GBE

Genome Biology and Evolution

SMBE

Evolution of the Vertebrate Pax4/6 Class of Genes with Focus on Its Novel Member, the *Pax10* Gene

Nathalie Feiner[1,2,3], Axel Meyer[1,2], and Shigehiro Kuraku[1,2,4,*]

[1]Department of Biology, University of Konstanz, Germany

[2]International Max-Planck Research School (IMPRS) for Organismal Biology, University of Konstanz, Germany

[3]Present address: Department of Zoology, University of Oxford, United Kingdom

[4]Present address: Genome Resource and Analysis Unit, RIKEN Center for Developmental Biology, Chuo-ku, Kobe, Hyogo, Japan

*Corresponding author: E-mail: shigehiro-kuraku@cdb.riken.jp.

Accepted: June 14, 2014

Data deposition: This project has been deposited at EMBL under the accessions HF567444–HF567455.

Abstract

The members of the *paired box* (*Pax*) family regulate key developmental pathways in many metazoans as tissue-specific transcription factors. Vertebrate genomes typically possess nine *Pax* genes (*Pax1–9*), which are derived from four proto-*Pax* genes in the vertebrate ancestor that were later expanded through the so-called two-round (2R) whole-genome duplication. A recent study proposed that *pax6a* genes of a subset of teleost fishes (namely, acanthopterygians) are remnants of a paralog generated in the 2R genome duplication, to be renamed *pax6.3*, and reported one more group of vertebrate *Pax* genes (*Pax6.2*), most closely related to the Pax4/6 class. We propose to designate this new member *Pax10* instead and reconstruct the evolutionary history of the Pax4/6/10 class with solid phylogenetic evidence. Our synteny analysis showed that *Pax4*, *-6*, and *-10* originated in the 2R genome duplications early in vertebrate evolution. The phylogenetic analyses of relationships between teleost *pax6a* and other *Pax4*, *-6*, and *-10* genes, however, do not support the proposed hypothesis of an ancient origin of the acanthopterygian *pax6a* genes in the 2R genome duplication. Instead, we confirmed the traditional scenario that the acanthopterygian *pax6a* is derived from the more recent teleost-specific genome duplication. Notably, *Pax6* is present in all vertebrates surveyed to date, whereas *Pax4* and *-10* were lost multiple times in independent vertebrate lineages, likely because of their restricted expression patterns: Among *Pax6*-positive domains, *Pax10* has retained expression in the adult retina alone, which we documented through in situ hybridization and quantitative reverse transcription polymerase chain reaction experiments on zebrafish, *Xenopus*, and anole lizard.

Key words: Pax6, Pax4, Pax10, gene loss, conserved synteny, gene duplication.

Introduction

The vertebrate gene repertoire was shaped by two rounds (2R) of whole-genome duplications (WGD) early in vertebrate evolution (Ohno 1970; Lundin 1993; Holland et al. 1994). These events initially generated four paralogs in vertebrates corresponding to a single invertebrate ortholog, and subsequent processes such as neofunctionalization, subfunctionalization, or complete loss of function modified this initial four-fold genetic abundance of evolutionary raw material. Genes that play crucial roles in development tend to be highly conserved and are therefore present in the genomes of diverse vertebrates. In contrast, genes that are less crucial for development and survival experience less selective pressure in the form of balancing selection, and are hence permitted to be differentially lost in vertebrate lineages (Lynch et al. 2001). Examples are the *Bmp16* gene, a sister gene of the highly conserved *Bmp2* and *-4* genes (Feiner et al. 2009), which has only been found in teleosts so far, and the *Pdx2* gene, a duplicate of the pancreatic key regulator *Pdx1*, which is only retained in cartilaginous fish and coelacanths (Mulley and Holland 2010).

Members of the *Pax* gene family encode transcription factors that play crucial regulatory roles in metazoan development (reviewed in Wehr and Gruss 1996). All vertebrate Pax proteins identified to date are characterized by the possession of a paired domain (Breitling and Gerber 2000; Underhill 2012). They are divided into four classes, namely Pax1/9, Pax3/7, Pax2/5/8, and Pax4/6, based on the completeness of a homeodomain and the presence of an octapeptide motif

(fig. 1*A*; Wehr and Gruss 1996; Chi and Epstein 2002). The last common ancestor of bilaterians, the so-called "Urbilateria," already possessed proto-orthologs of these four Pax classes, plus an additional class, the PaxA/Pox neuro class, restricted to invertebrates (Matus et al. 2007). Preceding the radiation of vertebrates, each of the four classes was quadruplicated by the 2R-WGD (Wada et al. 1998; Holland et al. 1999; Ogasawara et al. 1999; Manousaki et al. 2011).

The Pax4/6 class of genes consists of *Pax4* and *-6* genes as well as a recently identified gene named *Pax6.2* (Ravi et al. 2013). This novel gene has only been identified in a chondrichthyan (the elephant shark, *Callorhinchus milii*), several teleost fish, a reptile (the green anole, *Anolis carolinensis*), and an amphibian (the frog, *Xenopus tropicalis*; fig. 1*B*; Ravi et al. 2013). *Pax4*, the other close relative of *Pax6* (Manousaki et al. 2011), also shows a mosaic pattern of phylogenetic distribution, confined to mammals (Pilz et al. 1993; Tamura et al. 1994) and teleost fish (Hoshiyama et al. 2007; Manousaki et al. 2011), whereas the *Pax6* gene is identified in every vertebrate genome sequenced to date (fig. 1*B*). Ravi et al. (2013) also recently called the phylogeny of teleost fish *pax6* genes into question. Acanthopterygii is a group of teleost fish, and among the typical laboratory animals with the sequenced genomes, it includes medaka, Fugu, and stickleback, but not zebrafish. They concluded that acanthopterygian *pax6a* and *-6b* genes did not originate in the so-called "third-round" WGD in the teleost lineage (teleost-specific genome duplication, TSGD; Amores et al. 1998; Wittbrodt et al. 1998; Meyer and Van de Peer 2005). Instead, they proposed that acanthopterygian *pax6a* genes originated in a more ancient event, namely the 2R-WGD at the base of vertebrate evolution (fig. 2*A*; Ravi et al. 2013). Importantly, this would imply that the acanthopterygian *pax6a* genes are not orthologous to all other vertebrate *Pax6* genes. To demarcate acanthopterygian *pax6a* genes and emphasize their ancient origin, Ravi et al. (2013) therefore proposed to rename them *pax6.3*.

The *Pax6* gene is famous for its essential role as "master control gene" for eye development. Studies in the 1990s revealed the ability of an ectopically expressed mouse *Pax6* gene to induce ectopic eyes in *Drosophila* (Halder et al. 1995). Apart from this inductive role in eye development, the vertebrate *Pax6* gene is involved in the development of the central nervous system (CNS), including fore- and hindbrain, the neural tube, the pituitary, the nasal epithelium, and the endocrine part of the pancreas (Walther and Gruss 1991; St-Onge et al. 1997). In zebrafish, *pax6b*, but not *pax6a* ("*pax6.3*" in Ravi et al. 2013), is expressed in the endocrine pancreas (Delporte et al. 2008). The vertebrate paralog of *Pax6*, the *Pax4* gene, is necessary for the differentiation of insulin-producing β-cells in the endocrine part of the mammalian pancreas (Sosa-Pineda

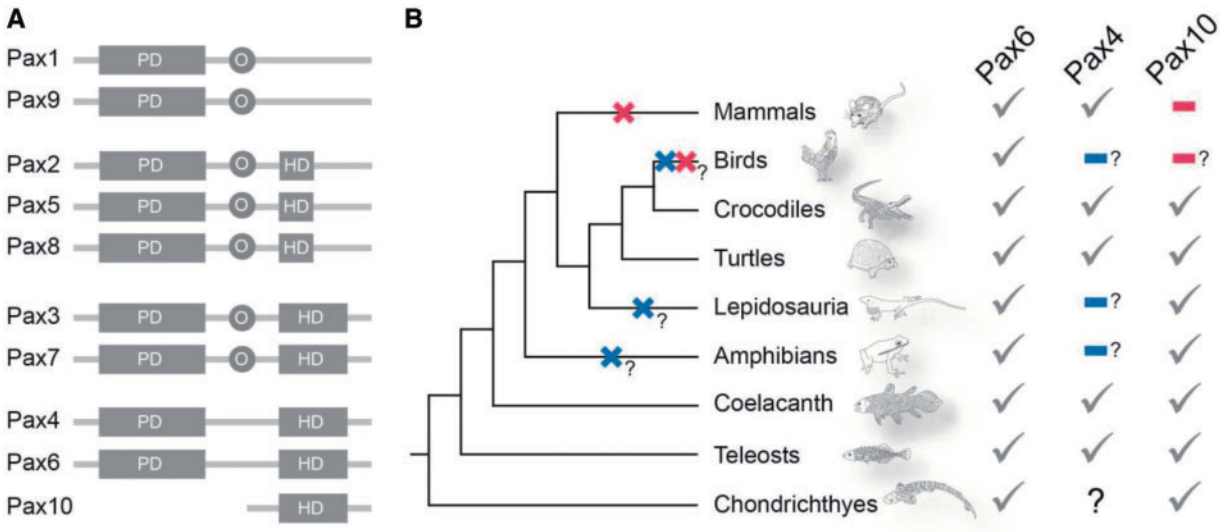

Fig. 1.—Domain structure of vertebrate Pax proteins and phylogenetic distribution of *Pax4*, *-6*, and *-10* genes across jawed vertebrates. (*A*) Presences of paired domains (PD), homeodomains (HD), and octapeptides (O) for all vertebrate Pax subtypes. No paired box has been identified in any of the *Pax10* genes, and thus, mature Pax10 proteins presumably lack a paired domain. (*B*) Phylogeny of major vertebrate taxa with indicated patterns of presence and presumed absence of *Pax4*, *-6*, and *-10* genes. The presence of these genes was investigated using exhaustive Blast searches in publicly available whole-genome sequences (see supplementary table S1, Supplementary Material online, for details). The chondrichthyan *Pax10* gene was reported by Ravi et al. (2013). Inferred secondary gene losses are indicated with red and blue crosses and mapped onto the generally accepted jawed vertebrate phylogeny. Question marks indicate uncertainties about the absence of genes because of insufficient sequence information of the respective taxa. The phylogenetic position of turtles is based on molecular phylogenetic studies (Zardoya and Meyer 1998; Rest et al. 2003; Iwabe et al. 2005; Chiari et al. 2012; Crawford et al. 2012; Wang et al. 2013).

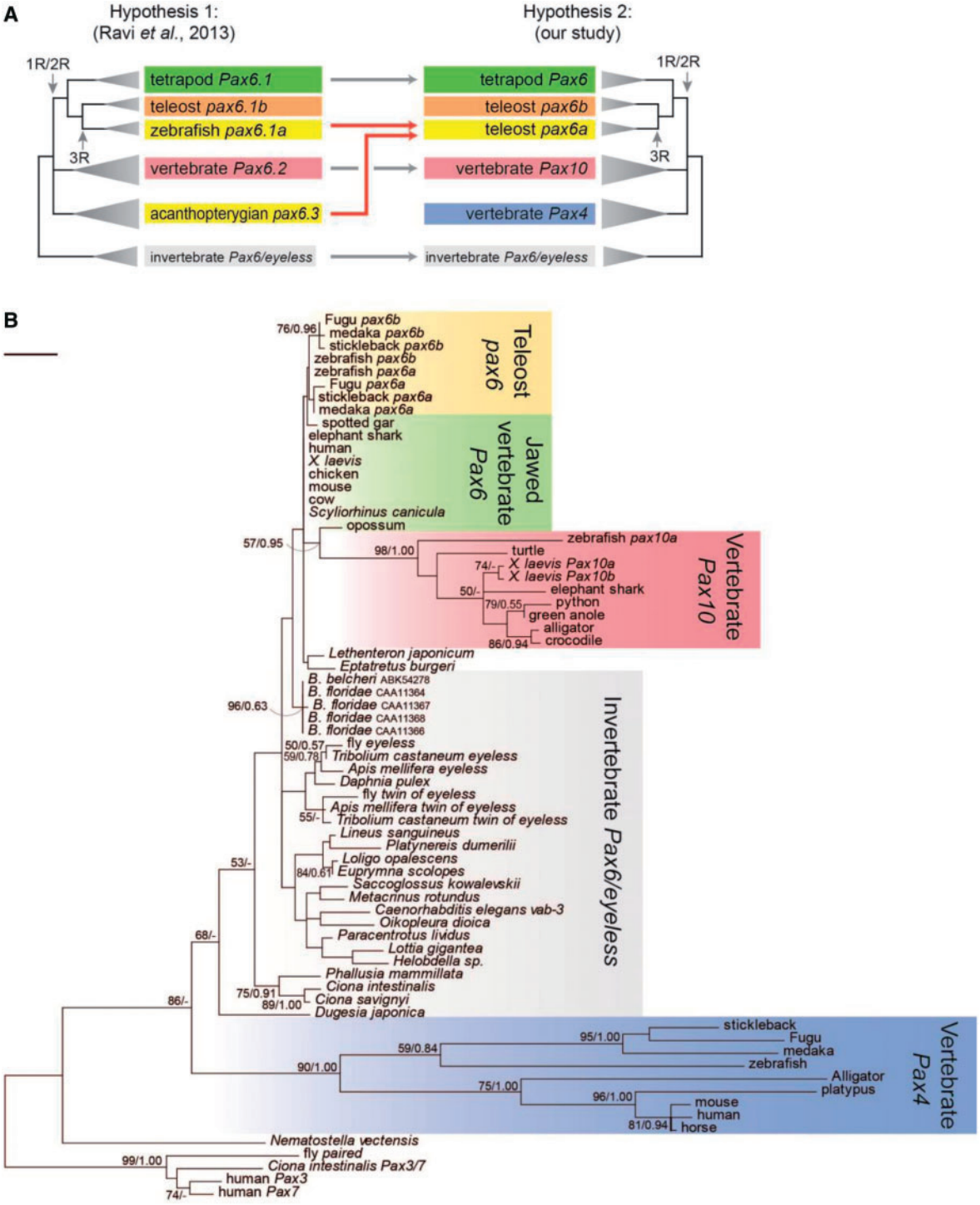

Fig. 2.—Phylogenetic relationships within the Pax4/6/10 class of genes. (*A*) Schematic presentations of two scenarios of the evolution of the Pax4/6/10 class of genes. Hypothesis 1, proposed by Ravi et al. (2013), assumes an ancient origin of one group of acanthopterygian *pax6* genes, namely *pax6.3*. In addition, this hypothesis does not take the group of *Pax4* genes into account. Hypothesis 2 is proposed based on our phylogenetic analysis, and it postulates the origin of both groups of teleost *pax6* genes, namely *pax6a* and *-6b*, in the TSGD. The gene nomenclature in Hypothesis 1 is adopted from Ravi et al. (2013). (*B*) ML tree showing phylogenetic relationships among vertebrate genes belonging to the Pax4/6/10 class, with *Pax3* and *-7* genes as outgroup. Exact names of the genes are included only when they experienced additional lineage-specific duplications. Support values are shown for each node in order, bootstrap probabilities in the ML and Bayesian posterior probabilities. Only bootstrap probabilities above 50 are shown. The inference is based on 99 amino acid residues, assuming the JTT+Γ_4 model (shape parameter of the gamma distribution $\alpha = 0.57$). The scale bar on the upper left indicates 0.2 substitutions per site.

et al. 1997), but not of the zebrafish pancreas (Djiotsa et al. 2012). This indicates evolutionarily conserved roles of *Pax6* and lineage-specific differences in the roles of *Pax4* during pancreas development among vertebrates. The *Pax4* genes of mammals and teleost fishes are not implicated in the development of eyes (Rath et al. 2009; Manousaki et al. 2011). However, mammalian *Pax4* genes are expressed in photoreceptors in the outer nuclear layer of the adult mammalian retina in a diurnal rhythm (Rath et al. 2009). The *Pax6* gene is expressed in other neuronal layers of the mature retina, namely the inner nuclear layer, the ganglion cell layer, and in several species including a shark, zebrafish, chicken, and mouse, and also in the horizontal cell layer (Belecky-Adams et al. 1997; Macdonald and Wilson 1997; Wullimann and Rink 2001; Ferreiro-Galve et al. 2011). The teleost *pax4* gene has an expression domain that has not been attributed to any other *Pax* gene, namely the stomodeum which corresponds to the developing lip (Manousaki et al. 2011).

Information on the developmental roles of the novel *Pax6* relative, *Pax6.2*, is limited because it was found only very recently (Ravi et al. 2013) and also because it is absent from the genomes of traditional model species. This recent study revealed the expression of the elephant shark *Pax6.2* gene in its adult eye, shared with *Pax4* and *-6* genes, and the adult kidney in which none of the other *Pax6* relatives have been shown to be expressed (Ravi et al. 2013). Its zebrafish ortholog is expressed in the head region during early developmental stages, and becomes restricted to the inner nuclear layer of the retina during late developmental stages (Ravi et al. 2013).

In this study, we assessed the molecular phylogeny of the recently identified relative of the *Pax6* gene (*Pax6.2* in Ravi et al. 2013) and propose to call it *Pax10* instead. By conducting rigorous molecular phylogenetic analyses and considering conserved synteny, we demonstrate that *Pax4*, *-6*, and *-10* originated in the 2R-WGD, and that the two *pax6* genes of teleost fishes (*pax6a* and *pax6b*) were duplicated in the TSGD. Our gene expression analyses in the zebrafish, a frog, and the green anole lizard revealed the *Pax10* expression in the adult retina and brain and suggested that *Pax10* presumably plays no role during early development. Our reanalysis provides a synthetic understanding of the evolution of the vertebrate Pax4/6/10 gene repertoire and the functional partitioning of these genes.

Materials and Methods

In Silico Identification of Novel *Pax* Genes

Using the green anole *Pax10* and human *Pax4* nucleotide sequences as queries, we performed BLASTn searches in the NCBI dbEST (National Center for Biotechnology Information) as well as in the nr/nt database and in nucleotide genomic sequences of species included in the Ensembl genome browser. Similarly, local BLASTn searches were performed in downloaded genome-wide or transcriptomic sequence resources of three cyclostomes (inshore hagfish, sea lamprey, and Japanese lamprey), three chondrichthyans (elephant shark, little skate, and cloudy catshark), a basal actinopterygian (spotted gar), the African coelacanth, reptiles (American alligator, Burmese python, Chinese soft-shell turtle, painted turtle, and saltwater crocodile), and birds (budgerigar, collared flycatcher, mallard, medium ground finch, turkey, and zebra finch; see supplementary table S1, Supplementary Material online, for details). Candidate sequences with *E* value of less than 1e-05 in these BLASTn searches were subjected to preliminary phylogenetic tree inferences. Sequences placed inside the Pax4/6/10 class were selected, and their longest open reading frames (ORFs) were curated either manually or by using the gene prediction program AUGUSTUS (Stanke et al. 2004). All sequences identified in this in silico screen, except for saltwater crocodile *Pax10* and Japanese lamprey *Pax6B* (supplementary tables S2 and S3, Supplementary Material online), are deposited in EMBL under accession numbers HF567444–HF567455.

Animals

Wild-type embryos and albino adults of the zebrafish (*Danio rerio*) and embryos (staging according to Nieuwkoop and Faber 1994) and adults of *Xenopus laevis* were obtained from captive breeding colonies. Eggs of the green anole lizard (*A. carolinensis*) were harvested from in-house captive breeding colonies and incubated at 28 °C and approximately 70% humidity until they reached required stages determined after Sanger et al. (2008). Animals that were subjected to sectioning or whole-mount in situ hybridizations were stored in methanol after fixation in 4% paraformaldehyde or Serra's fixative, respectively. All experiments were conducted in accordance with the animal use protocols of the University of Konstanz.

Reverse Transcription Polymerase Chain Reaction

Total RNA of embryos of zebrafish (24 hpf), *Xenopus* (11 individuals for each of stages 22, 30, and 40), and the green anole (stage 8.5) were extracted and reverse transcribed into cDNA using SuperScript III (Invitrogen), following the instructions of the 3′-RACE System (Invitrogen). Oligonucleotide primers were designed based on *Pax10* transcript sequences of zebrafish (ENSDART00000075395), *X. tropicalis* (ENSXETT00000065934), and the green anole (ENSACAT00000013868) to amplify full-length cDNAs of zebrafish *pax10a*, *X. laevis Pax10a* and *-10b*, and the green anole *Pax10* genes. 3′-ends of these four fragments were isolated by applying the 3′-RACE System (Invitrogen). 5′-ends of the green anole *Pax10* cDNA were obtained using the GeneRacer Kit (Invitrogen) and 5′-ends of the zebrafish *pax10a and X. laevis Pax10a* and *-10b* were obtained using the 5′-RACE System (Invitrogen). Full-length *Pax10* cDNA sequences of zebrafish, *X. laevis*, and green anole were assembled from

two fragments and are deposited in EMBL under accession numbers HF567440–HF567443.

In order to detect *Pax6* expression in the eyes of adult zebrafish, *Xenopus*, and the green anole, cDNA fragments covering the respective 3′-ends were isolated using the 3′-RACE System (Invitrogen). Primers were designed based on *Pax6* transcript sequences of zebrafish (Ensembl ID: ENSDART00000148420 for *pax6a* and ENSDART00000145946 for *pax6b*), *X. laevis* (NCBI: NM_001085944 for *Pax6a* and NM_001172195 for *Pax6b*), and the green anole (Ensembl ID: ENSACAT00000002377). Zebrafish *pax4* probe synthesis employed the cDNA fragment previously reported (Manousaki et al. 2011).

In order to analyze differential expression of *Pax6* and *-10* genes between various organs of adult zebrafish, *Xenopus* and the green anole, and between embryonic stages of developing zebrafish, the animals were dissected. Total RNA was extracted using TRIzol (Invitrogen) and treated with DNase I. The integrity of the extracted RNA was monitored using the Bioanalyzer 2100 (Agilent). Gene-specific primers to amplify approximately 200-bp-long cDNA fragments of *Pax6*, *-10*, and *GAPDH* were designed (supplementary table S4, Supplementary Material online). It should be noted that primers for the amplification of *X. laevis Pax6* and *-10* were designed to capture both "a" and "b" paralogs. The specificity of each primer pair was determined in a preliminary pilot polymerase chain reaction, and the amplification of comparable amounts of *GAPDH* cDNA fragments was used as proxy for similar quantity of cDNAs between samples. Semiquantitative reverse transcription (RT)-PCR runs were conducted using the DreamTaq™ DNA Polymerase (Fermentas). A predenaturing step at 95 °C for 3 min was followed by 35–50 cycles of three steps (95 °C for 30 s, 58 °C for 1 min, and 72 °C for 1 min). The amount of PCR product was visualized using standard agarose gel electrophoresis (supplementary fig. S1, Supplementary Material online). The intensity of individual bands was quantified using GelQuant.NET software (http://biochemlabsolutions.com, last accessed June 26, 2014). Expression levels of *Pax6* and *-10* genes were normalized to relative expression levels of *GAPDH*. The resulting values were further normalized to tissues with the highest expression level for the respective gene that was defined as 1. These values were shown in a heat map generated by CIMMiner (http://discover.nci.nih.gov/cimminer, last accessed June 26, 2014), which was also used for the clustering of genes based on expression patterns with an Euclidean distance algorithm.

In Situ Hybridization

The aforementioned 3′- and 5′-cDNA fragments of zebrafish, *Xenopus*, and green anole *Pax10*, 3′-fragments of zebrafish, *Xenopus*, and green anole *Pax6*, and 3′-fragment of zebrafish *pax4* were used as templates for riboprobe synthesis. Whole-mount in situ hybridizations for two *Pax10* probes and a *Pax6* probe as positive control were performed using embryonic zebrafish (Begemann et al. 2001), *Xenopus* (Gawantka et al. 1995), and green anole (Di-Poï N, personal communication). In situ hybridizations on paraffin-embedded sections for adult eyes of zebrafish, *Xenopus*, and green anole were performed using the aforementioned riboprobes as previously described (Kuraku et al. 2005).

Molecular Phylogenetic Analysis

An optimal multiple alignment of all retrieved cDNA sequences was constructed using MEGA5 (Tamura et al. 2011), in which the MUSCLE program (Edgar 2004) is implemented. Molecular phylogenetic trees were inferred using regions of selected vertebrate *Pax* amino acid sequences that were unambiguously aligned with no gaps (supplementary table S5, Supplementary Material online). Several sequence fragments of previously unidentified *Pax4/6/10* genes (supplementary table S6, Supplementary Material online) were excluded from the phylogenetic analysis depicted in figure 2*B*, but their affiliation to individual *Pax* subtypes was verified in separate phylogenetic analyses (data not shown). The human, *Cio. intestinalis*, and fly genes belonging to the Pax3/7 gene class served as outgroup because of the lack and truncation of the homeodomains of Pax1/9 and Pax2/5/8 class of genes, respectively. MEGA5 was used to determine the optimal amino acid substitution model and to reconstruct maximum likelihood (ML). Bayesian tree inference was performed using MrBayes 3.1 (Huelsenbeck and Ronquist 2001), with which we ran two independent chains with 5,000,000 generations for each, sampled every 100 generations, and excluded 25% of the samples as burnin. Convergence of the two chains was diagnosed using the Tracer v1.5 software (http://tree.bio.ed.ac.uk/software/tracer/, last accessed June 26, 2014). The main data set consisted of a broad selection of metazoan genes belonging to the Pax4/6/10 class of genes (fig. 2*B*). In addition, we performed a phylogenetic analysis focusing on vertebrate *Pax6* genes in which amphioxus *Pax6* genes served as outgroup (supplementary fig. S2, Supplementary Material online). To assess the statistical support for alternative hypotheses (fig. 2*A*), per site log likelihoods of two constrained tree topologies and the ML tree were calculated using Tree-Puzzle (Schmidt et al. 2002) and statistically assessed using CONSEL (Shimodaira and Hasegawa 2001).

Identification of Conserved Synteny

Using the Ensembl database, we identified genes within a 1-Mb region flanking *pax6* in the genome of the spotted gar. Orthologs of these genes in zebrafish, stickleback, and medaka were downloaded via the BioMart interface and plotted against the focal region of the spotted gar (fig. 3*A*). The three gene families that possess members both in the vicinity of *pax6a* and *-6b* genes were further analyzed. ML trees were reconstructed as described above to infer the evolutionary origin of the teleost duplicates (fig. 3*B*).

Genome Biology and Evolution

SMBE

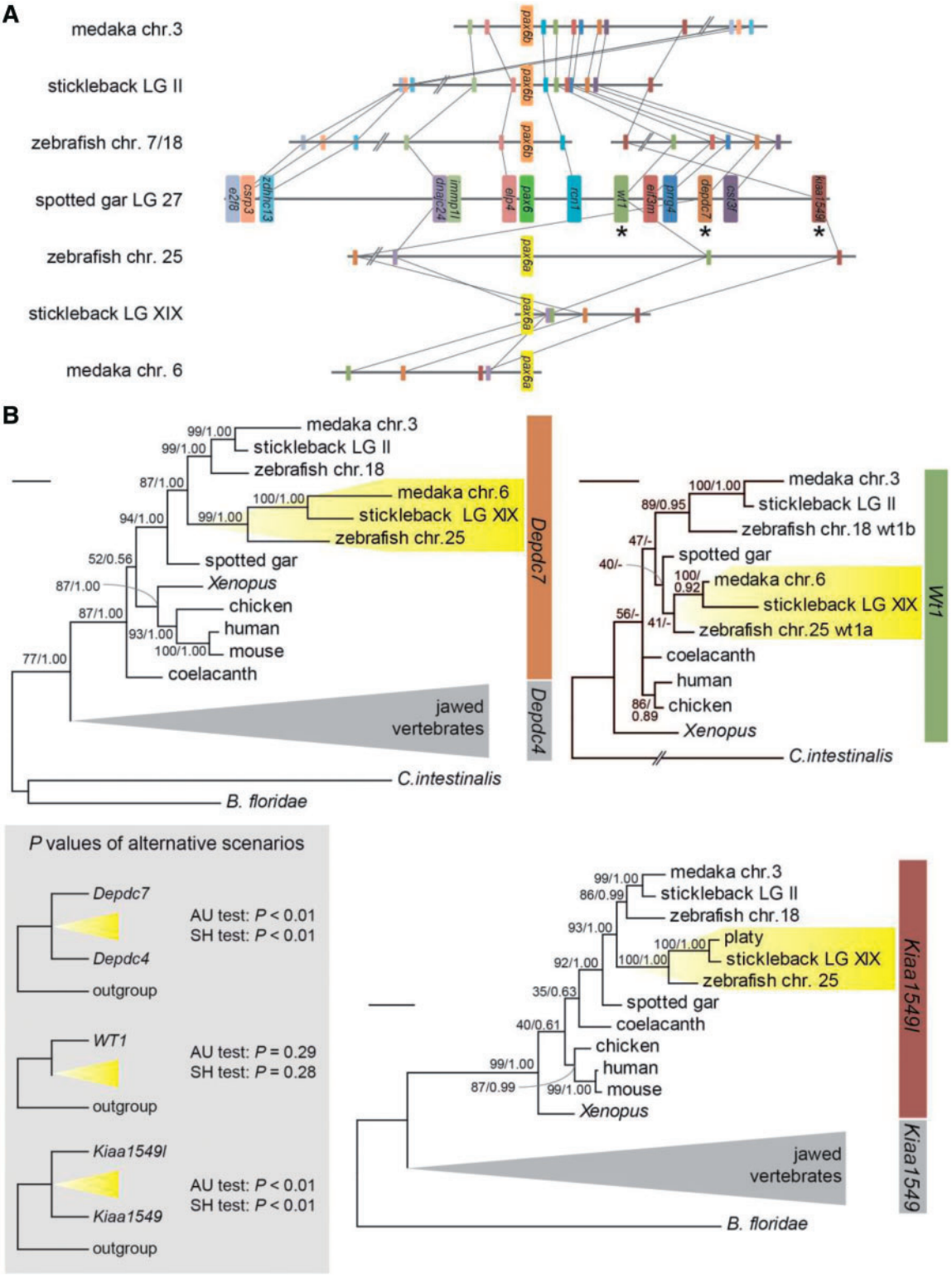

Fig. 3.—Evolutionary origins of *pax6*-neighboring genes in teleost genomes. (*A*) Conserved synteny between the genomic regions of the gar *Pax6* ortholog and the two teleost *pax6* subtypes. Co-orthologs to the spotted gar genes located within 1 Mb flanking the *pax6* gene that are harbored in the vicinity of *pax6a* and/or *-6b* of selected teleosts are shown in the same color (see supplementary table S9, Supplementary Material online, for accession IDs). Asterisks mark names of the selected gene families that retained both duplicates derived from the TSGD. (*B*) Molecular phylogenetic trees of selected gene families. ML trees of three gene families are shown, and statistical support values (*P* values) of alternative scenarios assuming a more ancient origin of

(continued)

The search for conserved intragenomic synteny (Kuraku and Meyer 2012) in the green anole focused on up to 10-Mb genomic regions flanking the *Pax6* and *-10* genes. Referring to the Ensembl "Gene Tree," the duplication patterns of neighboring gene families were investigated. We identified two additional chromosomal regions that harbor a similar gene array. A second search for paralogous relationships between these two identified chromosomal regions was conducted. The identified pairs, triplets, and quartets of paralogs located on the four chromosomes were plotted (fig. 4).

In order to analyze the mode of the putative loss of *Pax10* in the mammalian and avian lineages, we downloaded a list of Ensembl IDs of genes harbored in the 500-kb genomic regions flanking the green anole *Pax10* gene, together with IDs of opossum, human, and chicken orthologs of those genes via the BioMart interface. Opossum, human, and chicken orthologs located on chromosomes 4, 19, and 27, respectively, were plotted along the orthologous region on anole chromosome 6 (fig. 5 and supplementary tables S7 and S8, Supplementary Material online).

Results

Identification of Novel Pax4/6/10 Genes in Diverse Vertebrates Except for Mammals and Birds

We identified a group of protein-coding genes closely related to *Pax4* and *-6* in the Gene Tree view in release 54 of the Ensembl genome database (http://www.ensembl.org, last accessed June 26, 2014; released in spring 2009; Hubbard et al. 2009). In this tree at Ensembl, these uncharacterized genes possessed by zebrafish (ENSDARG00000053364) and stickleback (*Gasterosteus aculeatus*; ENSGACG00000007835) were placed basal to the vertebrate *Pax6* group of genes. More recent Ensembl releases (e.g., release 66) suggest that this gene is also possessed by *X. tropicalis* (ENSXETG00000032534), the green anole lizard (*A. carolinensis*; ENSACAG00000013797), Atlantic cod (*Gadus morhua*; ENSGMOG00000007260), and Nile tilapia (*Oreochromis niloticus*; ENSONIG00000020082). Following the conventional nomenclature of the *Pax* gene family (*Pax1–9*) and our intensive phylogenetic analysis (see below), we designated this group of genes as Pax10. Our survey of publicly available genome-wide sequence resources (supplementary table S1, Supplementary Material online) resulted in the identification of *Pax10* orthologs in the painted turtle (*Chrysemys picta bellii*; Shaffer et al. 2013), the Burmese python (*Python molurus*; Castoe et al. 2011), the American alligator (*Alligator mississippiensis*; St John et al. 2012), the saltwater crocodile (*Crocodylus porosus*; St John et al. 2012), the coelacanth (*Latimeria chalumnae*; Amemiya et al. 2013), and the elephant shark (*C. milii*; Ravi et al. 2013). Moreover, in *All. mississippiensis*, *Chr. picta bellii*, and *L. chalumnae*, we identified orthologs of the *Pax4* gene which was previously identified only in teleosts and mammals (see supplementary tables S1, S5, and S6, Supplementary Material online, for details). The *Pax4* genes of the Chinese soft-shell turtle (*Pelodiscus sinensis*; ENSPSIG00000006191) and the spotted gar (*Lepisosteus oculatus*; ENSGACP00000026154_1 at Ensembl Pre) are available in the Ensembl genome database since release 68 and in Pre-Ensembl, respectively. So far, no *Pax4* ortholog has been identified in transcriptomes and genomes of birds, lepidosaurs, amphibians, and chondrichthyans, whereas *Pax10* is most likely absent from the genomes of mammals and birds (fig. 1*B*). In addition to the previously described Japanese lamprey *Pax6* gene (AB061220 in NCBI nucleotide), we identified another *Pax6*-like sequence in the genome assembly of this species (Mehta et al. 2013) that we designated *Pax6B*. The putative ORF of this sequence is disrupted by a stretch of N's and thus is partial. Therefore, this Japanese lamprey *Pax6B* gene was not included in the phylogenetic analysis of the entire Pax4/6/10 class of genes (fig. 2), whereas it was included in the phylogenetic analysis of its subset, *Pax6* (supplementary fig. S2, Supplementary Material online). We would like to alert our readers to potential misidentification of these genes, as the novel relatives of *Pax6*, *Pax10* genes, are sometimes annotated as *Pax6* or *Pax6-like* in Ensembl.

Domain Structure of the Deduced Pax10 Protein Sequences

The deduced amino acid sequences of the *Pax10* genes, identified above in silico, contain a complete homeodomain preceded by a putative start codon, but lack the characteristic paired domain common to all other Pax proteins (fig. 1*A*). To rule out the possibility that incomplete annotations of genome assemblies led to the nonidentification of the paired domain, we performed extensive searches with tBLASTn in nucleotide sequences of the selected genomic contigs containing *Pax10* using peptide sequences of the paired domain as queries. However, this approach did not reveal any potential paired domain upstream of the identified *Pax10* ORFs. By means of RT-PCR, we isolated full-length cDNAs of a single Pax10 member in the zebrafish (designated *pax10a* by Ravi et al.

Fig. 3.—Continued

the *pax6a*-linked genes (compatible with Hypothesis 1 in fig. 2*A*) are given in the grey box on the lower left as inset. The *Depdc7* phylogeny is based on 321 amino acid residues and the JTT+Γ_4 model (shape parameter of the gamma distribution $\alpha = 1.39$) was assumed. The ML tree of *Wt1* genes was inferred from 331 amino acid residues assuming the JTT+F+Γ_4 model (shape parameter of the gamma distribution $\alpha = 1.45$). The *Kiaa1549l* phylogenetic analysis used an alignment of 295 amino acid residues and the JTT+Γ_4 model (shape parameter of the gamma distribution $\alpha = 1.30$). Bootstrap probabilities are provided for each node. The scale bars on the upper left of each phylogenetic tree indicate 0.2 substitutions per site.

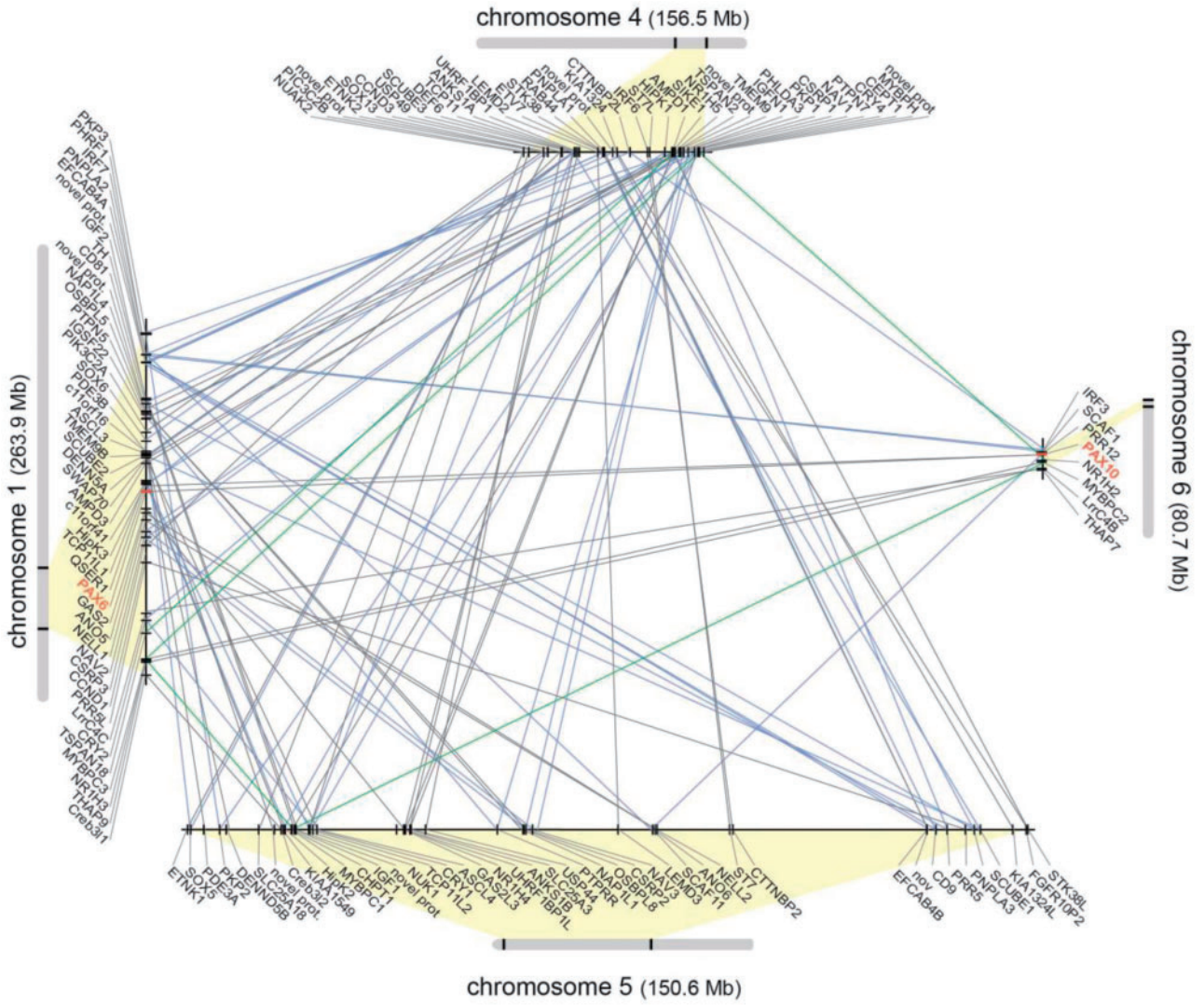

Fig. 4.—Intragenomic conserved synteny between *Pax6* and *-10* containing regions in the green anole lizard. Outer grey bars represent chromosomes 1 and 6 of the green anole that harbor *Pax6* and *-10* (shown in red), respectively, and their paralogous regions on chromosomes 4 and 5. Magnifications of the genomic regions indicated on the chromosomes (gray bars) are shown in the center. Gene-by-gene paralogies among the four members of the quartet are highlighted with diagonal lines: Gray lines for paralogy of gene families with two paralogs, blue lines for three paralogs, and green line for four paralogs.

2013 based on a comparison of syntenic relationships among teleosts) and *A. carolinensis*. We also isolated two full-length cDNAs of *X. laevis* (designated *Pax10a* and *-10b*)—evidence of transcription of these genes. In silico translation of these cDNA sequences confirmed the lack of a paired domain in the products of these *Pax10* genes.

Phylogenetic Relationships among *Pax4*, *-6*, and *-10*

We performed a molecular phylogenetic analysis employing a data set containing the members of the Pax4/6/10 class covering major metazoan taxa as well as human, *Cio. intestinalis*, and fly *Pax3/7* (paired) genes as outgroup (see supplementary table S5, Supplementary Material online). The inferred ML tree (fig. 2*B*) placed the putative *Nematostella vectensis Pax6* ortholog outside the monophyletic group containing bilaterian *Pax4*, *-6*, *and -10* genes. Within this group, gnathostome *Pax10* genes formed a monophyletic group (bootstrap probability in the ML analysis, 98). Except for the opossum *Pax6*, gnathostome *Pax6* genes formed a monophyletic group that was inferred to be a sister group of the *Pax10* group of genes (bootstrap probability in the ML analysis, 30). The two previously identified cyclostome *Pax* genes, namely *Pax6* of the inshore hagfish (*Eptatretus burgeri*) and *Pax6* of the Japanese lamprey (*Lethenteron japonicum*, also referred to as Arctic lamprey *Lethenteron camtschaticum*), form an exclusive cluster that is placed at the base of the gnathostome *Pax6* and *-10* subgroups (fig. 2*B*). It should be noted that the newly

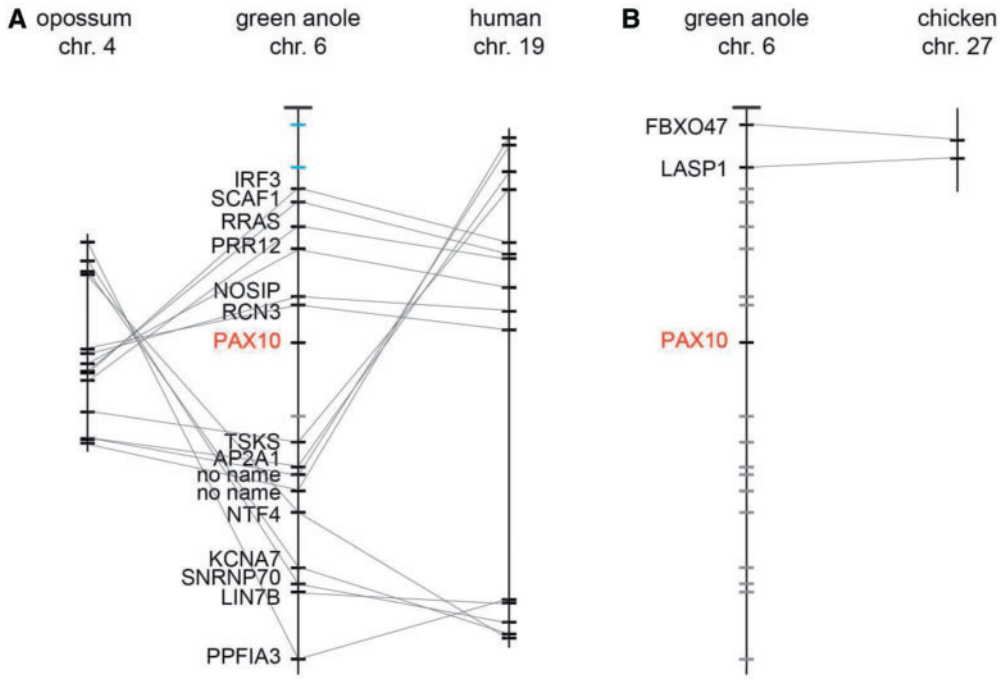

Fig. 5.—Conserved synteny between the *Pax10*-containing region in the green anole and its orthologous regions in mammals and birds. A 1-Mb region flanking the green anole *Pax10* gene, shown in red, was analyzed and gene-by-gene orthologies are indicated with gray lines. (*A*) Conserved synteny between the green anole *Pax10*-containing region and its orthologous regions in the opossum and human genomes. The dense pattern of one-to-one orthologies suggests a small-scale deletion of *Pax10* in the lineage leading to eutherians, before the split of the marsupial lineage. (*B*) Synteny between the green anole *Pax10*-containing region and the chicken genome. The lack of one-to-one orthologies in the region around the green anole *Pax10* gene is best explained by a large-scale deletion in the avian lineage. Green anole genes whose orthologs are located elsewhere in the opossum, human, or chicken genome are indicated with blue bars, whereas green anole genes lacking orthologs in these genomes are shown with grey bars. Exact genomic locations and accession IDs of the identified orthologs are included in supplementary tables S7 and S8, Supplementary Material online. chr., chromosome.

identified *L. japonicum Pax6B* gene and a *Pax6* gene (AAM18642.1 in NCBI) of a different lamprey species (*Lampetra fluviatilis*) were excluded from this analysis because of their incomplete sequence. However, a phylogenetic analysis focusing on vertebrate *Pax6* genes including the *L. japonicum Pax6B* gene supported its grouping with the inshore hagfish *Pax6* gene (bootstrap probability in the ML analysis, 73; supplementary fig. S2, Supplementary Material online). The previously identified *L. japonicum Pax6* gene showed higher affinity to gnathostome *Pax6* genes (bootstrap probability in the ML analysis, 54; supplementary fig. S2, Supplementary Material online). The phylogenetic relationships among invertebrate *Pax6/eyeless* genes barely reflect the generally accepted species phylogeny (fig. 2*B*). The group of *Pax4* genes forms a monophyletic cluster that is placed outside all other bilaterian *Pax6* and *-10* genes. Support values for this inferred tree are generally low. This poorly resolved tree topology prevents us from deriving more clear-cut conclusions about the phylogenetic relationships among the individual *Pax4*, *-6*, and *-10* subgroups. In particular, the phylogenetic position of the acanthopterygian *pax6a* (*pax6.3* in Ravi et al. 2013) genes in relation to other vertebrate *Pax6* genes could not be reliably inferred by this analysis. Although the ML analyses suggest an origin of acanthopterygian *pax6a* and *-6b* genes through the TSGD (fig. 2*B*), the support for this duplication node (bootstrap probability in the ML analysis, 18) is poor.

To test the hypothesis postulating an ancient origin of acanthopterygian *pax6a* (*pax6.3*) genes (fig. 2*A*; Ravi et al. 2013), we complemented the heuristic ML tree reconstruction with statistical tests. We simulated tree topologies for alternative scenarios supporting an early (Hypothesis 1; Ravi et al. 2013) or late (Hypothesis 2) origin of acanthopterygian *pax6a* (*pax6.3*) genes (fig. 2*A*) and computed likelihoods for them. Although the latter hypothesis yielded larger likelihood values, no significant difference was observed in the levels of support among the two hypotheses: The *P* values of Hypothesis 1 are 0.12 (approximately unbiased [AU] test) and 0.42 (Shimodaira–Hasegawa [SH] test), and those of Hypothesis 2 are 0.38 (AU test) and 0.44 (SH test). Thus, the question about the evolutionary origin of acanthopterygian *pax6a* (*pax6.3*) genes cannot be solely answered with a molecular phylogenetic approach. Therefore, we investigated gene orders in the genomic regions flanking teleost *pax6* genes.

We detected conserved synteny between the genomic region containing the single *Pax6* gene in the spotted gar

and those containing *pax6a* and *-6b* genes in medaka, stickleback and zebrafish (fig. 3*A* and supplementary table S9, Supplementary Material online). The TSGD presumably occurred after the split of teleost fish from the lineage of Lepisosteiformes, including the spotted gar (Hoegg et al. 2004; Crow et al. 2006; Hurley et al. 2007; Amores et al. 2011). Thus, this one-to-two relationship between the spotted gar and teleost fish matches the pattern resulting from the TSGD. We also investigated the molecular phylogeny of the neighboring gene families (fig. 3*B*). We sought to determine the timing of duplication that gave rise to pairs of genes flanking acanthopterygian *pax6a* and *-6b*, namely in the *Depdc7*, *Wt1*, and *Kiaa1549l* gene families. The inferred ML trees suggested a duplication that gave rise to the teleost paralogs in question in the actinopterygian fish lineage (fig. 3*B*). Two of the gene families (*Depdc7* and *Kiaa1549l*) indicated duplications after the split of the spotted gar from the teleost stem lineage, coinciding with the TSGD. To further test this hypothesis, we computed likelihoods of alternative, non-ML tree topologies supporting the ancient origin of *pax6a*-linked (*pax6.3*) genes at the base of vertebrate evolution. This analysis rejected an ancient origin of *pax6a*-linked genes at the base of vertebrate evolution for two of the three gene families (*Depdc7* and *Kiaa1549l*) with *P* values below 0.1 in the AU and SH tests. For the *Wt1* gene family, this hypothesis was not significantly rejected (*P* value of 0.29 in the AU test and 0.28 in the SH test), but was less likely than a more recent duplication in the TSGD (fig. 3*B*). Although not unambiguously inferred by our analysis, the timing of duplication of *Wt1* genes in the TSGD has been demonstrated previously (Kluever et al. 2009). The synchrony of the duplications of these flanking genes suggests that they were duplicated in a large-scale event, probably in the TSGD. Hence, the paralogy between *pax6a* and *-6b* genes, embedded in this region, should have originated through the TSGD (see fig. 2*A*).

Are *Pax4*, *-6*, and *-10* Derived from the 2R-WGD?

Because our molecular phylogenetic analysis did not yield high confidence about the timing of diversification of the Pax4/6/10 gene class, we employed a synteny analysis (see Materials and Methods). An intragenomic synteny analysis of the green anole lizard showed a dense pattern of gene-by-gene paralogies among four chromosomal regions including those containing *Pax6* and *-10* (fig. 4). The gene families whose duplications coincide with the 2R-WGD show a dense pattern of gene-by-gene paralogy among the identified regions on chromosomes 1, 4, 5, and 6 (fig. 4). In addition, five of the identified gene families were previously shown to have diversified early in vertebrate evolution (Manousaki et al. 2011). Interestingly, the orthologous regions in the human and chicken genomes were suggested to be derived from a single ancestral vertebrate chromosome, namely "D," reported by Nakatani et al. (2007). These results suggest that the quartet of the identified regions, on chromosomes 1, 4, 5, and 6 in the green anole genome, originated from a large-scale quadruplication, probably as a result of the 2R-WGD.

Mode of Secondary Loss of *Pax10* in Mammals

Our search for *Pax10* genes in public databases and subsequent phylogenetic analyses suggested the absence of this gene in mammals and birds (see figs. 1*B* and 2). We employed intergenomic synteny analyses in order to determine whether the absence was caused by a large-scale deletion of a chromosomal segment or a small-scale gene loss. If it is caused by a major deletion, simultaneous loss of multiple neighboring genes should be observed. Our comparison between the 1-Mb genomic region flanking the *Pax10* gene in the green anole, the human, and opossum genomes revealed conserved synteny (fig. 5*A*). Of the 19 protein-coding genes contained in this region in the green anole, 15 and 14 genes, respectively, have orthologs in particular regions on human chromosome 19 and opossum chromosome 4 (see supplementary table S7, Supplementary Material online, for exact positions of the orthologs). Notably, orthologs of one immediate neighbor of *Pax10* (an uncharacterized gene encoding 776 amino acid-long putative peptide; ENSACAG00000013620) are also absent from mammalian genomes. Thus, the absence of the *Pax10* gene from mammalian genomes was presumably caused by a small-scale deletion early in mammalian evolution before the divergence between the marsupial and eutherian lineages. If there was no massive chromosomal rearrangement that could have hindered our synteny analyses, this deletion event in the mammalian lineage could have involved as few as two genes.

We also performed a synteny comparison between the green anole and the chicken. Of the 19 genes in the *Pax10*-containing region in the green anole genome, only two were unambiguously shown to have orthologs in the chicken genome (see supplementary table S8, Supplementary Material online, for positions of the orthologs). Those two anole genes are more than 230 kb away from the *Pax10* gene (fig. 5*B*). The absence of the chicken orthologs of *Pax10* and other genes in this genomic region in anole suggests that its counterpart region got lost in the lineage leading to chicken. Compared with a small-scale loss in the mammalian lineage, the loss in the avian lineage could have been larger in size, possibly involving more than ten genes.

Expression Patterns of *Pax10* in Zebrafish, *Xenopus*, and Green Anole

We conducted semiquantitative RT-PCR using a developmental series of zebrafish to characterize the temporal expression profile of *pax10a*. This experiment suggested an onset of *pax10a* expression at 25 h postfertilization (hpf) with the maximal expression level at 5 days postfertilization (dpf; fig. 6*A* and supplementary fig. S1, Supplementary Material online).

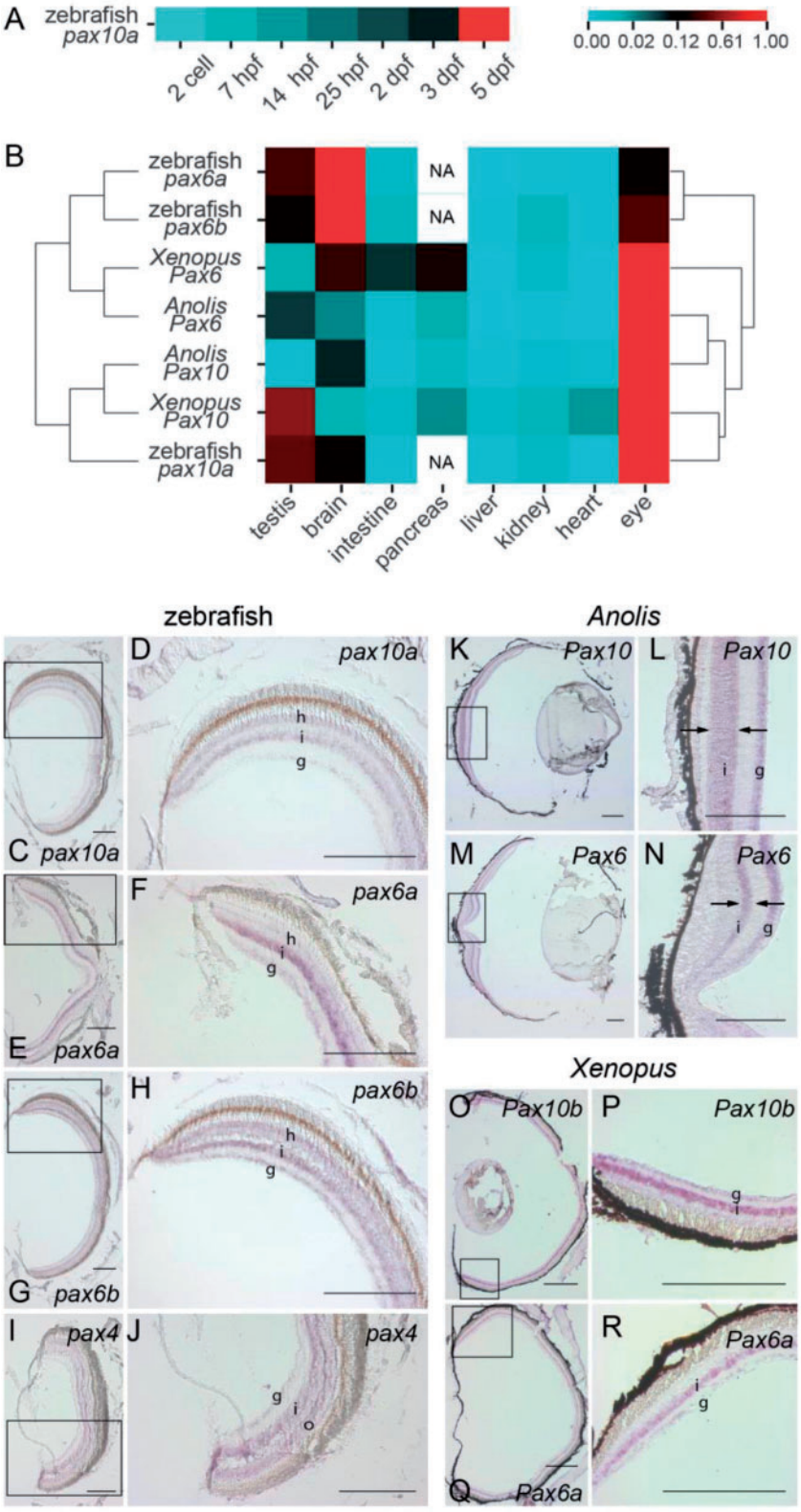

Fig. 6.—Expression profiles of *Pax4*, *-6*, and *-10* in zebrafish, *Xenopus*, and green anole. (*A*) Expression levels of *pax10a* in a developmental series of zebrafish with semiquantitative RT-PCRs. Heat map indicates upregulation of *pax10a* in late developmental stages reaching a plateau at 5 dpf. Color scale at the right indicates the relative expression levels normalized to values between 0 and 1 for no expression (blue) and for the highest observed expression level

(continued)

This result prompted us to examine whether *Pax10* is expressed in adult animals. Therefore, we analyzed *Pax10* expression levels, as well as those of *Pax6*, in individual organs of zebrafish, *Xenopus*, and green anoles. Our comparison revealed strong *Pax10* expressions in the eyes of all three examined animals and also subtle expression in testes of *Xenopus* and brain tissue of zebrafish and the green anole (fig. 6*B*). It also revealed intensive *Pax6* expressions in the eyes and weaker expressions in the brain, pancreas, and testis, which was common to all the animals examined (fig. 6*B*). We performed clustering of the analyzed genes based on their similarity in expression levels in the examined organs (fig. 6*B*). It partly recovered their phylogenetic relationship—closely related gene pairs (e.g., zebrafish *pax6a* and *-6b*, as well as *Xenopus Pax10* and zebrafish *pax10a*) also showed the highest similarity in expression patterns.

In order to scrutinize the expression patterns of these genes, we performed in situ hybridizations on sectioned eyes of adult zebrafish, *Xenopus*, and green anoles. We performed expression analyses of *pax4*, *-6a, -6b*, and *-10a* genes in zebrafish, *Pax6* and *-10* in the green anole, and *Pax6a* and *-10b* in *X. laevis* and identified distinct expression signals of all investigated genes in specific layers of the mature retina, namely, intensively in the inner nuclear layer and weakly in the ganglion cell layer (fig. 6*C–R*). In addition, we identified transcripts of zebrafish *pax6a*, *-6b*, and *-10a* in the horizontal cell layer (fig. 6*D*, *F*, and *H*) and those of zebrafish *pax4* in the outer nuclear layer (fig. 6*J*). In the retina of the green anole, the expression domain of *Pax6* in the inner nuclear layer was revealed to be nested within the broader area of *Pax10*-expressing cells (fig. 6*L* and *N*).

Discussion

Phylogeny of the Pax4/6/10 Class of Genes

Only recently, a novel relative of *Pax6*, designated *Pax6.2* (*Pax10* in this study), was reported (Ravi et al. 2013). This study ignored *Pax4*, the other relative previously revealed to have been split from *Pax6* in the 2R-WGD (Manousaki et al. 2011). Therefore, in this study, we included *Pax4* genes in the data set and sought to reconstruct the evolutionary history of the entire Pax4/6/10 gene class. The group of *Pax4* genes, with long branches, was placed outside the clade containing vertebrate and invertebrate *Pax6* genes as well as *Pax10* genes (fig. 2*B*). Monophyly of *Pax6* could not be unambiguously inferred (fig. 2*B*), and the group of *Pax10* genes was placed inside the vertebrate *Pax6* genes. As shown for many other genes (Qiu et al. 2011; Kuraku 2013), the phylogenetic positions of the cyclostome *Pax6* homologs were not unambiguously determined (fig. 2*B*). Still, Japanese lamprey *Pax6B*, newly reported in this study, was suggested to be an ortholog of the previously reported hagfish *Pax6* gene (supplementary fig. S2, Supplementary Material online). The previously reported Japanese lamprey *Pax6* was instead suggested to be orthologous to gnathostome *Pax6* genes (supplementary fig. S2, Supplementary Material online). As for the position of acanthopterygian *pax6b* genes questioned by Ravi et al. (fig. 2*A*), our phylogenetic analysis (fig. 2*B*) supported the scenario that they originated in a duplication in the teleost fish lineage (Hypothesis 2) instead of the scenario that it originated in the 2R-WGD (Hypothesis 1).

Synteny analyses have been widely used as reliable tool to identify traces of WGD (Kuraku and Meyer 2012 and references therein). It is particularly useful when molecular phylogenies of the involved genes cannot be reliably reconstructed. Our analysis demonstrated conserved synteny among four chromosomal regions including the two containing *Pax6* and *-10*, respectively (fig. 4). The duplications giving rise to this quartet, inferred by the diversification patterns of neighboring gene families (Manousaki et al. 2011), occurred after the split of the tunicate and cephalochordate lineages but before the cyclostome–gnathostome split. This timing coincides with the 2R-WGD early in vertebrate evolution (Kuraku et al. 2009).

Origin of Acanthopterygian *pax6a* Genes

The study by Ravi et al. suggested the origin of acanthopterygian *pax6a* (*pax6.3*) genes in the 2R-WGD (Hypothesis 1 in

Fig. 6.—Continued

(red), respectively. (*B*) Heat map visualizing expression levels of *Pax6* and *-10* in individual organs of adult zebrafish, *Xenopus*, and green anole. In adult zebrafish, *pax6a*, *-6b*, and *-10a* transcripts were detected in the brain, testis, and eye. High levels of *Pax6* expression were detected in *Xenopus laevis* brain and eye, and low levels in pancreas and intestine, whereas *Pax10* expression signals were found in the eye and testis. In green anole, *Pax6* transcripts were detected in the eye, brain, and testis and at low concentrations in pancreas, and *Pax10* expression was detected in the eye and brain. It should be noted that the zebrafish pancreas was not analyzed (NA) because its anatomical structure was not precisely identified. The phylogram on the left reflects the phylogenetic relationships of the genes inferred in figure 2*B*, whereas that on the right shows the clustering based on their expression levels in various organs. (*C–R*) In situ hybridizations of *Pax4*, *-6*, and *-10* orthologs in the retinas of adult zebrafish, *Xenopus*, and the green anole. *C–J* show expression patterns in the retinas of an albino zebrafish, *K–N* of a green anole, and *O–R* of a *X. laevis*. *D*, *F*, *H*, *J*, *L*, *N*, *P*, and *R* are magnifications of the rectangles in *A*, *C*, *E*, *G*, *I*, *K*, *M*, and *O*, respectively. All investigated genes were strongly expressed in the inner nuclear layer (i) and weakly in the ganglion cell layer (g). Zebrafish *pax6a*, *-6b*, and *-10a* also showed weak expressions in the horizontal cell layer (h in *D*, *F*, and *H*) and zebrafish *pax4* transcripts were detected in the outer nuclear layer (o in *J*). It was evident from the results for *Anolis carolinensis* that the *Pax6* expression is nested within that of *Pax10* in the inner nuclear layer of the retina (arrows in *L* and *N*). Scale bar: 200 μm.

fig. 2*A*). Evidence for this hypothesis was: 1) "Subpartitioning" of conserved noncoding elements, 2) the absence of exon 5a from *pax6a* (*pax6.3*) genes, and 3) a simple molecular phylogenetic analysis with the neighbor-joining method and excluding *Pax4* genes (Ravi et al. 2013). The first evidence, that is, the divergence or loss of *cis*-elements between acanthopterygian *pax6a* (*pax6.3*) and other gnathostome *Pax6* genes, can be explained by subfunctionalization and asymmetric retention of expression domains after the gene duplication. It is a recurrent pattern in gene family evolution that one duplicate, here the acanthopterygian *pax6b*, retains the ancestral role, whereas the other duplicate, *pax6a* (*pax6.3*) in this case, is more prone to changes in their functions (reviewed in Prince and Pickett 2002). The sole reliance on presence and absence of conserved noncoding elements, without any tree-based approach, might therefore mislead the reconstruction of phylogenetic relationships. Regarding the second evidence, Ravi et al. regarded the lack of exon 5a in acanthopterygian *pax6a* (*pax6.3*) genes as the evidence of exclusion of those genes from the canonical *Pax6* genes (Ravi et al. 2013). On the other hand, however, they report "N-terminal protein extension" of all teleost *pax6* genes (including both *pax6a* and *-6b*), which can be regarded as evidence of phylogenetic proximity between *pax6a* and *-6b*. Thus, the comparison of gene structures does not add to our understanding of the evolution of *Pax6* genes. As for the third evidence mentioned above, we reanalyzed the phylogeny of the entire Pax4/6/10 class by employing a more rigorous methodology in light of the standard procedure of molecular phylogenetics (Anisimova et al. 2013). Although our molecular phylogenetic analysis focusing on the *Pax* gene family did not unambiguously solve the question about the timing of the acanthopterygian *pax6a–6b* split, expansion of the scope of our study to a broader genomic region enhanced the power of the analysis (fig. 3). Two of the three gene families whose members reside near *pax6a* and *-6b* (*Depdc7* and *Kiaa1549l*) unambiguously rejected their more ancient origins in the 2R-WGD, and one gene family (*WT1*) also strongly favored their duplication in the TSGD (fig. 3*B*). This strongly suggests that these genomic regions, including *pax6a* and *-6b* genes, originated in the TSGD (Hypothesis 2 in fig. 2*A*). Ravi et al. also analyzed conserved synteny (fig. 6*B* and *C* of Ravi et al. 2013) but did not recognize the copresence of the *Depdc7*, *Wt1*, and *Dnajc24* family members as the evidence of the orthology between the zebrafish *pax6a* (*pax6.3* in their hypothesis) and the *pax6b* of the other teleost fishes. Another argument supporting Hypothesis 2 is the phylogenetic distribution of teleost *pax6* genes. According to Ravi et al., acanthopterygian lost one of the TSGD-derived duplicates orthologous to *Pax6* and is the only taxon retaining the 2R-derived *pax6.3* gene (Hypothesis 1 in fig. 2*A*). This hypothesis assumes independent losses of its orthologs in all other vertebrate lineages (cyclostomes, chondrichthyans, nonteleost actinopterygians, coelacanth, and tetrapods as depicted in fig. 9 in Ravi et al. 2013). Taken together, we strongly argue that the genes designated as *pax6.3* by Ravi et al. are indeed the second TSGD-derived gene, orthologous to *Pax6*, which should continue to be called *pax6a* (Hypothesis 2 in fig. 2*A*).

Asymmetric Gene Retention Rates between *Pax6* and *Pax4/-10*

In the course of vertebrate evolution, *Pax4*, *-6*, and *-10* genes experienced different patterns of retention between vertebrate lineages. Although the *Pax6* gene is present in all vertebrates investigated to date, *Pax4* and *-10* were lost multiple times secondarily in independent lineages (fig. 1*B*). This asymmetry of gene retention was also manifested after the TSGD—in each teleost genome that we analyzed, we did not identify more than one *pax10* gene (fig. 2*B*). Using a synteny-based approach, it was suggested that the zebrafish *pax10a* gene is a TSGD-derived paralog of the identified acanthopterygian *pax10b* genes (Ravi et al. 2013). The tetraploid *X. laevis* is the only species possessing two copies of both *Pax6* and *-10* genes (see fig. 2*B*).

Although the general trend of unequal patterns of gene retention is clear, the exact number of secondary gene losses of *Pax4* and *-10* remains difficult to determine. The certainty with which one can infer a secondary gene loss in a given lineage depends on the quality of available sequence resources and the density of taxon sampling. In Ensembl release 71, the mammalian lineage is represented with 40 genomes, including those sequenced with low coverage, whereas only one amphibian and six sauropsid genomes are available. However, the fact that we identified a *Pax6* gene in every genome we investigated suggests that the absences of *Pax4* and *-10* genes are not caused by insufficient sequence information, but rather by secondary gene losses. In contrast, the absence of *Pax10* genes from the mammalian lineage can be inferred with a high degree of confidence. By investigating synteny between the anole *Pax10*-containing genomic region and orthologous regions in mammalian and avian genomes, we demonstrated that different modes of gene loss caused the absence of *Pax10* in these two lineages. Although we identified a trace of small-scale loss in mammalian genomes, loss of a larger genomic segment was suggested in the avian lineage (fig. 5).

Functional Diversification of *Pax4, -6,* and *-10* Genes

An explanation for the asymmetric gene retention between *Pax6* and the other two paralogs, *Pax4* and *-10* genes, might be rapid subfunctionalization soon after the 2R-WGD in which only *Pax6* retained the essential role as master control gene for eye development. This crucial role could have imposed a high selection pressure on the *Pax6* gene making it indispensable. Comparison of *Pax10* expression patterns among a teleost fish (zebrafish), an amphibian (*X. laevis*), and a reptile (the green anole lizard) allowed us to corroborate this hypothesis and infer the ancestral expression profile of the euteleostome

Pax10 gene. We did not identify significant *Pax10* expression in early embryos of the three examined species, and quantification of *Pax10* transcripts of a developmental series of zebrafish suggested the onset of *pax10a* expression during later developmental stages (fig. 6*A*). This finding is in accordance with a previous study reporting an increasing expression level of zebrafish *pax10a* from 1 dpf to 5 dpf (Ravi et al. 2013). Morpholino depletion for *pax10a* in zebrafish causes a reduction in eye size (Ravi et al. 2013), indicating a restricted role of *pax10a* to eye growth. The crucial role in the initiation of eye development is presumably retained solely by the *Pax6* gene.

Our semiquantitative RT-PCR of adult zebrafish, *Xenopus*, and green anoles confirmed the *Pax6* expression in the eye, pancreas, and brain (fig. 6*B*), as described previously (Nornes et al. 1998; Moreno et al. 2008). Exceptions are the identified expression signals in the testes of *Pax6* as well as of *Pax10*. This expression domain has not been described for any *Pax4* or *-6* gene before. Excessive gene expression in testes can be explained by a permissive state of chromatin in testes owing to the peculiar properties of transcription in meiotic and postmeiotic cells ("out of testis" hypothesis; Kaessmann 2010). In light of this hypothesis, these identified transcripts might merely be caused by leaky transcription. Results of our semiquantitative RT-PCR revealed expressions of *Pax10* genes in the eyes of all three investigated species and also in the brain tissue of zebrafish and the green anole, but not *X. laevis* (fig. 6*B*). Our inference based on the maximum parsimony principle suggests that the *Pax10* gene in the ancestral euteleostome was expressed in the eyes, testes, and brains of adults. The elephant shark *Pax10* gene also shows expression in mature eyes (Ravi et al. 2013), indicating that the *Pax10* expression in the adult eyes was established before the split between the euteleostome and chondrichthyan lineages. On the other hand, we observed among-lineage differences in the *Pax10* expression: Expression in the adult kidney is unique to the elephant shark, whereas those in the adult brains and testes are confined to euteleostomes (fig. 6*B*; Ravi et al. 2013).

The expression signals of *Pax4*, *-6*, and *-10* in the adult eyes are detected in specific layers in the retina (fig. 6*C*–*R*). The retinal expression of *Pax6* genes in the inner and the ganglion cell layer identified in this study is equivalent to that in other vertebrate species (reviewed in Osumi et al. 2008). However, we did not detect *Anolis* and *Xenopus Pax6* transcripts in the horizontal cell layer (fig. 6*K*–*R*), an expression domain previously reported for several vertebrates, including an elasmobranch shark, zebrafish, chicken, and mouse (Belecky-Adams et al. 1997; Macdonald and Wilson 1997; Wullimann and Rink 2001; Ferreiro-Galve et al. 2011). We report for the first time expression of a *Pax4* gene, namely zebrafish *pax4*, in the inner nuclear and the ganglion cell layer in addition to the outer nuclear layer (fig. 6*J*) as reported in mammals (Rath et al. 2009). The expression of *Pax10* is almost identical to *Pax6*, except that the *Pax6* expression is nested within the broader area of *Pax10*-positive cells in *Anolis* (fig. 6*L*–*N*). The expression of zebrafish *pax10a* appears to alter in the course of development. Although the *pax10a* expression is restricted to the inner nuclear layer at 2 dpf (Ravi et al. 2013), it expands to the horizontal and ganglion cell layer in the adult retina (fig. 6*C* and *D*). A parsimonious reconstruction of the nondevelopmental expression profile indicates a role of *Pax10* in the mature retina and brain of the euteleostome ancestor (fig. 7).

Although the proto-*Pax4/6/10* gene in the vertebrate ancestor had a dual role in the visual system, namely an early one in eye morphogenesis and a late one in mature photoreceptors, these roles could have been subdivided during vertebrate evolution. The vertebrate *Pax6* gene retained the crucial role in eye morphogenesis, whereas *Pax4* and *-10* as well as *Pax6* are coexpressed in the mature vertebrate retina (fig. 7).

The redundancy resulting from the quadruplication of a proto-*Pax4/6/10* gene in the 2R-WGD presumably led to a functional partitioning among the four paralogs early in vertebrate evolution. In order to understand these secondary changes in associated developmental pathways, the ancestral expression profile of the proto-*Pax4/6/10* gene needs to be reconstructed and compared with those of individual *Pax4, -6*, and *-10* genes in extant vertebrates. This reconstruction of the ancestral state should be reinforced by comparing expression patterns of their orthologs of nonvertebrate deuterostomes with those of protostomes. So far, sparse information of expression patterns in adults for most of invertebrates does not allow a reliable reconstruction of nondevelopmental roles of the ancestral *Pax4/6/10* gene. However, expression of an *eyeless* isoform in adult *Drosophila* eyes, more precisely in photoreceptors, has been reported (Sheng et al. 1997). The reconstruction of the ancestral expression profile of Urbilateria suggests a role of the proto-*Pax4/6/10* gene in the developing visual and olfactory systems, the CNS, and mature photoreceptors (fig. 7).

Pancreatic expression of *Pax4* and *-6* was reported in mammals (Turque et al. 1994; Sosa-Pineda et al. 1997) and teleost fish (Thisse and Thisse 2004; Manousaki et al. 2011). The single *Pax4/6/10* proto-ortholog in amphioxus lacks expression in the possible pancreas homolog (Reinecke 1981; Glardon et al. 1998; Sun et al. 2010). After the split of cephalochordate and tunicate lineages, the proto-*Pax4/6/10* should have gained the pancreatic expression domain in the vertebrate ancestor before the 2R-WGD. Thus, expression of *Pax4* and *-6* in the pancreas presumably represents a synapomorphy of vertebrates. On the other hand, the vertebrate *Pax10* gene, the third paralog, presumably did not retain the pancreatic expression domain (fig. 7).

Another interesting feature of the newly identified *Pax10* gene is the lack of a paired domain in the deduced protein sequence (fig. 1*A*; Ravi et al. 2013). Although no other *Pax* gene has ever been shown to completely lack the eponymous paired box, there are reports of "paired-less" *Pax6* isoforms

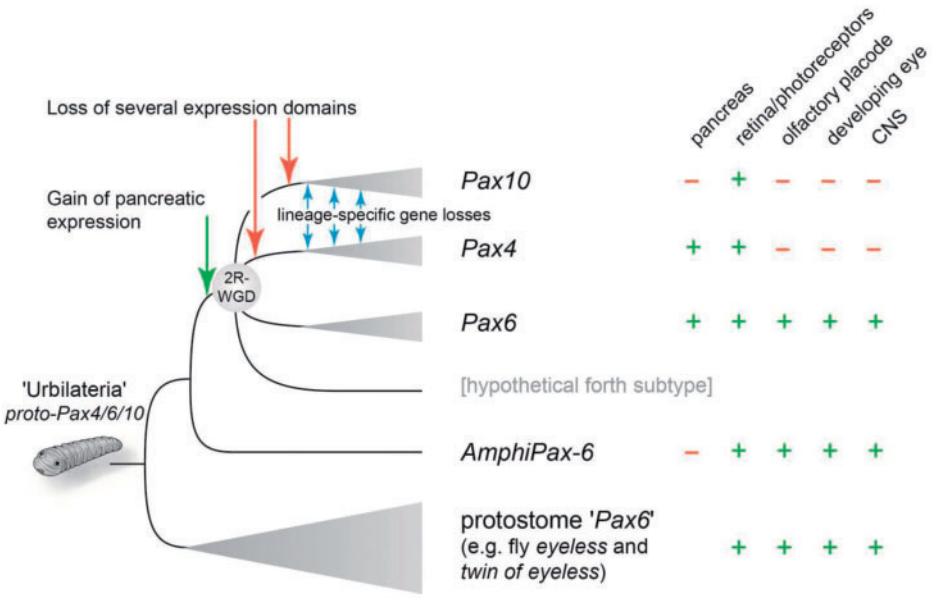

Fig. 7.—Evolutionary scenario focusing on the functional diversification of the Pax4/6/10 class of genes. Expression domains identified in this study are mapped with those previously described onto a simplified gene tree of the Pax4/6/10 class. Based on parsimonious reconstruction, a proto-*Pax4/6/10* gene at the last common ancestor of protostomes and deuterostomes, the so-called Urbilateria, was most likely expressed in photoreceptors, olfactory placode, developing the eye and CNS. Secondary modification, such as the gain of a pancreatic expression before the 2R-WGD or the loss of several *Pax4* and *-10* expression domains, led to the functional differentiation among *Pax4*, *-6*, *and -10*.

through alternative splicing in quail (Carriere et al. 1993), mouse (Kammandel et al. 1999), and zebrafish (Lakowski et al. 2007), as well as the ortholog in *Caenorhabditis elegans*, *vab-3* (Zhang and Emmons 1995). It was shown in mice that paired-less *Pax6* is one of the three *Pax6* isoforms whose transcriptions are initiated from three alternative promoters (Kammandel et al. 1999) and is expressed in a tissue-specific manner (Mishra et al. 2002). The structural similarity between paired-less *Pax6* isoforms and the *Pax10* gene is intriguing. It will be interesting to disentangle the functional interplay on the cellular level among the paired-less *Pax10* gene, isoforms of *Pax6* and also *Pax4*. Teleosts, in particular zebrafish, would be an ideal system as they are amenable to genetic experiments and have retained orthologs of *Pax4*, *-6*, and *-10*. Chromatin immunoprecipitation combined with sequencing (ChIP-seq) experiments could potentially reveal differences and overlaps in the target DNA of these transcription factors. Knockout (or -down) experiments of *pax10* could reveal to what extent its sister genes might compensate for its loss of function and thereby reveal potential redundancy. This experiment might allude to possible consequences that a secondary gene loss of *Pax4* and/or *-10* might have had on associated regulatory pathways.

Supplementary Material

Supplementary figures S1 and S2 and tables S1–S9 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors thank Gerrit Begemann, Nicola Blum, and Dominique Leo for the supply of zebrafish; Thomas Meyer and Thomas Tischer for providing *Xenopus laevis* specimens; and Ursula Topel, Lenia Beck, Mario Hupfeld, and Hans Recknagel for technical support in RNA extraction, cDNA cloning, and in situ hybridizations. This work was supported by the Young Scholar Fund of the University of Konstanz and the research grant (KU2669/1-1) from German Research Foundation (DFG) to S.K., support of the University of Konstanz to A.M., and by the International Max Planck Research School (IMPRS) for Organismal Biology to N.F. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Literature Cited

Amemiya CT, et al. 2013. The African coelacanth genome provides insights into tetrapod evolution. Nature 496:311–316.

Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. 2011. Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. Genetics 188:799–808.

Amores A, et al. 1998. Zebrafish hox clusters and vertebrate genome evolution. Science 282:1711–1714.

Anisimova M, et al. 2013. State-of the art methodologies dictate new standards for phylogenetic analysis. BMC Evol Biol. 13:161.

Begemann G, Schilling TF, Rauch GJ, Geisler R, Ingham PW. 2001. The zebrafish neckless mutation reveals a requirement for raldh2 in mesodermal signals that pattern the hindbrain. Development 128: 3081–3094.

Belecky-Adams T, et al. 1997. Pax-6, Prox 1, and Chx10 homeobox gene expression correlates with phenotypic fate of retinal precursor cells. Invest Ophthalmol Vis Sci. 38:1293–1303.
Breitling R, Gerber JK. 2000. Origin of the paired domain. Dev Genes Evol. 210:644–650.
Carriere C, et al. 1993. Characterization of quail Pax-6 (Pax-QNR) proteins expressed in the neuroretina. Mol Cell Biol. 13:7257–7266.
Castoe TA, et al. 2011. Sequencing the genome of the Burmese python (*Python molurus bivittatus*) as a model for studying extreme adaptations in snakes. Genome Biol. 12:406.
Chi N, Epstein JA. 2002. Getting your Pax straight: Pax proteins in development and disease. Trends Genet. 18:41–47.
Chiari Y, Cahais V, Galtier N, Delsuc F. 2012. Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). BMC Biol. 10:65.
Crawford NG, et al. 2012. More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biol Lett. 8:783–786.
Crow KD, Stadler PF, Lynch VJ, Amemiya C, Wagner GP. 2006. The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. Mol Biol Evol. 23:121–136.
Delporte FM, et al. 2008. Expression of zebrafish pax6b in pancreas is regulated by two enhancers containing highly conserved cis-elements bound by PDX1, PBX and PREP factors. BMC Dev Biol. 8:53.
Djiotsa J, et al. 2012. Pax4 is not essential for beta-cell differentiation in zebrafish embryos but modulates alpha-cell generation by repressing arx gene expression. BMC Dev Biol. 12:37.
Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.
Feiner N, Begemann G, Renz AJ, Meyer A, Kuraku S. 2009. The origin of bmp16, a novel Bmp2/4 relative, retained in teleost fish genomes. BMC Evol Biol. 9:277.
Ferreiro-Galve S, Rodriguez-Moldes I, Candal E. 2011. Pax6 expression during retinogenesis in sharks: comparison with markers of cell proliferation and neuronal differentiation. J Exp Zool B Mol Dev Evol. 318: 91–108.
Gawantka V, Delius H, Hirschfeld K, Blumenstock C, Niehrs C. 1995. Antagonizing the Spemann organizer: role of the homeobox gene Xvent-1. EMBO J. 14:6268–6279.
Glardon S, Holland LZ, Gehring WJ, Holland ND. 1998. Isolation and developmental expression of the amphioxus Pax-6 gene (AmphiPax-6): insights into eye and photoreceptor evolution. Development 125: 2701–2710.
Halder G, Callaerts P, Gehring WJ. 1995. Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. Science 267: 1788–1792.
Hoegg S, Brinkmann H, Taylor JS, Meyer A. 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol. 59:190–203.
Holland LZ, Schubert M, Kozmik Z, Holland ND. 1999. AmphiPax3/7, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. Evol Dev. 1:153–165.
Holland PW, Garcia-Fernandez J, Williams NA, Sidow A. 1994. Gene duplications and the origins of vertebrate development. Dev. Sppl. 125–133.
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.
Hubbard TJ, et al. 2009. Ensembl 2009. Nucleic Acids Res. 37:D690–D697.
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.
Hurley IA, et al. 2007. A new time-scale for ray-finned fish evolution. Proc Biol Sci. 274:489–498.
Iwabe N, et al. 2005. Sister group relationship of turtles to the bird-crocodilian clade revealed by nuclear DNA-coded proteins. Mol Biol Evol. 22:810–813.
Kaessmann H. 2010. Origins, evolution, and phenotypic impact of new genes. Genome Res. 20:1313–1326.
Kammandel B, et al. 1999. Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. Dev Biol. 205:79–97.
Kluever N, Herpin A, Braasch I, Driessle J, Schartl M. 2009. Regulatory back-up circuit of medaka Wt1 co-orthologs ensures PGC maintenance. Dev Biol. 325:179–188.
Kuraku S. 2013. Impact of asymmetric gene repertoire between cyclostomes and gnathostomes. Semin Cell Dev Biol. 24:119–127.
Kuraku S, Meyer A. 2012. Detection and phylogenetic assessment of conserved synteny derived from whole genome duplications. Methods Mol Biol. 855:385–395.
Kuraku S, Meyer A, Kuratani S. 2009. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? Mol Biol Evol. 26:47–59.
Kuraku S, Usuda R, Kuratani S. 2005. Comprehensive survey of carapacial ridge-specific genes in turtle implies co-option of some regulatory genes in carapace evolution. Evol Dev. 7:3–17.
Lakowski J, Majumder A, Lauderdale JD. 2007. Mechanisms controlling Pax6 isoform expression in the retina have been conserved between teleosts and mammals. Dev Biol. 307:498–520.
Lundin LG. 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16:1–19.
Lynch M, O'Hely M, Walsh B, Force A. 2001. The probability of preservation of a newly arisen gene duplicate. Genetics 159:1789–1804.
Macdonald R, Wilson SW. 1997. Distribution of Pax6 protein during eye development suggests discrete roles in proliferative and differentiated visual cells. Dev Genes Evol. 206:363–369.
Manousaki T, Feiner N, Begemann G, Meyer A, Kuraku S. 2011. Co-orthology of Pax4 and Pax6 to the fly eyeless gene: molecular phylogenetic, comparative genomic, and embryological analyses. Evol Dev. 13:448–459.
Matus DQ, Pang K, Daly M, Martindale MQ. 2007. Expression of Pax gene family members in the anthozoan cnidarian, *Nematostella vectensis*. Evol Dev. 9:25–38.
Mehta TK, et al. 2013. Evidence for at least six Hox clusters in the Japanese lamprey (*Lethenteron japonicum*). Proc Natl Acad Sci U S A. 110: 16044–16049.
Meyer A, Van de Peer Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 27:937–945.
Mishra R, Gorlov IP, Chao LY, Singh S, Saunders GF. 2002. PAX6, paired domain influences sequence recognition by the homeodomain. J Biol Chem. 277:49488–49494.
Moreno N, Retaux S, Gonzalez A. 2008. Spatio-temporal expression of Pax6 in *Xenopus* forebrain. Brain Res. 1239:92–99.
Mulley JF, Holland PW. 2010. Parallel retention of Pdx2 genes in cartilaginous fish and coelacanths. Mol Biol Evol. 27:2386–2391.
Nakatani Y, Takeda H, Kohara Y, Morishita S. 2007. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. Genome Res. 17:1254–1265.
Nieuwkoop PD, Faber J. 1994. Normal table of *Xenopus laevis* (Daudin). New York: Garland Publishing Inc.
Nornes S, et al. 1998. Zebrafish contains two pax6 genes involved in eye development. Mech Dev. 77:185–196.
Ogasawara M, Wada H, Peters H, Satoh N. 1999. Developmental expression of Pax1/9 genes in urochordate and hemichordate gills: insight into function and evolution of the pharyngeal epithelium. Development 126:2539–2550.

Ohno S. 1970. Evolution by gene duplication. Berlin-Heidelberg-New York: Springer.
Osumi N, Shinohara H, Numayama-Tsuruta K, Maekawa M. 2008. Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator. Stem Cells 26:1663–1672.
Pilz AJ, Povey S, Gruss P, Abbott CM. 1993. Mapping of the human homologs of the murine paired-box-containing genes. Mamm Genome. 4:78–82.
Prince VE, Pickett FB. 2002. Splitting pairs: the diverging fates of duplicated genes. Nat Rev Genet. 3:827–837.
Qiu H, Hildebrand F, Kuraku S, Meyer A. 2011. Unresolved orthology and peculiar coding sequence properties of lamprey genes: the *KCNA* gene family as test case. BMC Genomics 12:325.
Rath MF, et al. 2009. Developmental and daily expression of the Pax4 and Pax6 homeobox genes in the rat retina: localization of Pax4 in photoreceptor cells. J Neurochem. 108:285–294.
Ravi V, et al. 2013. Sequencing of Pax6 loci from the elephant shark reveals a family of Pax6 genes in vertebrate genomes, forged by ancient duplications and divergences. PLoS Genet. 9:e1003177.
Reinecke M. 1981. Immunohistochemical localization of polypeptide hormones in endocrine cells of the digestive tract of *Branchiostoma lanceolatum*. Cell Tissue Res. 219:445–456.
Rest JS, et al. 2003. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. Mol Phylogenet Evol. 29:289–297.
Sanger TJ, Losos JB, Gibson-Brown JJ. 2008. A developmental staging series for the lizard genus *Anolis*: a new system for the integration of evolution, development, and ecology. J Morphol. 269:129–137.
Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502–504.
Shaffer HB, et al. 2013. The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. Genome Biol. 14:R28.
Sheng G, Thouvenot E, Schmucker D, Wilson DS, Desplan C. 1997. Direct regulation of rhodopsin 1 by Pax-6/eyeless in *Drosophila*: evidence for a conserved function in photoreceptors. Genes Dev. 11: 1122–1131.
Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17:1246–1247.
Sosa-Pineda B, Chowdhury K, Torres M, Oliver G, Gruss P. 1997. The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. Nature 386:399–402.
St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, Gruss P. 1997. Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. Nature 387:406–409.
St John JA, et al. 2012. Sequencing three crocodilian genomes to illuminate the evolution of archosaurs and amniotes. Genome Biol. 13:415.
Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 32: W309–W312.
Sun T, Zhang S, Ji G. 2010. Identification and expression of an elastase homologue in *Branchiostoma belcheri* with implications to the origin of vertebrate pancreas. Mol Biol Rep. 37:3303–3309.
Tamura K, et al. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 28:2731–2739.
Tamura T, et al. 1994. Assignment of the human PAX4 gene to chromosome band 7q32 by fluorescence in situ hybridization. Cytogenet Cell Genet. 66:132–134.
Thisse B, Thisse C. 2004. Fast release clones: a high throughput expression analysis [Internet]. ZFIN Direct Data Submission. [cited 2014 Jun 26]. Available from: http://zfin.org.
Turque N, Plaza S, Radvanyi F, Carriere C, Saule S. 1994. Pax-QNR/Pax-6, a paired box- and homeobox-containing gene expressed in neurons, is also expressed in pancreatic endocrine cells. Mol Endocrinol. 8: 929–938.
Underhill DA. 2012. PAX proteins and fables of their reconstruction. Crit Rev Eukaryot Gene Expr. 22:161–177.
Wada H, Saiga H, Satoh N, Holland PW. 1998. Tripartite organization of the ancestral chordate brain and the antiquity of placodes: insights from ascidian Pax-2/5/8, Hox and Otx genes. Development 125: 1113–1122.
Walther C, Gruss P. 1991. Pax-6, a murine paired box gene, is expressed in the developing CNS. Development 113:1435–1449.
Wang Z, et al. 2013. The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nat Genet. 45:701–706.
Wehr R, Gruss P. 1996. Pax and vertebrate development. Int J Dev Biol. 40: 369–377.
Wittbrodt J, Meyer A, Schartl M. 1998. More genes in fish? Bioessays 20: 511–515.
Wullimann MF, Rink E. 2001. Detailed immunohistology of Pax6 protein and tyrosine hydroxylase in the early zebrafish brain suggests role of Pax6 gene in development of dopaminergic diencephalic neurons. Brain Res Dev Brain Res. 131:173–191.
Zardoya R, Meyer A. 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. Proc Natl Acad Sci U S A. 95:14226–14231.
Zhang Y, Emmons SW. 1995. Specification of sense-organ identity by a *Caenorhabditis elegans Pax-6* homologue. Nature 377:55–59.

Associate editor: Soojin Yi