). Finally, the outermost region is the external pinacodermal layer lining the surface of the sponge facing the environment. *PaxB* and *Six1/2* are transcribed in the pinacodermal lining of the atrial opening of the oscular chimney (black arrows in Figs. 8B–D) as well as in the immediately adjacent choanosome (Figs. 8C and D). However, it is unclear whether they are transcribed in the same cells. Transcripts of *PaxB* appear more abundant in the pinacodermal cell layer lining the oscular chimney than those of *Six1/2*. Although transcripts of both *PaxB* and *Six1/2* were

Fig. 4. Phylogenetic tree of Pax genes. The cladogram, inferred by maximum likelihood, was obtained by the analysis of 133 paired domains (Fig. S2). Next to the branches, their statistical support is indicated as approximate likelihood ratio test (aLRT) values. For branches with aLRT values below 0.50 forking is not resolved. For NCBI accession numbers of paired domain sequences, see Fig. S2. Eight paired domains from Ctenophora, Platyhelminthes, Annelida, Nematoda, Echinodermata, and Hemichordata formed a clade but could not be grouped unambiguously with any of the paired domain subfamilies and hence were labeled '*Pax?*' (see also legend to Fig. S2). To the right of the names of the Pax proteins and their cognate phyla, Pax protein subfamilies are schematically indicated with their characteristic domains, the paired domain (red), octapeptide (blue), and homeodomain (green).

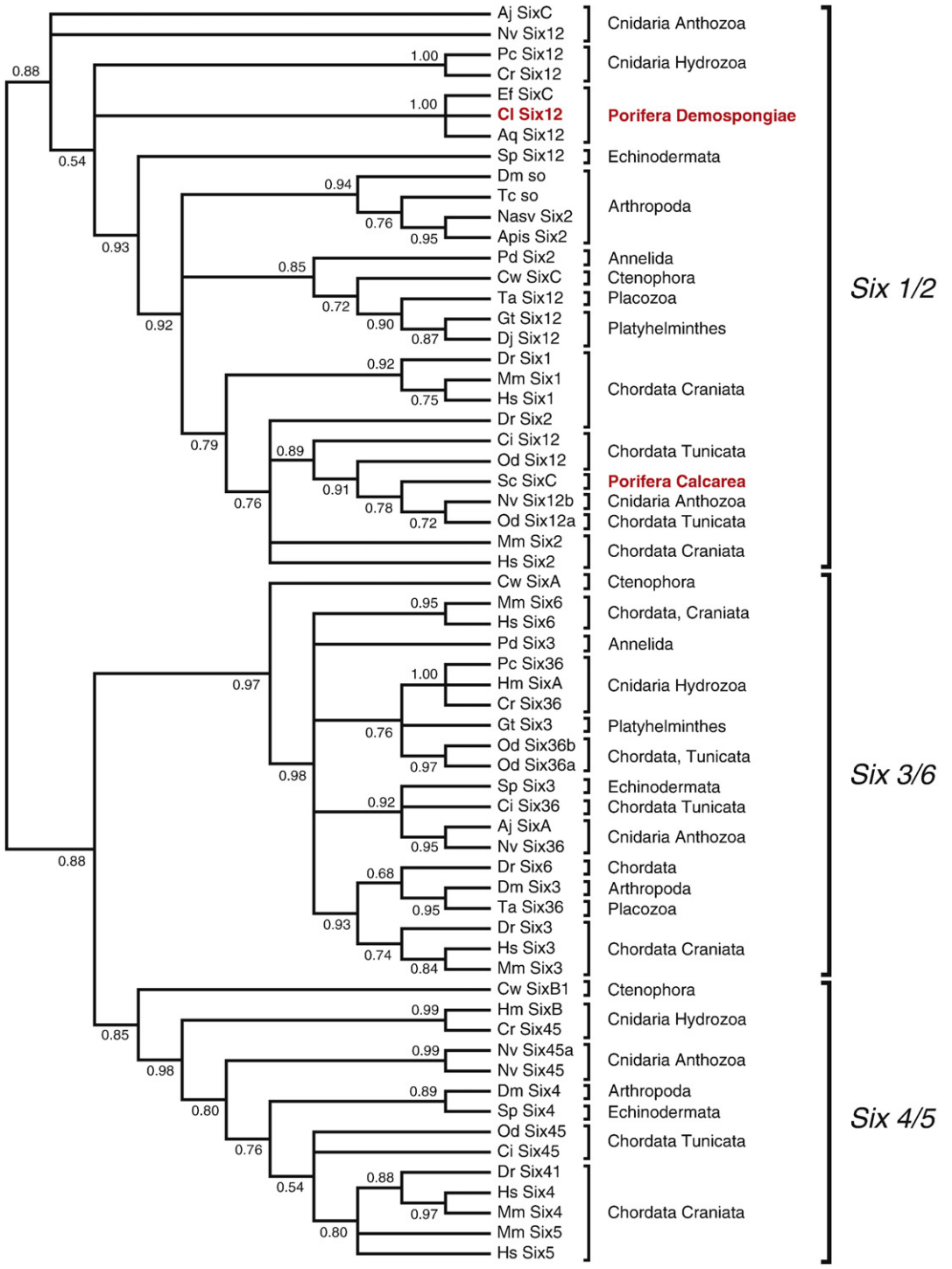


Fig. 6. Phylogenetic tree of Six genes. The cladogram, inferred by maximum likelihood, was obtained by the analysis of the 62 Six domains and Six-type homeodomains shown in Fig. S3. Next to the branches, their statistical support is indicated as approximate likelihood ratio test (aLRT) values. For branches with aLRT values below 0.50 forking is not resolved. For accession numbers of Six domains and Six-type homeodomains, see Fig. S3. The Six proteins, whose names and phyla are indicated to the right of the cladogram, are grouped into three subfamilies (Seo et al., 1999) labeled at the right.

detected within the choanosome, transcripts of neither gene were detectable in a narrow band of the outer choanosome and the adjacent external pinacodermal layer (inset in Fig. 8C). It is not probable that either gene was expressed at high levels in choanocytes because choanocyte chambers have a defined, circular shape. Since it was not possible to determine whether *PaxB* and *Six1/2* were transcribed in the same nor in which cell type because of tissue damage caused by spicules during sectioning, this level of resolution will require another sponge model whose tissue can be better preserved for *in situ* analysis. We are currently pursuing these questions in the freshwater sponge *Ephydatia muelleri*, where the global expression pattern we observe is similar to that of *Chalinula* (data not shown). Although global coexpression is observed for *PaxB* and *Six1/2* in some regions of adult tissue (Figs. 8B–D), given the nonoverlapping expression profiles during embryonic development (Figs. 8A and B), larval development (Fig. 7), and cell reaggregation (Fig. 7), it is difficult to reconcile these

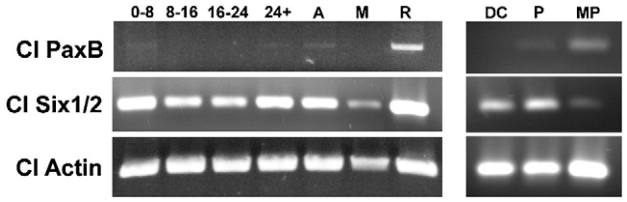


Fig. 7. Transcript levels of *Cl-PaxB* and *Cl-Six1/2* in developmentally staged *Chalinula* tissues. *Chalinula PaxB* (Cl-PaxB) and *Six1/2* (Cl-Six1/2) transcript levels were analyzed by RT–PCR in free swimming larvae (0–8 h, 8–16 h, 16–24 h, and more than 24 h old), attached larvae (A), metamorphosing larvae (M), rhagon/juvenile adults (R), dissociated adult cells from disaggregated tissue (DC), primmorph aggregates (P), and metamorphosing primmorph tissue (MP). Transcript levels of actin (Cl-Actin), which are constant throughout development, are shown as control below.

results with exclusive roles of *PaxB* and *Six1/2* in a Pax–Six network of *Chalinula*. Our results rather suggest that many cells express *Six1/2* in which *PaxB* is inactive or transcribed at much lower levels during development and in adults.

Cl-PaxB partially rescues the spa^{pol} eye phenotype

Previous experiments have shown that a cnidarian PaxB can substitute for Pax2 functions in the *Drosophila* eye (Kozmik et al., 2003). Similarly, we tested the ability of a sponge PaxB to substitute for the D-Pax2 function in *spa^{pol}* mutants of *Drosophila*, in which the eye-specific enhancer *spa* of *D-Pax2* is deleted (Fu and Noll, 1997). These mutants exhibit severe eye defects resulting from the missing *D-Pax2* expression in cone and primary pigment cells of developing eye discs (Fu and Noll, 1997). Thus, Cl-PaxB protein was expressed in *spa^{pol}* mutants under the indirect control of the *spa* enhancer by *spa-Gal4* driving the expression of *UAS-Cl-PaxB*. Whereas wild-type eyes have regular hexagonal facets with mechanosensory bristles projecting from alternate facet vertices (Fig. 9F), *spa^{pol}* eyes are reduced in size, their ommatidia form irregular arrays of variable size with broken or missing bristles, and their corneal lenses are irregular, often fused, and contain necrotic pits (Fig. 9A). Each of eight independent transgenic *UAS-Cl-PaxB* lines was able to partially rescue the *spa^{pol}* phenotype when combined with *spa-Gal4* (Figs. 9B–D). The rescued eyes seemed to be normal in size, showed more regularly arranged ommatidia, and bristles were present over the entire surface of the eye. However, necrotic pits were still observed, mainly in the anterior half of the eyes, affecting in different transgenic lines between 1% and 10% of the ommatidia (Figs. 9B–D). By contrast, rescue with D-Pax2, also driven by *spa-Gal4*, was nearly complete, and no necrotic pits were apparent (Fig. 9E).

Sponge PaxB protein is unable to induce ectopic eyes in Drosophila

A cnidarian PaxB is also able to substitute to some extent for eye-specific Pax6 functions by inducing ectopic eyes in *Drosophila* (Kozmik et al., 2003), which are, however, smaller than those induced by the Pax6 homologs, Ey or Toy (Halder et al., 1995; Czerny et al., 1999). Accordingly, we tested whether sponge PaxB can similarly substitute for

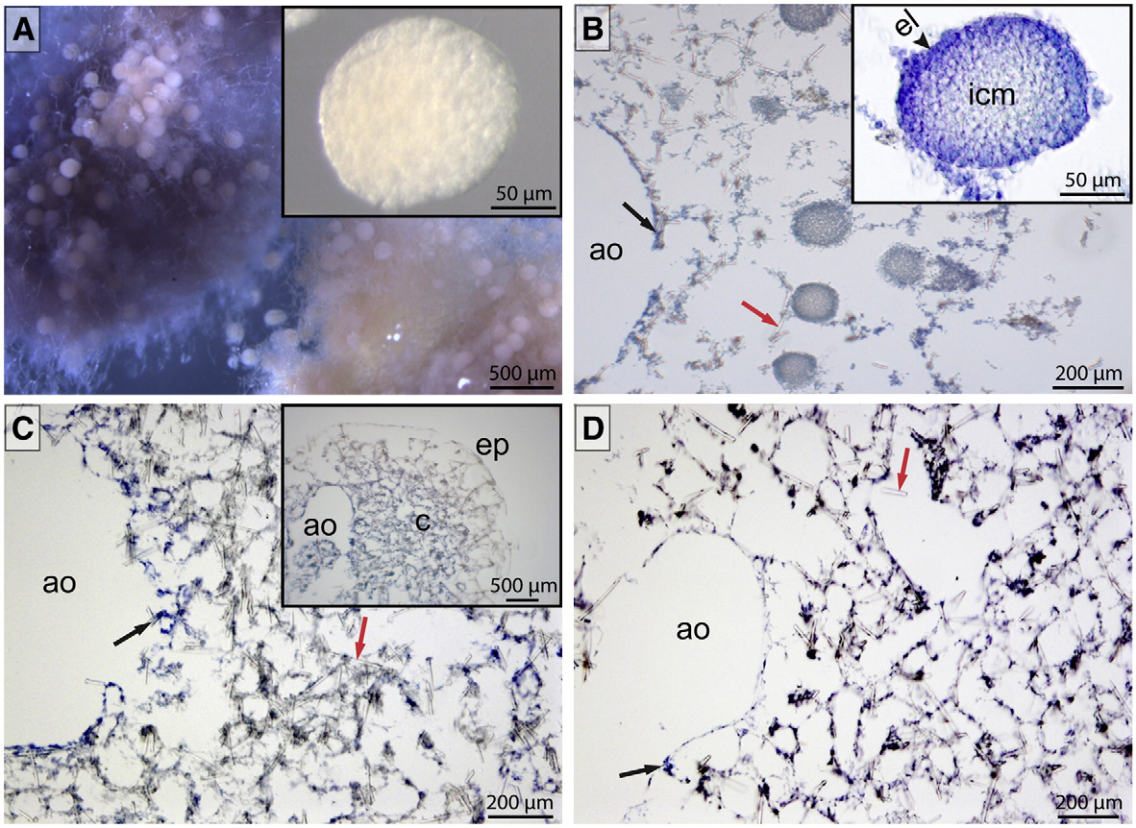


Fig. 8. Patterns of *Cl-PaxB* and *Cl-Six1/2* transcripts in prerelease larvae and in reproductive and nonreproductive adults of *C. loosanoffi*. *PaxB* (A, C) and *Six1/2* (B, D) transcripts are visualized by immunohistochemical staining with alkaline phosphatase after *in situ* hybridization of DIG-labeled riboprobes to whole mount (A) and sections (B–D) of reproductive (A, B) and nonreproductive (C, D) adult *C. loosanoffi* tissues. Reproductive adult tissues include many embryos and prerelease larvae (A, B). (A) The left half showing abundant hybridization with *PaxB* transcripts was hybridized with an antisense probe, while the right part, hybridized with a sense probe, exhibits no staining above background. The enlarged view in the inset demonstrates the absence of *PaxB* transcripts in an embryo/prerelease larva hybridized with antisense probe. (B) Reproductive adult tissue, hybridized with antisense probe, reveals *Six1/2* transcripts in embryos and prerelease larvae. The inset shows an enlarged view of an embryo/prerelease larva with transcripts evident in the epithelial layer (el) but not the inner cell mass (icm). (C, D) Hybridization with antisense probes of *PaxB* (C) or *Six1/2* (D) to nonreproductive adult tissues. The inset in C shows a lower magnification of the sponge body to highlight the absence of *PaxB* transcripts in the external pinacoderm (ep) and adjacent choanosome (c). Black and red arrows (B–D) point to pinacodermal cells that line the atrial opening of the oscular chimney (ao) and to some of the abundant spicules, respectively.

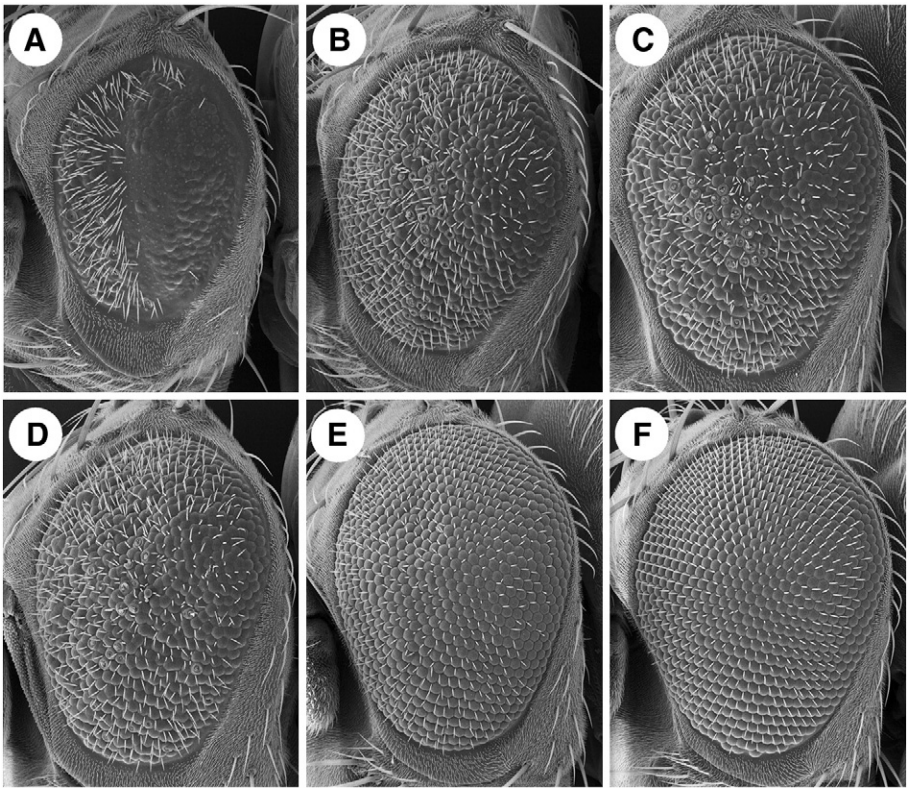


Fig. 9. Rescue of the spa^{pol} eye phenotype by expression of Cl-PaxB under control of the *spa* enhancer. Scanning electron micrographs of left eyes (anterior to the left) of females with the following genotypes: (A) *y w*; *UAS-Cl-PaxB* (line CR6)/+; spa^{pol}, (B–D) *y w spa-Gal4*; *UAS-Cl-PaxB* (lines CR3, CR5, and CR6)/+; spa^{pol}, (E) *y w spa-Gal4*; *UAS-D-Pax2*/+; spa^{pol}, and (F) *y w*; *UAS-Cl-PaxB* (line CR6)/+.

these Pax6 functions by expressing *UAS-Cl-PaxB* under the control of a *dpp-Gal4* driver in the leg, antenna, and wing disc. In control flies, expression of *UAS-ey* (Halder et al., 1995) under the same control generated many ectopic eye structures on legs, wings, and antennae. Ectopic expression of jellyfish PaxB led to substantially smaller eye structures or single ommatidia on legs (Kozmik et al., 2003). By contrast, none of the *UAS-Cl-PaxB* transgenic lines was able to induce ectopic eye morphogenesis under the same conditions. Neither did ectopic expression of *UAS-Cl-PaxB* in imaginal discs lead to deformed appendages, as observed with *UAS-ey* (Punzo et al., 2001) or *UAS-Tc-PaxB* transgenes (Kozmik et al., 2003; data not shown), an effect we attribute to developmental pathway interference (Jiao et al., 2001). To potentially increase the binding specificity of Cl-PaxB for enhancers of Ey target genes, the three amino acids Q, R, and H at positions 42, 44, and 47 of the Cl-PaxB paired domain, which are conserved in the PaxB/2/5/8 subfamily, were replaced by the Pax6-specific amino acids I, Q, and N (Czerny and Busslinger, 1995), a modification that strongly enhanced in Tc-PaxB its efficiency to induce ectopic eyes (Kozmik et al., 2003). In Cl-PaxB, however, none of eight independent *UAS-Cl-PaxB(IQN)* transgenic lines produced offspring with ectopic eyes or appendage defects when combined with the *dpp-Gal4* driver.

Discussion

Gene networks have been highly conserved in metazoans, as predicted by the gene network hypothesis (Frigerio et al., 1986; Noll, 1993). A striking example is the conservation of the Pax–Six–Eya–Dac network that determines eye development (for a review see Treisman, 1999), but which is also instrumental in myogenesis, nephrogenesis, and in the development of other organs (Kawakami et al., 2000). Here we investigated the origin of two gene families that participate in this network by examining their presence in the basal phylum of sponges. We have shown that the demosponge *C. loosanoffi* has only one member each of these gene families, *PaxB* and *Six1/2*, which thus are representatives of the founders of these gene families. The *Six3/6* and *Six4/5* subfamilies and the 'archetypal' Six gene cluster (Boucher et al., 2000), therefore, evolved in metazoans only after the poriferan lineage split but were both already present when bilaterians diverged from cnidarians (Stierwald et al., 2004). The Pax gene subfamilies, on the other hand, have emerged only by the time of the protostome–deuterostome split, as predicted (Noll, 1993).

In addition, we have shown that, like PaxB of the cubomedusan jellyfish *Tripedalia* (Kozmik et al., 2003), *Chalinula* PaxB protein is able to substitute for Pax2 functions in *Drosophila* eye development but, unlike the cnidarian PaxB, cannot perform Pax6 functions in *Drosophila*. As *Tripedalia* has complex eyes with lenses, this might reflect that the Pax–Six–Eya–Dac network has evolved in medusozoans, but not in sponges. Indeed, the temporal expression patterns of *PaxB* and *Six1/2* in developing embryonic and larval tissue and reaggregating adult cells of *Chalinula* do not suggest the existence of an exclusive primordial Pax–Six network in the ancestor of sponges and eumetazoans. However, our results do not exclude that such a network could exist given overlapping expression patterns in adult sponges. Furthermore, it is possible that a network could play a role in a small fraction of cells of developing sponges, such as in the photoresponsive cells of parenchymella larvae (Leys and Degnan, 2001; Maldonado et al., 2003). If the generation of these photoresponsive larval cells depends on a Pax–Six network, the hypothesis that this network is a precursor to metazoan sensory systems (Maldonado et al., 2003) and/or gave rise to eyes of Cnidaria or Chordata would receive considerable support. The *Chalinula* larvae used in this study are small, fragile, and difficult to collect, which made it impossible to determine whether *PaxB* and *Six1/2* are expressed at the site of the future or existing pigmented ring at the posterior pole where the photoresponsive ciliated epithelial cells are located (Maldonado et al., 2003). Should however a Pax–Six network exist in sponges, it seems more likely to be

active in pinacodermal cells and their associated choanosome that may play a role in the contractile activity of the adult rather than in photoresponsive larval cells.

An alternative possibility, consistent with the absence of a temporally correlated expression of *PaxB* and *Six1/2* in *Chalinula*, is that the Pax–Six gene network evolved in metazoans only after their divergence from sponges. An attractive hypothesis is that homologous gene networks evolved and diversified in parallel with the diversification of the gene families of a particular network, here those of Pax and Six genes. For example, *Pax2* and *Pax6* evolved by duplication of *PaxB*, as discussed below, both taking over different tasks in eye development of vertebrates and flies, illustrating how gene duplication and functional diversification accelerated the diversification of eyes (Kozmik, 2005). Such a hypothesis is consistent with the many homologous Pax–Six gene networks, in which the replacement of a member of one subfamily by a member of a different subfamily leads to a different integrated function of the network (Noll, 1993). Among these diverse functions of homologous Pax–Six gene networks are crucial roles in the development of the eye, inner ear, skeletal muscles, kidney, limb, and brain (Kawakami et al., 2000).

Origin and evolution of Pax genes

Based on their paired domains, modern Pax genes can be divided into four or five subfamilies, *PaxB/2/5/8*, *PaxD/3/7*, *Pax1/9*, *Pax6*, and *Poxn*, depending on whether or not *Poxn*-like genes are included in the *PaxB/2/5/8* subfamily (Noll, 1993). Pax genes have been found in metazoans but not in plants, lower eukaryotes, or bacteria, as predicted (Burri et al., 1989). While in anthozoans clearly more than one subfamily of Pax genes exists (Miller et al., 2000), medusozoans may include only a single type, the PaxB subfamily (Sun et al., 1997; Kozmik et al., 2003). This raises the question as to the origin of Pax genes, i.e., whether more primitive metazoans, such as Porifera or Placozoa, possess only a single subfamily of Pax genes. While a *PaxB* gene was found in the only characterized placozoan species *T. adhaerens* (Hadrys et al., 2005) and in the demosponges *E. fluviatilis* and *Reniera* sp. (Hoshiyama et al., 1998; Larroux et al., 2006), it remained unclear whether additional subfamilies of Pax genes were present in these animals. We have shown here that only one Pax gene is present in the genome of the demosponge *C. loosanoffi* and that it belongs to the *PaxB* subfamily.

In an attempt to trace the evolution of Pax genes back to their origin, we have analyzed 136 paired domain sequences, covering a wide range of phyla, i.e., parazoans and eumetazoans, diploblasts and triploblasts, as well as protostomes and deuterostomes (Fig. S2). To increase the coverage of lower phyla, we have annotated a number of Pax genes from the genome sequencing projects of the placozoan *T. adhaerens*, the planarian *S. mediterranea*, the annelid *C. capitata*, and the hemichordate *S. kowalevskii*. A dendrogram including 133 of these paired domains (Fig. 4) is consistent with the grouping of Pax genes into five subfamilies among which the *Poxn/PaxA/PaxC* clade is assumed as separate subfamily, in addition to the previously established subfamilies of *Pax1/9*, *PaxD/3/7*, *PaxB/2/5/8*, and *Pax4/6* (Noll, 1993; Sun et al., 1997; Galliot et al., 1999; Miller et al., 2000). Eight paired domains could not be

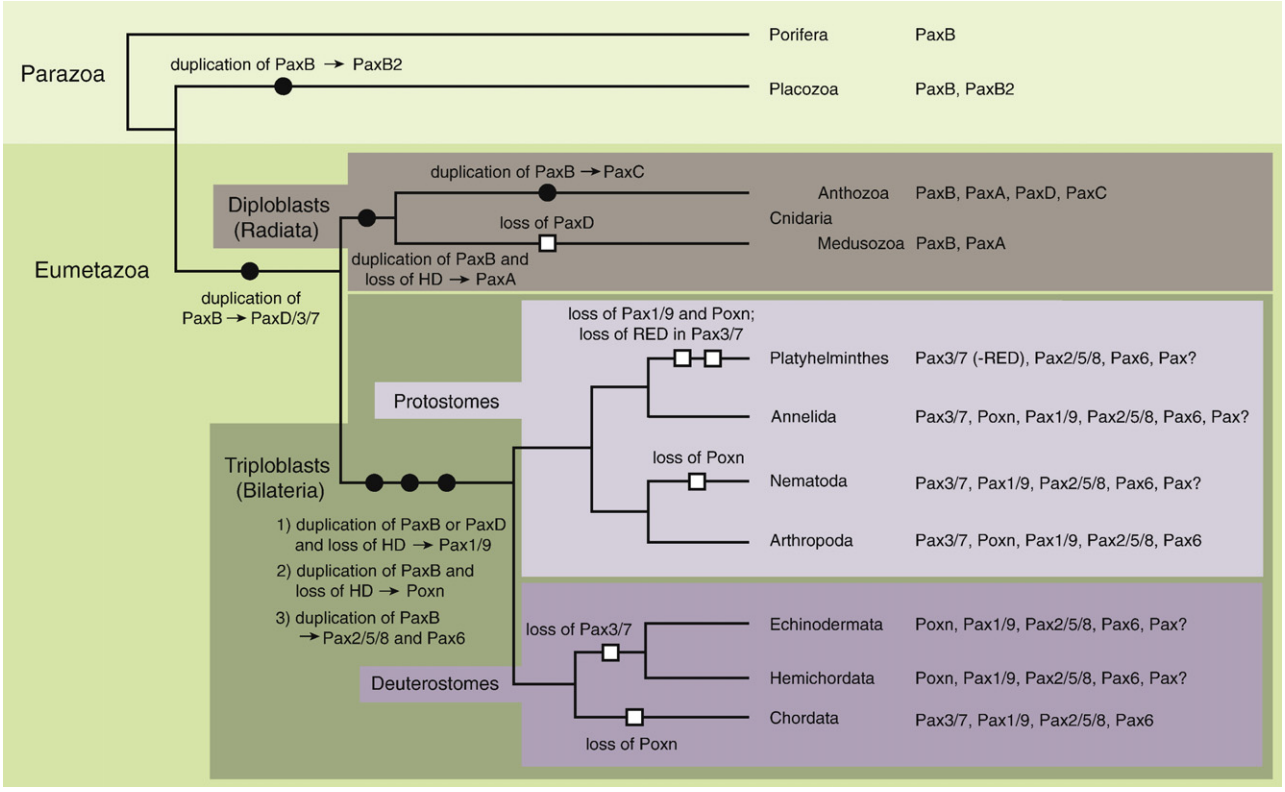


Fig. 10. Evolution of Pax genes in metazoans. The scheme was derived from the 136 paired domains shown in Fig. S2. Pax genes evolved from a single PaxB-type ancestral Ur-Pax gene whose modern descendant, PaxB, has been maintained in Porifera. Subsequent duplications (filled circles) of Pax genes before the radiation of Bilateria generated the modern Pax gene subfamilies *Pax3/7*, *Pax1/9*, *Poxn*, *Pax2/5/8*, and *Pax6*. In several lines of Diploblasts and Triploblasts Pax genes were lost (open squares), resulting in their spectrum of subfamilies indicated behind the phyla at the right. The phylum of Cnidaria is characterized by different subfamily spectra in the class of Anthozoa and the remaining classes of Medusozoa. Note that the four duplications giving rise to all Pax subfamilies have occurred by the time of the Cambrian explosion, about 500 Mya, whereas most losses of subfamilies occurred later. Duplication of Pax genes within subfamilies continued, as exemplified by the members of the *Drosophila prd/gsb/gsbn* subfamily or the vertebrate *Pax2/5/8* and *Pax6* subfamilies. Although paired domains that formed a single cluster in Fig. 4 but could not be assigned unambiguously to a particular subfamily, their proteins, indicated as 'Pax?', do not belong to the Pax3/7 or Pax6 subfamily because they do not include a homeodomain. The unassigned Platyhelminthean proteins do not belong to the Poxn or the Pax1/9 subfamily since, unlike Pax1/9 that have no introns in their paired domains, they are interrupted by an intron in α3 but not by one close to α4, which is diagnostic for Poxn. Not included in the scheme are Pax genes that have lost one of the two subdomains of the paired domain, PAI or RED (Hobert and Ruvkun, 1999), except in the newly discovered case where the RED domain was lost in the *Pax3/7* subfamily of Platyhelminthes.

assigned to any of these subfamilies and formed a separate group (Fig. 4). However, an assignment of these paired domains to a new subfamily is presently not sufficiently supported.

In a more perspicuous approach to assess to which subfamily a certain paired domain belongs, we have examined which amino acid positions in paired domains are diagnostic for a paired domain subfamily. An extension of this analysis, previously used for the PaxB/2/5/8 and Pax6 subfamilies (Kozmik et al., 2003), is shown for all paired domain subfamilies in Fig. 5A. It has been derived from the 125 complete paired domains of Fig. S2 that have been assigned unambiguously to a specific subfamily in the phylogenetic tree of Fig. 4. According to these 41 sites diagnostic for paired domain subfamilies (Fig. 5A), the paired domain of Poxn is clearly closer to that of PaxB than those of PaxA and PaxC. Thus, Poxn differs from PaxB at only 9 but from PaxA and PaxC at 14 and 13 of these positions (Fig. 5B). In addition, with as few as 4 differences at the 41 diagnostic sites, the paired domains of the PaxB subfamily are closest to those of the Pax2/5/8 subfamily, while the paired domains of PaxD are most similar to those of the Pax3/7 subfamily (Fig. 5B), in agreement with the dendrogram shown in Fig. 4.

Based on these considerations, we have derived a pedigree of paired domain subfamilies that originates with the single Pax gene found in Porifera (Fig. 10). For this pedigree, we have adopted a classical phylogeny in which Porifera are basal to Placozoa (Philippe et al., 2009), although this has been questioned recently by the proposal that Placozoa are basal to all Metazoa (Dellaporta et al., 2006; Schierwater et al., 2009). Clearly, our analysis of Pax and Six gene families supports the classical pedigree with sponges at its root better because only one subfamily of each gene family is found in sponges, whereas Placozoa have two subfamilies of Six genes, *Six1/2* and *Six3/6* (Figs. 3B and 6). However, as we shall see, these alternative proposals of an urmetazoon, regardless of whether a sponge or placozoon, will not seriously affect the pedigree of Pax genes in Metazoa.

The demosponge *C. loosanoffi* has only a single *PaxB* gene, including an octapeptide and a homeodomain. While the dendrogram of paired domains indicates that *T. adhaerens* PaxB2 belongs to the Poxn subfamily (Fig. 4), analysis of its paired domain (Fig. S2) by use of diagnostic paired domain sites (Fig. 5A) shows that it is closer to the PaxB (11 differences) than the Poxn subfamily (15 differences). This view is supported by the fact that the placozoan PaxB2 includes a homeodomain, like PaxB but unlike Poxn. Thus, the placozoan *T. adhaerens* genome contains two *PaxB* genes (Hadrys et al., 2005; Fig. S2), both encoding a homeodomain and one of them an octapeptide, but no Pax genes of a subfamily different from *PaxB*. Therefore, the Ur-Pax gene that appeared in Porifera and Placozoa was a *PaxB* gene, including an octapeptide and a homeodomain, as previously proposed (Noll, 1993). This gene duplicated in the placozoan lineage where one of the duplicated genes lost its octapeptide. As the *PaxD/3/7* subfamily is found in triploblasts and anthozoan diploblasts (Figs. 4, S2, and 9), its precursor was generated by duplication of *PaxB* between the parazoan–eumetazoan and the diploblast–triploblast splits. This argument rests on the conclusion that it is improbable that the *PaxD* and *Pax3/7* subfamilies arose independently by convergent evolution, as they differ in a consistent manner at 16 diagnostic positions from the *PaxB* subfamily (Fig. 5A). Since at least some cnidarian Medusozoa have no Pax genes of the *PaxD* subfamily, *PaxD* was lost again in these cnidarians (Fig. 10). Another duplication of *PaxB* in which the homeodomain was lost produced the precursor of the modern *PaxA* gene as observed in cnidarians (Fig. 10). Since bilaterian *Poxn* is clearly closer to *PaxB* than *PaxA* (Fig. 5B) and also shares the octapeptide with *PaxB* (Fig. S2), which is absent from *PaxA*, we propose that both *PaxA* and *Poxn* were derived from *PaxB* by independent duplications rather than *Poxn* from *PaxA*, as previously proposed (Miller et al., 2000). Therefore, *PaxA* was generated in diploblasts before the cnidarian radiation (Fig. 10). By contrast, *PaxC* originated by a later duplication from *PaxB* in anthozoans. Thus, the origin of anthozoan *PaxD* preceded that of *PaxA*, which in turn occurred before that of *PaxC*, in agreement with the increasing number of deviations from *PaxB* at the diagnostic paired domain sites (Fig. 5B).

In the triploblast lineage, a similar argument can be made for the origin of the *Pax1/9* subfamily. Since these genes are found in protostomes and deuterostomes, a duplication of one of the ancestral Pax genes occurred before the protostome–deuterostome split in triploblasts to generate the precursor of the *Pax1/9* subfamily (Fig. 10). At this time, the ancestral *PaxB* and *PaxD* genes existed, either of which could have produced the *Pax1/9* subfamily by duplication and loss of the homeodomain in the precursor of *Pax1/9* (Fig. 10) because both show a similar number of deviations from the *Pax1/9* subfamily at the diagnostic sites of their paired domains (Fig. 5B). In the same time interval before the protostome–deuterostome split, *PaxB* rather than *PaxD* generated the ancestral *Pax6* gene because *Pax6* is much closer to *PaxB* (Fig. 5B). Since during this very period *Pax2/5/8* emerged from *PaxB*, which was losing part of its homeodomain, we propose that *Pax6* and *Pax2/5/8* were generated by the same duplication event of the ancestral *PaxB* (Fig. 10). Thus, one of the duplicated genes lost the octapeptide and evolved into *Pax6*, the other evolved into *Pax2/5/8*, losing the C-terminal half of its homeodomain in vertebrates and Platyhelminthes and retaining only the N-terminal fifth of its homeodomain in arthropods (Fig. 2B). Interestingly, the amino acid sequence following the conserved part of the homeodomain has been highly conserved in arthropods although with no recognizable similarity to homeodomains (Fig. 2B). A similar conservation of the sequence following the N-terminal half of the homeodomain was observed in chordates (Fig. 2B).

Remarkably, the genomes of the two species of the phylum Platyhelminthes that have been sequenced have Pax genes with a complete paired box that can be assigned unambiguously to only two subfamilies, *Pax2/5/8* and *Pax6*. The two Pax genes *Pax2b,c* of the Platyhelminthean species *S. mediterranea* that cannot be assigned to a particular subfamily have an intron in α3, unlike *Pax1/9* that have no introns interrupting their paired domain, but no intron close to α4 (Fig. 1C), which is diagnostic for *Poxn* (s. below). Hence, in this phylum *Pax1/9* and *Poxn* were lost, while the *Pax3/7* subfamily was retained but lost the RED domain (Fig. 10). Similarly, *Poxn* appears to have been lost in nematodes and chordates, whereas *Pax3/7* was lost in echinoderms and hemichordates before their divergence from chordates, as their unassigned Pax genes do not encode a homeodomain (Fig. 10). Pax genes that could not be assigned to a specific subfamily but grouped in a single clade ('Pax?' in Fig. 4) were found in Platyhelminthes, Annelida, Nematoda, Echinodermata, and Hemichordata (Fig. 10). It is conceivable that these are members of an additional subfamily, although at present, such a hypothesis is not sufficiently supported. If this was the case, this subfamily originated by an additional duplication of *PaxB* and concomitant loss of its homeodomain before the split of deuterostomes from protostomes with its subsequent loss in Chordata and Ecdysozoa.

Origin and conservation of intron positions in Pax genes

A discussion of the evolution of Pax genes would be incomplete without a consideration of the positions of introns interrupting their conserved domains. These may be regarded as fingerprints of the history of a gene family and permit an estimate of the time at which introns must have been present during evolution. Consequently, some of them might be diagnostic for Pax gene subfamilies. All Pax genes have an intron in the first codon, or only a few nucleotides upstream, of the paired domain. Since this position is always close to the N-terminus of the coding region, its precise location might have been under less stringent selection, provided the paired domain remained intact, and could have varied its position by a sliding during evolution (Stoltzfus et al., 1997; Rogozin et al., 2000). Another intron the position of which is highly conserved, from *PaxB* in sponges to *Pax2/5/8* in Platyhelminthes, arthropods, sea urchins, and chordates

(Fig. 1C), splits the third α-helix of the paired domain. This intron is also present in two Platyhelminthean Pax genes, *Sm Pax2b,c* (Fig. 1C), that could not be assigned to a particular subfamily but lack a homeodomain and hence are probably diverged members of the *Pax2/5/8* subfamily. The position of this intron is also conserved in *Pax6* of arthropods and the urochordate *Ciona intestinalis*, but is absent in *Pax6* of sea urchins and vertebrates, which instead have an intron shifted to a position between the second and third α-helix (Fig. 1C). Similarly, in *Caenorhabditis elegans* this intron that splits α3 is absent from both *Pax2* and *Pax6*. In *Pax2*, it is replaced by an intron located between α2 and α3 at the same position as in *Pax6* of sea urchins and vertebrates; and in *Pax6*, by an intron at the N-terminus of α3, a position shifted with regard to an intron of *Drosophila Poxn* by one nucleotide (Fig. 1C). The positions of all these introns may be explained by rare events of intron sliding over very short distances during evolution (Stoltzfus et al., 1997; Rogozin et al., 2000) as well as by nearby hot spots to which introns may slide (Fig. 1C).

It is striking that the position of the first intron interrupting α6 of the paired domain is conserved in *Pax6* while that of the second intron in α6 is conserved and restricted to *Pax2* (Fig. 1C). This suggests that these introns were present in the ancestral *Pax6* and *Pax2* gene, respectively. As we propose that these genes originated in a duplication event of their ancestral *PaxB* gene, the simplest model would predict that the positions of introns were also fixed during this event, which would imply that the ancestral *PaxB* gene had an intron at only one of these two positions. Most *Pax3/7* genes have lost this intron, but it was retained in *Pax3/7* of vertebrates, although its position was shifted by one nucleotide as compared to its position in *Pax6* (Fig. 1C). The intron found in α5 of the paired domains of *Drosophila gsbn* and *C. elegans Pax2* may again have resulted by intron sliding to a hot spot from the respective positions in α6 of *Pax3/7* and *Pax2/5/8*.

Similar arguments can be made for the presence at earlier times in evolution of the two introns that split α1 of the homeodomain in the precursor of *Pax6* genes and α3 of the homeodomain in the precursors of *PaxB/C/D*, *Pax3/7*, and *Pax6* (Fig. 1C). In general, many introns have been lost during evolution that split the paired domain and homeodomain of Pax genes. Thus, all introns interrupting the paired domain of the *Pax1/9* gene subfamily were lost. By contrast, evidence supporting the gain of introns by Pax genes is scarce. An example, however, may be the intron interrupting α1 of the *Pax6* homeodomain (Fig. 1C), which appears to be present in *Pax6* genes of most, if not all, bilaterian phyla and hence may have been acquired during the duplication of *PaxB* giving rise to the ancestral *Pax6* and *Pax2/5/8* genes (Fig. 10). Another example of an intron gain may be the intron interrupting the linker between the PAI and RED subdomains in *Pax2* of *C. elegans* (Fig. 1C).

An interesting intron position is conserved between *Pax3/7* of vertebrates, *Pax6* of *C. elegans*, and *Poxn* of all phyla examined, which is located between the coding regions of the PAI and RED subdomains of the paired domain (Fig. 1C). At this position, small triplet shifts of the intron–exon boundary are apparently tolerated or may be even advantageous, as long as they result in the insertion of only few amino acids in the linker between the PAI and RED domains (Fig. S2). Hence, it is possible that this intron was present in *PaxB* when its duplications generated the ancestral *PaxD/3/7*, *Poxn*, and *Pax6* genes, i.e., before the separation of triploblasts from diploblasts (Fig. 10). If the paired domain was generated by fusion of the two subdomains, PAI and RED, a homeodomain and a homeodomain-like fold (Xu et al., 1999), it is possible that this intron is a relic of this event and hence was present already in the first *PaxB* gene. Alternatively, this intron may have arisen independently at hot spots in various Pax gene subfamilies and phyla. In *Pax2* of *C. elegans*, this intron may again have shifted to a position closer to the third α-helix of the paired domain (Fig. 1C).

In summary, these considerations suggest that the first Pax gene in metazoans was of the *PaxB*-type, as previously proposed (Gröger et al., 2000; Kozmik et al., 2003), and included, in addition to its paired domain, an octapeptide and homeodomain (Noll, 1993; Balczarek et al., 1997). In addition, it included probably at least five introns, if the intron separating the PAI and RED subdomains was present, which were located at the N-terminus of the paired domain, in α3 and α6 of the paired domain, between the PAI and RED domain, and in α3 of the homeodomain. These introns were lost in some lineages, whereas other introns were probably added later (for a review, see Jeffares et al., 2006). It appears that after the Cambrian explosion of triploblasts intron loss of Pax genes was much more frequent than intron gain. Finally, new Pax genes not only originated by gene duplication but were also lost, such as *Pax3/7* in Echinodermata and Hemichordata or *Poxn* in Chordata. Moreover, no new subfamilies of Pax genes were generated after the Cambrian explosion, although duplication of Pax genes continued to occur.

Relation to earlier models of Pax gene evolution

While our model of the evolution of Pax genes is largely consistent with previous models (Noll, 1993; Breitling and Gerber, 2000; Galliot and Miller, 2000; Miller et al., 2000; Hadrys et al., 2005; Matus et al., 2007), it is much more specific and deviates in a number of salient features. In addition to the paired domain, the first Pax gene of Porifera or Placozoa encoded a homeodomain and an octapeptide and thus was not a *PaxA* but a *PaxB* gene. Rather *PaxA* has been derived from the ancestral *PaxB* gene by duplication and loss of the homeobox and octapeptide (Fig. 10). Moreover, in contrast to the proposal of others (Galliot and Miller, 2000; Miller et al., 2000), we have argued that *Poxn* was derived by duplication from *PaxB* rather than *PaxA*. Therefore, contrary to a recent report (Hoshiyama et al., 2007), the diversification of the Pax gene family began after the divergence of Eumetazoa from Parazoa (Fig. 10).

The *PaxD* subfamily of Pax genes originated by duplication of the ancestral *PaxB* gene, including its octapeptide and homeodomain, before the diploblast–triploblast split and gave rise to the modern *Pax3/7* subfamily (Miller et al., 2000; Fig. 10). It lost its octapeptide in the anthozoan lineage but retained it in many modern *Pax3/7* genes (Hadrys et al., 2005) where it was discovered (Burri et al., 1989). Surprisingly, this subfamily of Pax genes was lost at least twice, first in Medusozoa and again in Echinodermata and Hemichordata (Fig. 10). It appears that loss of Pax gene subfamilies occurred repeatedly during evolution. Thus, *Poxn* and *Pax1/9* were lost in Platyhelminthes while *Poxn* was lost in nematodes and chordates (Fig. 10).

The remaining Pax gene subfamilies originated by gene duplication before the protostome–deuterostome split, *Pax1/9* from the ancestral *PaxB* or *PaxD* gene, where it lost the homeobox (Fig. 10), and *Pax2* and *Pax6* from the ancestral *PaxB* gene, during which *Pax2* lost most of the homeobox, and *Pax6*, the octapeptide (Matus et al., 2007; Fig. 10).

The enigma of diagnostic paired domain sites

From the expanded list of paired domains (Fig. S2), a table was derived for those sites that are diagnostic for the various paired domain subfamilies (Fig. 5A). Astonishingly, most of the 41 amino acids at these positions do not contact the DNA of an optimal Pax6-binding site, and of those that do, about half are located in the region that links the two paired subdomains (residues 61–76) and interact with the minor groove while only N47 interacts with a base in the major groove (Xu et al., 1999; Fig. 5C). Moreover, of the three amino acids, 42, 44, and 47, known to distinguish between DNA binding sites (Czerny and Busslinger, 1995), the first two do not contact the DNA either (Xu et al., 1999). Conversely, of the 38 amino acids that contact the DNA binding site (Xu et al., 1999), only 9 are at diagnostic sites (Fig. 5C). Indeed, a plot of the number of amino acid changes at diagnostic sites versus their position within the paired domain (Fig. 5C) shows that 6 of the 9 contacts by diagnostic amino acids occur at minimal frequencies. As the diagnostic sites probably

influence the selection of DNA target sites, they may affect the precise structure of the paired domain and thereby the binding of other amino acids to the DNA binding site and/or determine directly or indirectly the interaction of the paired domains with other transcription factors, such as Ets proteins (Fitzsimmons et al., 1996), thus influencing the selection of DNA target sites, regardless of whether these act as enhancers or silencers (Eberhard et al., 2000). Some of the diagnostic site changes may also compensate for each other. For example, R64 interacts with D20 (Xu et al., 1999), which is reversed in the paired domain of PaxD but not Pax3/7 (Fig. 5A). Considering the paired domain DNA interaction model derived from crystallographic studies (Xu et al., 1999), we note that the diagnostic sites generally cluster in regions oriented away from the DNA, as expected if the sites would interact with other transcription factors (Garvie et al., 2001). In agreement with this view, paired domain proteins of the same subfamily can largely substitute for each other's functions *in vivo* and hence recognize the same target sites (Li and Noll, 1994; Xue and Noll, 1996; Bouchard et al., 2000). To test the significance of the diagnostic paired domain sites, it will be important to determine the target sites of different subfamilies of paired domains as well as the amino acids within paired domains that mediate interactions with other transcriptional regulators important for target site selection.

Acknowledgments

This work has been supported by a grant from the Jeffress Memorial Trust (to A.H.) and the Swiss National Science Foundation grant 3100-56817 (to M.N.) and by the Kanton Zürich, Switzerland.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2010.03.010.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.
Adell, T., Grebenjuk, V.A., Wiens, M., Müller, W.E.G., 2003. Isolation and characterization of two T-box genes from sponges, the phylogenetically oldest metazoan taxon. Dev. Genes Evol. 213, 421–434.
Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. 55, 539–552.
Balczarek, K.A., Lai, Z.-C., Kumar, S., 1997. Evolution and functional diversification of the paired box (*Pax*) DNA-binding domains. Mol. Biol. Evol. 14, 829–842.
Bebenek, I.G., Gates, R.D., Morris, J., Hartenstein, V., Jacobs, D.K., 2004. *sine oculis* in basal Metazoa. Dev. Genes Evol. 214, 342–351.
Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., Noll, M., 1986. Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*. Cell 47, 1033–1040.
Bopp, D., Jamet, E., Baumgartner, S., Burri, M., Noll, M., 1989. Isolation of two tissue-specific *Drosophila* paired box genes, *Pox meso* and *Pox neuro*. EMBO J. 8, 3447–3457.
Bouchard, M., Pfeffer, P., Busslinger, M., 2000. Functional equivalence of the transcription factors Pax2 and Pax5 in mouse development. Development 127, 3703–3713.
Boucher, C.A., Winchester, C.L., Hamilton, G.M., Winter, A.D., Johnson, K.J., Bailey, M.E.S., 2000. Structure, mapping and expression of the human gene encoding the homeodomain protein, SIX2. Gene 247, 145–151.
Brand, A.H., Perrimon, N., 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401–415.
Breitling, R., Gerber, J.-K., 2000. Origin of the paired domain. Dev. Genes Evol. 210, 644–650.
Bruun, J.-A., Thomassen, E.I.S., Kristiansen, K., Tylden, G., Holm, T., Mikkola, I., Bjørkøy, G., Johansen, T., 2005. The third helix of the homeodomain of paired class homeodomain proteins acts as a recognition helix both for DNA and protein interactions. Nucleic Acids Res. 33, 2661–2675.
Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G., Noll, M., 1989. Conservation of the paired domain in metazoans and its structure in three isolated human genes. EMBO J. 8, 1183–1190.
Chen, R., Amoui, M., Zhang, Z., Mardon, G., 1997. Dachshund and Eyes absent proteins form a complex and function synergistically to induce ectopic eye development in *Drosophila*. Cell 91, 893–903.
Cheyette, B.N.R., Green, P.J., Martin, K., Garren, H., Hartenstein, V., Zipursky, S.L., 1994. The *Drosophila sine oculis* locus encodes a homeodomain-containing protein required for the development of the entire visual system. Neuron 12, 977–996.
Chi, N., Epstein, J.A., 2002. Getting your Pax straight: Pax proteins in development and disease. Trends Genet. 18, 41–47.
Cullings, K.W., 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1, 233–240.
Custodio, M.R., Prokic, I., Steffen, R., Koziol, C., Borojevic, R., Brümmer, F., Nickel, M., Müller, W.E.G., 1998. Primmorphs generated from dissociated cells of the sponge *Suberites domuncula*: a model system for studies of cell proliferation and cell death. Mech. Ageing Dev. 105, 45–59.
Czerny, T., Busslinger, M., 1995. DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). Mol. Cell. Biol. 15, 2858–2871.
Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W.J., Busslinger, M., 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.
Dellaporta, S.L., Xu, A., Sagasser, S., Jakob, W., Moreno, M.A., Buss, L.W., Schierwater, B., 2006. Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum. Proc. Natl. Acad. Sci. U. S. A. 103, 8751–8756.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.
Eberhard, D., Jimenez, G., Heavey, B., Busslinger, M., 2000. Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. EMBO J. 19, 2292–2303.
Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S.-M., Wasylyk, B., Hagman, J., 1996. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. Genes Dev. 10, 2198–2211.
Frigerio, G., Burri, M., Bopp, D., Baumgartner, S., Noll, M., 1986. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. Cell 47, 735–746.
Fu, W., Noll, M., 1997. The *Pax2* homolog *sparkling* is required for development of cone and pigment cells in the *Drosophila* eye. Genes Dev. 11, 2066–2078.
Funayama, N., Nakatsukasa, M., Hayashi, T., Agata, K., 2005. Isolation of the choanocyte in the fresh water sponge, *Ephydatia fluviatilis* and its lineage marker, *Ef annexin*. Dev. Growth Differ. 47, 243–253.
Gallardo, M.E., Lopez-Rios, J., Fernaud-Espinosa, I., Granadino, B., Sanz, R., Ramos, C., Ayuso, C., Seller, M.J., Brunner, H.G., Bovolenta, P., et al., 1999. Genomic cloning and characterization of the human homeobox gene *SIX6* reveals a cluster of *SIX* genes in chromosome 14 and associates *SIX6* hemizygosity with bilateral anophthalmia and pituitary anomalies. Genomics 61, 82–91.
Galliot, B., Miller, D., 2000. Origin of anterior patterning. Trends Genet. 16, 1–5.
Galliot, B., de Vargas, C., Miller, D., 1999. Evolution of homeobox genes: Q_{50} paired-like genes founded the paired class. Dev. Genes Evol. 209, 186–197.
Garvie, C.W., Hagman, J., Wolberger, C., 2001. Structural studies of Ets-1/Pax5 complex formation on DNA. Mol. Cell 8, 1267–1276.
Gehring, W.J., Ikeo, K., 1999. *Pax6* mastering eye morphogenesis and eye evolution. Trends Genet. 15, 371–377.
Gröger, H., Callaerts, P., Gehring, W.J., Schmid, V., 2000. Characterization and expression analysis of an ancestor-type *Pax* gene in the hydrozoan jellyfish *Podocoryne carnea*. Mech. Dev. 94, 157–169.
Gruss, P., Walther, C., 1992. Pax in development. Cell 69, 719–722.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Hadrys, T., DeSalle, R., Sagasser, S., Fischer, N., Schierwater, B., 2005. The Trichoplax *PaxB* gene: a putative proto-*PaxA*/*B*/*C* gene predating the origin of nerve and sensory cells. Mol. Biol. Evol. 22, 1569–1578.
Halder, G., Callaerts, P., Gehring, W.J., 1995. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. Science 267, 1788–1792.
Halder, G., Callaerts, P., Flister, S., Walldorf, U., Kloter, U., Gehring, W.J., 1998. Eyeless initiates the expression of both *sine oculis* and *eyes absent* during *Drosophila* compound eye development. Development 125, 2181–2191.
Hanson, I., van Heyningen, V., 1995. Pax6: more than meets the eye. Trends Genet. 11, 268–272.
Hartman, W.D., 1958. Natural history of the marine sponges of Southern New England. Bull. Peabody Mus. Nat. Hist. 12, 1–155.
Heanue, T.A., Reshef, R., Davis, R.J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A.B., Tabin, C.J., 1999. Synergistic regulation of vertebrate muscle development by *Dach2*, *Eya2*, and *Six1*, homologs of genes required for *Drosophila* eye formation. Genes Dev. 13, 3231–3243.
Hill, A., Tetrault, J., Hill, M., 2004. Isolation and expression analysis of a poriferan *Antp*-class *Bar*-/*Bsh*-like homeobox gene. Dev. Genes Evol. 214, 515–523.
Hobert, O., Ruvkun, G., 1999. *Pax* genes in *Caenorhabditis elegans*. Trends Genet. 15, 214–216.
Hordijk, W., Gascuel, O., 2005. Improving the efficiency of SPR moves in phylogenetic tree search methods based on maximum likelihood. Bioinformatics 22, 4338–4347.
Hoshiyama, D., Suga, H., Iwabe, N., Koyanagi, M., Nikoh, N., Kuma, K.-i., Matsuda, F., Honjo, T., 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.
Hoshiyama, D., Iwabe, N., Miyata, T., 2007. Evolution of the gene families forming the *Pax*/*Six* regulatory network: isolation of genes from primitive animals and molecular phylogenetic analyses. FEBS Lett. 581, 1639–1643.
Jeffares, D.C., Mourier, T., Penny, D., 2006. The biology of intron gain and loss. Trends Genet. 22, 16–22.
Jiao, R., Daube, M., Duan, H., Zou, Y., Frei, E., Noll, M., 2001. Headless flies generated by developmental pathway interference. Development 128, 3307–3319.
Kawakami, K., Sato, S., Ozaki, H., Ikeda, K., 2000. *Six* family genes—structure and function as transcription factors and their roles in development. BioEssays 22, 616–626.
King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., et al., 2008. The genome of the choanoflagellate *Monsiga brevicollis* and the origin of metazoans. Nature 451, 783–788.

Kozmik, Z., 2005. Pax genes in eye development and evolution. Curr. Opin. Gen. Dev. 15, 430–438.
Kozmik, Z., Daube, M., Frei, E., Norman, B., Kos, L., Dishaw, L.J., Noll, M., Piatigorsky, J., 2003. Role of Pax genes in eye evolution: a cnidarian *PaxB* gene uniting Pax2 and Pax6 functions. Dev. Cell 5, 773–785.
Krasko, A., Schröder, H.C., Batel, R., Grebenjuk, V.A., Steffen, R., Müller, I.M., Müller, W.E.G., 2002. Iron induces proliferation and morphogenesis in primmorphs from the marine sponge *Suberites domuncula*. DNA Cell Biol. 21, 67–80.
Kronhamn, J., Frei, E., Daube, M., Jiao, R., Shi, Y., Noll, M., 2002. Headless flies produced by mutations in the paralogous *Pax6* genes *eyeless* and *twin of eyeless*. Development 129, 1015–1026.
Kumar, J.P., Moses, K., 2001. EGF receptor and Notch signaling act upstream of Eyeless/Pax6 to control eye specification. Cell 104, 687–697.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Larroux, C., Fahey, B., Liubicich, D., Hinman, V.F., Gauthier, M., Gongora, M., Green, K., Wörheide, G., Leys, S.P., Degnan, B.M., 2006. Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. Evol. Dev. 8, 150–173.
Larroux, C., Luke, G.N., Koopman, P., Rokhsar, D.S., Shimeld, S.M., Degnan, B.M., 2008. Genesis and expansion of metazoan transcription factor gene classes. Mol. Biol. Evol. 25, 980–996.
Laughon, A., 1991. DNA binding specificity of homeodomains. Biochemistry 30, 11357–11367.
Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320.
Leith, A., 1979. Role of aggregation factor and cell type in sponge cell adhesion. Biol. Bull. 156, 212–223.
Leys, S.P., Degnan, B.M., 2001. Cytological basis of photoresponsive behavior in a sponge larva. Biol. Bull. 201, 323–338.
Li, X., Noll, M., 1994. Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different *cis*-regulatory regions. Nature 367, 83–87.
Maldonado, M., Durfort, M., McCarthy, D.A., Young, C.M., 2003. The cellular basis of photobehavior in the tufted parenchymella larva of demosponges. Mar. Biol. 143, 427–441.
Matus, D.Q., Pang, K., Daly, M., Martindale, M.Q., 2007. Expression of Pax gene family members in the anthozoan cnidarian *Nematostella vectensis*. Evol. Dev. 9, 25–38.
Miller, D.J., Hayward, D.C., Reece-Hoyes, J.S., Scholten, I., Catmull, J., Gehring, W.J., Callaerts, P., Larsen, J.E., Ball, E.E., 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. U. S. A. 97, 4475–4480.
Müller, W.E.G., Wiens, M., Batel, R., Steffen, R., Schröder, H.C., Borojevic, R., Custodio, M.R., 1999. Establishment of a primary cell culture from a sponge: primmorphs from *Suberites domuncula*. Mar. Ecol. Prog. Ser. 178, 205–219.
Nichols, S.A., Dirks, W., Pearse, J.S., King, N., 2006. Early evolution of animal cell signaling and adhesion genes. Proc. Natl. Acad. Sci. U.S.A. 103, 12451–12456.
Noll, M., 1993. Evolution and role of Pax genes. Curr. Opin. Genet. Dev. 3, 595–695.
Ochman, H., Gerber, A.S., Hartl, D.L., 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120, 621–623.
Ozaki, H., Nakamura, K., Funahashi, J.-i., Ikeda, K., Yamada, G., Tokano, H., Okamura, H.-o., Kitamura, K., Muto, S., Kotaki, H., et al., 2004. *Six1* controls patterning of the mouse otic vesicle. Development 131, 551–562.
Perovic, S., Schröder, H.C., Sudek, S., Grebenjuk, V.A., Batel, R., Stifanic, M., Müller, I.M., Müller, W.E.G., 2003. Expression of one sponge *Iroquois* homeobox gene in primmorphs from *Suberites domuncula* during canal formation. Evol. Dev. 5, 240–250.
Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault, N., Vacelet, J., Renard, E., Houliston, E., Quéinnec, E., et al., 2009. Phylogenomics revives traditional views on deep animal relationships. Curr. Biol. 19, 706–712.
Pignoni, F., Hu, B., Zavitz, K.H., Xiao, J., Garrity, P.A., Zipursky, S.L., 1997. The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development. Cell 91, 881–891.
Pomerantz, J.L., Sharp, P.A., 1994. Homeodomain determinants of major groove recognition. Biochemistry 33, 10851–10858.
Punzo, C., Kurata, S., Gehring, W.J., 2001. The *eyeless* homeodomain is dispensable for eye development in *Drosophila*. Genes Dev. 15, 1716–1723.
Quiring, R., Walldorf, U., Kloter, U., Gehring, W.J., 1994. Homology of the *eyeless* gene of *Drosophila* to the *Small eye* gene in mice and *Aniridia* in humans. Science 265, 785–789.
Robb, S.M.C., Ross, E., Sánchez Alvarado, A., 2008. SmedGD: the *Schmidtea mediterranea* genome database. Nucl. Acids Res. 36, D599–D606.
Rogozin, I.B., Lyons-Weiler, J., Koonin, E.V., 2000. Intron sliding in conserved gene families. Trends Genet. 16, 430–432.
Ruiz-Trillo, I., Inagaki, Y., Davis, L.A., Sperstad, S., Landfald, B., Roger, A.J., 2004. *Capsaspora owczarzaki* is an independent opisthokont lineage. Curr. Biol. 14, R946–R947.
Ruiz-Trillo, I., Burger, G., Holland, P.W.H., King, N., Lang, B.F., Roger, A.J., Gray, M.W., 2007. The origin of multicellularity: a multi-taxon genome initiative. Trends Genet. 23, 113–118.
Schierwater, B., Eitel, M., Jakob, W., Osigus, H.-J., Hadrys, H., Dellaporta, S.L., Kolokotronis, S.-O., DeSalle, R., 2009. Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoan" hypothesis. PLoS Biol. 7, 36–44.
Seo, H.-C., Curtiss, J., Mlodzik, M., Fjose, A., 1999. *Six* class homeobox genes in *Drosophila* belong to three distinct families and are involved in head development. Mech. Dev. 83, 127–139.
Shen, W., Mardon, G., 1997. Ectopic eye development in *Drosophila* induced by directed *dachshund* expression. Development 124, 45–52.
Staehling-Hampton, K., Jackson, P.D., Clark, M.J., Brand, A.H., Hoffmann, F.M., 1994. Specificity of bone morphogenetic protein-related factors: cell fate and gene expression changes in *Drosophila* embryos induced by *decapentaplegic* but not *60A*. Cell Growth Diff. 5, 585–593.
Stierwald, M., Yanze, N., Bamert, R.P., Kammermeier, L., Schmid, V., 2004. The *Sine oculis*/*Six* class family of homeobox genes in jellyfish with and without eyes: development and eye generation. Dev. Biol. 274, 70–81.
Stoltzfus, A., Logsdon, J.M., Palmer, J.D., Doolittle, W.F., 1997. Intron "sliding" and the diversity of intron positions. Proc. Natl. Acad. Sci. U. S. A. 94, 10739–10744.
Sun, H., Rodin, A., Zhou, Y., Dickinson, D.P., Harper, D.E., Hewett-Emmett, D., 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. U. S. A. 94, 5156–5161.
Sun, H., Dickinson, D.P., Costello, J., Li, W.-H., 2001. Isolation of *Cladonema Pax-B* genes and studies of the DNA-binding properties of cnidarian Pax paired domains. Mol. Biol. Evol. 18, 1905–1918.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Treisman, J.E., 1999. A conserved blueprint for the eye? BioEssays 21, 843–850.
Xu, H.E., Rould, M.A., Xu, W., Epstein, J.A., Maas, R.L., Pabo, C.O., 1999. Crystal structure of the human Pax6 paired domain–DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes Dev. 13, 1263–1275.
Xu, P.-X., Zheng, W., Huang, L., Maire, P., Laclef, C., Silvius, D., 2003. *Six1* is required for the early organogenesis of mammalian kidney. Development 130, 3085–3094.
Xue, L., Noll, M., 1996. The functional conservation of proteins in evolutionary alleles and the dominant role of enhancers in evolution. EMBO J. 15, 3722–3731.