# Asymmetric Evolution in Two Fish-Specifically Duplicated Receptor Tyrosine Kinase Paralogons Involved in Teleost Coloration

*Ingo Braasch,*[1] Walter Salzburger,*† and Axel Meyer**

*Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Konstanz, Germany; and †Center for Junior Research Fellows, University of Konstanz, Konstanz, Germany

The occurrence of a fish-specific genome duplication (FSGD) in the lineage leading to teleost fishes is widely accepted, but the consequences of this event remain elusive. Teleosts, and the cichlid fishes from the species flocks in the East African Great Lakes in particular, evolved a unique complexity and diversity of body coloration and color patterning. Several genes involved in pigment cell development have been retained in duplicate copies in the teleost genome after the FSGD. Here we investigate the evolutionary fate of one of these genes, the type III receptor tyrosine kinase (RTK) *colony-stimulating factor 1 receptor* (*csf1r*). We isolated and shotgun sequenced two paralogous *csf1r* genes from a bacterial artificial chromosome library of the cichlid fish *Astatotilapia burtoni* that are both linked to paralogs of the *pdgfrβ* gene, another type III RTK. Two *pdgfrβ-csf1r* paralogons were also identified in the genomes of pufferfishes and medaka, and our phylogenetic analyses suggest that the *pdgfrβ-csf1r* locus was duplicated during the course of the FSGD. Comparisons of teleosts and tetrapods suggest asymmetrical divergence at different levels of genomic organization between the teleost-specific *pdgfrβ-csf1r* paralogons, which seem to have evolved as coevolutionary units. The high-evolutionary rate in the teleost B-paralogon, consisting of *csf1rb* and *pdgfrβb*, further suggests neofunctionalization by functional divergence of the extracellular, ligand-binding region of these cell-surface receptors. Finally, we hypothesize that genome duplications and the associated expansion of the RTK family might be causally linked to the evolution of coloration in vertebrates and teleost fishes in particular.

## Introduction

Gene and genome duplications are considered major evolutionary forces adding to the genetic material that may be necessary for increasing the genomic and phenotypic complexity of vertebrates (Ohno 1970; Holland et al. 1994). While the occurrence of two rounds of whole-genome duplication early in the vertebrate lineage (2R hypothesis) is still debated (Holland et al. 1994; Spring 1997; Furlong and Holland 2002; Larhammar, Lundin, and Hallbook 2002; McLysaght, Hokamp, and Wolfe 2002; Hughes and Friedman 2003; Dehal and Boore 2005), an additional round of fish-specific genome duplication (FSGD) in the lineage of ray-finned fishes (3R-hypothesis) (Amores et al. 1998; Meyer and Schartl 1999) is now widely accepted (Postlethwait et al. 2004; Meyer and Van de Peer 2005; Volff 2005). Strong evidence for the FSGD comes from the combination of synteny data and genome-wide phylogenetic reconstructions of gene duplicates in zebrafish, pufferfishes, medaka, and other teleost fishes (Postlethwait et al. 2000; Taylor et al. 2001, 2003; Christoffels et al. 2004; Jaillon et al. 2004; Naruse et al. 2004; Vandepoele et al. 2004; Woods et al. 2005). These studies also demonstrated that many fish-specifically duplicated genes are organized in paralogous chromosomal blocks that contain two or more duplicated genes, so-called paralogons. The multiple Hox clusters in teleosts that arose from a single cluster by successive duplication in the course of the three whole-genome duplications are the most prominent examples of such paralogons (reviewed in Hoegg and Meyer 2005). While global analyses of teleost genomes mainly described syntenic relationships of chromosomes within and between genomes (Jaillon et al. 2004; Naruse et al. 2004; Woods et al. 2005), detailed studies on postduplication evolution of paralogons on different genomic levels are sparse and restricted to Hox clusters (e.g., Santini, Boore, and Meyer 2003; Wagner et al. 2005). Important questions are if duplicated genes coevolve with their genomic environment and whether or not paralogons evolve at similar rates after duplication. Insights from Hox clusters, however, might not be representative of global genomic patterns due to their exceptional organized structure. Thus, fine-scale genomic analyses of other paralogons are required.

Here, we present a comparative genomic study of two *pgdfrβ-colony-stimulating factor 1 receptor* (*csf1r*) paralogons in the teleost lineage. These genes as well as the anciently paralogous gene tandem *pdgfrα-kit* belong to the family of type III receptor tyrosine kinases (RTKs) (Rousset et al. 1995). We hypothesized that the type III RTK tandems represent an excellent setup for the study of the genomic consequences of large-scale duplication events in general and the FSGD in particular. Specifically, we were interested in the evolutionary fates of the two *pgdfrβ-csf1r* paralogons after the FSGD with respect to evolutionary rates, exon/intron structure, synteny, genomic expansions/compactions, and the degree of conservation in noncoding regions. The reasons for choosing the two *pgdfrβ-csf1r* paralogons were: (i) all four gene copies of the tandem, and the gene order within each paralogon, have been retained since the FSGD; (ii) type III RTK paralogons are linked to the ParaHox clusters, which contain developmentally important genes and which arose by the same succession of genome duplications in vertebrates (Spring 2002; Ferrier et al. 2005); (iii) comparative genomic data from other organisms, particularly from other teleost species, are available; and (iv) the genes have been well characterized, i.e., functional data are available for zebrafish and mouse demonstrating that members of the type III RTKs are highly interesting with respect to the evolution of pigmentation in the vertebrate lineage. For

[1] Present address: BioFuture Research Group ''Evolutionary Fish Genomics'', Physiological Chemistry I, Biozentrum, University of Würzburg, Germany.

Key words: genome duplication, fish, pigmentation, *csf1r*, *pdgfrβ*, cichlid.

E-mail: axel.meyer@uni-konstanz.de.

*Mol. Biol. Evol.* 23(6):1192–1202. 2006
doi:10.1093/molbev/msk003
Advance Access publication March 17, 2006

example, *kit* and *csf1r* have similar functions for the development of different neural crest-derived pigment cell types. While *kit* is essential for the development of the dark melanocytes in mammals and zebrafish (Parichy et al. 1999), *csf1r* promotes the development of yellow xanthophores in zebrafish (Parichy et al. 2000; Parichy and Turner 2003). In view of several examples of other duplicated "pigmentation genes" in teleosts such as *mitf* (Altschmied et al. 2002), *sox10* (Lang et al. 2005*b*), and *tyrosinase* (Hoegg et al. 2004), the fact that teleosts possess more pigment cell types than tetrapods (Bagnara 1998), the unique diversity of coloration and color patterning in teleosts, and the importance of pigmentation for speciation in teleosts as signaling cue for mate and species recognition (e.g., Salzburger et al. 2006), we also investigated the *pdgfrβ-csf1r* locus with regard to pigmentation. We chose a cich-lid fish for this study, as cichlids are well-suited models to examine teleost coloration within an evolutionary framework. For example, sexual selection on diverse male color patterns is considered to be a major factor promoting the explosive speciation and adaptive radiation in East African cichlids (Kocher 2004; Salzburger and Meyer 2004).

We isolated and shotgun sequenced two clones representing different *pdgfrβ-csf1r* paralogons from a bacterial artificial chromosome (BAC) library of the haplochromine cichlid *Astatotilapia burtoni* (Lang et al. 2005*b*). We identified two *pdgfrβ-csf1r* paralogons in the genomes of medaka and pufferfishes, and phylogenetic reconstructions led us to conclude that the *pdgfrβ-csf1r* locus was duplicated during the FSGD. We found pronounced differences in rates of molecular evolution between the two *pdgfrβ-csf1r* paralogons prior to the divergence of cichlid, medaka, and pufferfishes at different levels of their genomic organization. Our results suggest asymmetrical coevolution of these paralogons resulting, most likely, in the evolution of new gene functions (neofunctionalization) in the entire B-paralogon (*pdgfrβb* and *csf1rb*). Furthermore, our results provide implications for the evolution of teleost coloration.

## Materials and Methods

### Identification of *csf1r* Genes in the Genomes of Cichlid and Other Teleosts

Based on the alignment of all available teleost *csf1r* sequences, we designed degenerate primer sets (Supplementary Table S1, Supplementary Material online) to amplify fragments of *csf1r* paralogs from genomic DNA of the cichlid *A. burtoni* as well as from other teleost fishes.

### BAC Library Screening

Fragments of *csf1r* paralogs were polymerase chain reaction (PCR) amplified from genomic DNA of *A. burtoni* using specific primer sets for *csf1ra* (Fms-F-2961/Fms-R-3199 and Burt-Fms-F-1986/Burt-Fms-R-2186; Supplementary Table S1, Supplementary Material online) and specific PCR conditions for *csf1rb* (Fms2-F-1962/Fms2-R-2269; annealing temperature, 66°C; Supplementary Table S1, Supplementary Material online). Purified PCR products were labeled with the NEBlot Phototope Kit (NEB). A BAC library of *A. burtoni* (Lang et al. 2005*b*) was then screened by hybridization of high-density filters with the probes according to the Phototope-Star Detection Kit (NEB). Positive clones were further characterized by PCR screening, sequencing, and restriction fragment patterning. Insert sizes were determined by *Not*I digestions.

### Shotgun Sequencing and Contig Assembly

BAC DNA was obtained using the QIAGEN Large Construct Kit and sheared by sonication. The size fraction of 2–3 kb was blunt-end ligated into pUC18 vector (Roche) and electrotransformed into ElectroMAX DH10B T1 phage resistant cells (Invitrogen, Karlsruhe, Germany). Subclones were grown in selective medium; plasmid DNA was recovered by standard methods and sequenced directly from both ends with M13 standard primers on an ABI3100 sequencer (Applied Biosystems, Darmstadt, Germany). Sequences were quality-trimmed with PHRED (http://www.phrap.org). The Sequencher software (Gene Codes) was used to trim vector ends and for sequence assembly. Large-scale sequencing was performed until coverage of approximately 4× was obtained. Remaining gaps were closed with gap spanning primer sets so that we finally obtained a continuous sequence assembly with each base pair sequenced at least twice.

### Sequence Annotation

Genomic structures of type III RTKs in cichlid were determined by sequence comparisons with the pufferfish *Takifugu rubripes* (Williams, Brenner, and Venkatesh 2002) and through cDNA sequencing. Additional genes were predicted with GENSCAN (http://genes.mit.edu/GENSCAN.html) and repetitive elements were identified with RepeatMasker (http://www.repeatmasker.org/). We found orthologous loci in the genomes of the spotted green pufferfish (*Tetraodon nigroviridis*) (http://www.genoscope.cns.fr/externe/tetranew/) and medaka (*Oryzias latipes*) (http://dolphin.lab.nig.ac.jp/medaka/) by BLAT/Blast searches and deduced the sequences of their *csf1r* and *pdgfrβ* paralogs by sequence comparisons to *A. burtoni* and *Takifugu* under inclusion of available express sequence tag sequences. The *csf1ra*-locus of zebrafish could not be included in genomic analyses due to an apparent misassembly of that region in the current zebrafish genome release (Zv4; http://www.ensembl.org/Danio_rerio/).

### Sequence Analyses

Phylogenetic reconstructions of type III RTKs were performed with PAUP* (Swofford 2000) using the maximum likelihood method (general time-reversible model, with a gamma substitution correction and a proportion of invariable sites; model parameters estimated from the data). For relative rate analyses, we applied the nonparametric rate test as implemented in MEGA3 (Kumar, Tamura, and Nei 2004) and compared paralogs of *csf1r* as well as *pdgfrβ* from teleosts with their murine ortholog as described in Van de Peer et al. (2001). Additionally, we performed local molecular clock tests with HyPhy (Pond, Frost, and Muse 2005) using the GY94 codon model. To compare the BAC contigs of cichlid with those of medaka, pufferfishes, frog, chicken, mouse, and human, we downloaded the respective genomic sequences (Supplementary Table S2,

Supplementary Material online) and aligned them applying the MLAGAN algorithm in mVISTA (http://genome.lbl.gov/vista/index.shtml). Conserved noncoding sequences (CNSs) were then analyzed using the MatInspector software (http://www.genomatix.de/) for the presence of putative binding sites for transcription factors that might be involved in pigment cell development such as Wnt (Dorsky, Moon, and Raible 1998), Sox10 (Dutton et al. 2001), Tfap2a (Knight et al. 2003), Pax3 (Lang et al. 2005*a*), and Mitf (Béjar, Hong, and Schartl 2003). We used the DIVERGE software (Gu and Vander Velden 2002) to identify amino acid residues that are likely to be involved in functional divergence after the duplication of Csf1r and Pdgfrβ paralogs. Based on neighbor-joining trees constructed with DIVERGE, paralogous clades were selected. Posterior probabilities of functional divergence were estimated for each alignment position and assigned to the protein domain structure as obtained from PFAM (http://www.sanger.ac.uk/Software/Pfam/).

### Analysis of Gene Expression

Total mRNA from various tissues of adult male and female of *A. burtoni* were extracted with TRIZOL LS (Invitrogen, Karlsruhe, Germany). First strand cDNA synthesis was mediated by SUPERSCRIPT II reverse transcriptase (Invitrogen, Karlsruhe, Germany). Reverse transcriptase-PCR (RT-PCR) experiments were performed with specific primer sets for *csf1ra*, *csf1rb*, *pdgfrβa*, and *pdgfrβb* (Supplementary Table S1, Supplementary Material online). RT-PCRs targeting *gapdh* were used as a control for the quality of the cDNA. All primers spanned at least one intron.

## Results

### Presence of Two *pdgfrβ-csf1r* Loci in the Cichlid Genome

As two *csf1r* paralogs were found previously in the pufferfish *T. rubripes* (Williams, Brenner, and Venkatesh 2002), but only a single copy of *csf1r* has been identified so far in zebrafish (Mellgren and Johnson 2002), we were interested in the number of *csf1r* paralogs in the cichlid lineage. With a particular degenerate primer set spanning three introns (Fms2-F-1962/Fms2-R-2269; Supplementary Table S1, Supplementary Material online), two fragments of different lengths were amplified from genomic DNA of the haplochromine cichlid *A. burtoni*. Sequencing of the two fragments and phylogenetic reconstruction confirmed that these were indeed two different copies of *csf1r* termed *csf1ra* and *csf1rb* in the following. The length difference of the fragments was due to variation in intron lengths. By screening a BAC library of *A. burtoni* (Lang et al. 2005*b*) separately for both *csf1r* copies, we identified two nonoverlapping populations of BAC clones from different genomic regions containing either *csf1ra* or *csf1rb*. After shotgun sequencing of representative BAC clones and sequence assembly, we obtained two contigs representing the "A-paralogon" containing *csf1ra* (BAC 26-M7; 96.6 kb) and the "B-paralogon" containing *csf1rb* (BAC 20-D21; 144.6 kb) (see also Supplementary Table S2, Supplementary Material online). These contigs (GenBank accession numbers DQ386647–8) represent the first published genomic sequences of any significant length of a haplochromine cichlid.

### Duplication of the *pdgfrβ-csf1r* Locus in Teleost Fish

Both *csf1r* paralogs are located downstream of *pdgfrβ* genes, termed *pdgfrβa* and *pdgfrβb* in the following, just as in *Takifugu* and in the single locus of human and mouse (Williams, Brenner, and Venkatesh 2002). A maximum likelihood phylogeny of type III RTKs including representative members of this gene family confirms that the obtained RTKs are duplicated copies of *csf1r* and *pdgfrβ* (fig. 1*A*). The topology of the tree further suggests that the duplications of *csf1r* and *pdgfrβ* in cichlid and pufferfish are the results of the FSGD (fig. 1*A*).

We also found two *pdgfrβ-csf1r* loci in the current genome assemblies of the spotted green pufferfish and of medaka, but not in zebrafish, and obtained additional *csf1r* sequences from other teleosts from GenBank and using our degenerate primer sets for *csf1r* (Supplementary Tables 1–3, Supplementary Material online). The phylogenies of *csf1r* (fig. 1*B*) and *pdgfrβ* genes (fig. 1*C*) are consistent with a duplication of these genes in the course of the FSGD.

### Faster Evolution of B-Paralogon than of A-Paralogon RTKs

Nonparametric relative rate tests with MEGA revealed that both type III RTK genes from the B-paralogons evolved with a significantly faster rate of molecular evolution than their paralogs from the A-paralogons, both at the nucleotide level as well as at the protein level (table 1). Local molecular clock tests performed with HyPhy confirmed a faster rate in B-paralogon RTKs ($P < 0.001$) but suggested that the synonymous mutation rate is similar between A- and B-paralogons. Thus, it appears that not the neutral mutation rate but the selection pressure is differing between the two paralogons.

### Intron Accumulation in B-Paralogon RTKs

In cichlid, medaka, and *Tetraodon*, *pdgfrβb* and *csf1rb* have additional introns compared to the A-paralogon genes. *pdgfrβa* and *csf1ra* both contain 21 exons, while *pdgfrβb* and *csf1rb* have 26 and 22 exons, respectively (Supplementary Figure S1, Supplementary Material online). An identical situation has been reported for *Takifugu* (Williams, Brenner, and Venkatesh 2002). Furthermore, the teleost *csf1ra* genes have the same exon-intron structure as their mammalian semiortholog (as inferred from human and mouse). The exon-intron structure of the teleost *pdgfrβb* genes is also highly similar to the situation found in mammals, where *Pdgfrβ* consists of 22 coding exons (data not shown). These results suggest that the type III RTKs of the A-paralogon most likely resemble the ancestral exon-intron structure and that introns have accumulated preferentially in the B-paralogs after the FSGD but before the split of the four teleost species.

### Less Synteny in the B-Paralogon with the Single Tetrapod Locus

In the four teleost A-paralogons, several genes were predicted in the surrounding of the *pdgfrβa-csf1ra* gene tandem (fig. 2). In all B-paralogons, an *enteropeptidase* gene

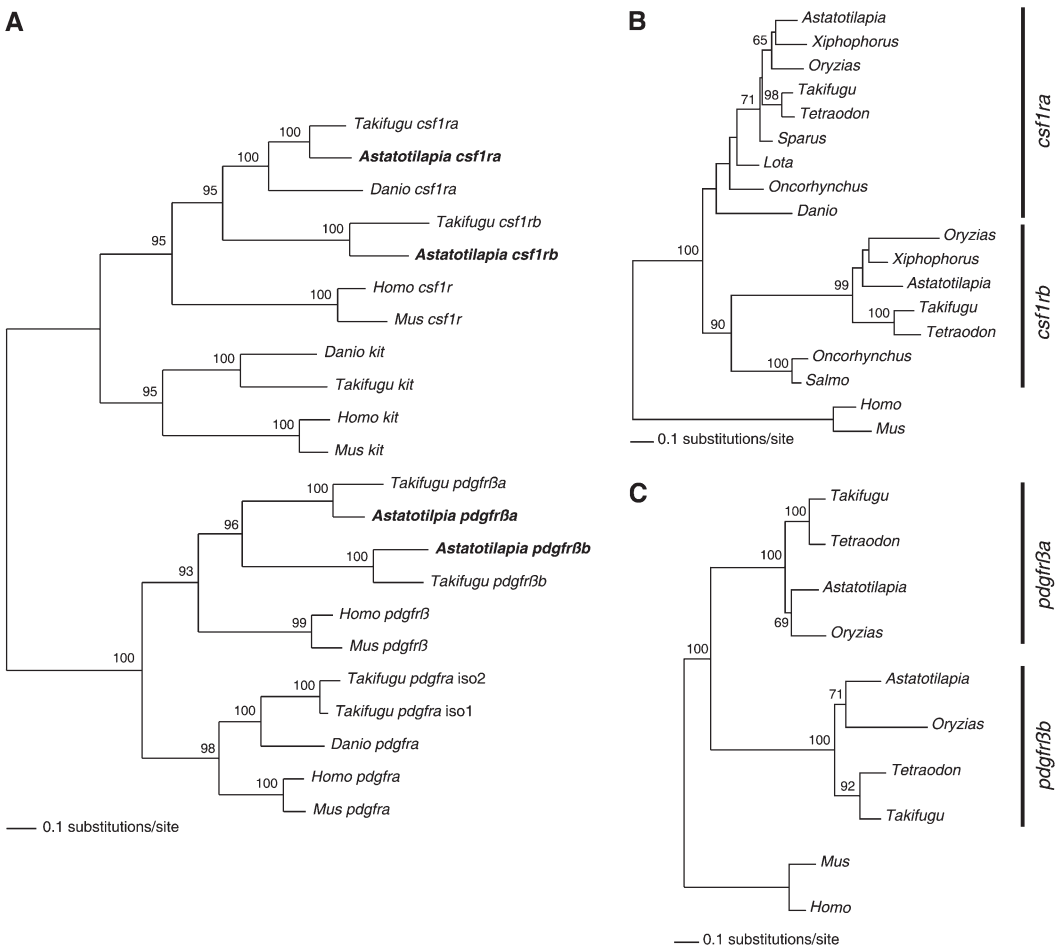


FIG. 1.—Maximum likelihood phylogenies of type III RTKs. (*A*) Phylogeny of vertebrate type III RTKs (unrooted). Phylogenies of *csf1r* (*B*) and *pdgfrβ* (*C*) genes in teleosts rooted with mouse (*Mus*) and human (*Homo*) sequences. Numbers on the branches represent percentages of 100 bootstrap replicates; bootstrap values below 50 are not shown. The duplication of *csf1r* and *pdgfrβ* in the common ancestor of all species under examination is supported with high-bootstrap values and coincides with the FSGD. See Supplementary Table S3 (Supplementary Material online) for accession numbers.

was predicted downstream of *csf1rb*. The comparison of the teleost paralogons to the single tetrapod loci shows that the teleost A-paralogon has retained a higher level of synteny with its tetrapod semiorthologs (fig. 2). The array of the four genes *cdx1-pdgfrβ-csf1r*-A630042L21RIK was found for the tetrapod loci and for all teleost A-paralogons. In contrast, the synteny of the teleost B-paralogons with the tetrapod loci is restricted to the tandem array of *pdgfrβ-csf1r*. Thus, it seems that more genomic rearrangements have occurred in the B-paralogon after duplication in teleosts, while the A-paralogon has retained more similarity with the ancestral region. Furthermore, in *Tetraodon*, the *pdgfrβ-csf1r* paralogons are located on chromosome 1 and chromosome 7 (fig. 2). These two chromosomes are highly syntenic, containing over 25 paralogous gene pairs (Jaillon et al. 2004; Woods et al. 2005), which provides further evidence for the duplication of the *pdgfrβ-csf1r* in the course of the FSGD.

## Preferential Accumulation of Repetitive Elements in the B-Paralogon

By comparing both paralogons of *A. burtoni*, it became obvious that the B-paralogon is substantially larger compared to the A-paralogon, which is likely to result from the accumulation of repetitive elements in the B-paralogon (18.6% vs. 9.2% of their length is made up of repetitive elements; Supplementary Table S2, Supplementary Material online). Accumulation of repetitive elements particularly in the B-paralogon and an increase in length was also found in medaka but not in pufferfishes.

## Less Conservation of Noncoding Regions in the B-Paralogon

Phylogenetic footprinting using VISTA revealed no sequence similarity in noncoding regions between both cichlid paralogons (not shown). Yet, no differences in gene expression among duplicated type III RTKs were apparent in cichlids (Supplementary Figure S2, Supplementary Material online). Next, we compared each cichlid paralogon individually with its orthologous and semiorthologous loci in other fish species and tetrapods, respectively (fig. 3). There was virtually no sequence similarity detected in noncoding regions between fishes and tetrapods for both paralogons. Interestingly, the overall sequence similarity among fishes was higher in the A-paralogon than in the B-paralogon

**Table 1**
**Nonparametric Relative Rate Tests of Teleost Type III RTKs**

| | | Nucleotide Sequence (first and second codon position) | | | Amino Acid Sequence | | |
|---|---|---|---|---|---|---|---|
| Species | Gene[a] | Sites | Unique Differences | Significance[b] | Sites | Unique Differences | Significance[b] |
| *Astatotilapia* | *csf1ra* | 1854 | 190 | * | 926 | 53 | ** |
| | ***csf1rb*** | | 236 | | | 94 | |
| | *pdgfrβa* | 2031 | 226 | * | 1014 | 69 | *** |
| | ***pdgfrβb*** | | 298 | | | 150 | |
| *Oryzias* | *csf1ra* | 1776 | 198 | *** | 886 | 65 | ** |
| | ***csf1rb*** | | 277 | | | 111 | |
| | *pdgfrβa* | 1970 | 243 | *** | 988 | 89 | *** |
| | ***pdgfrβb*** | | 387 | | | 191 | |
| *Takifugu* | *csf1ra* | 1848 | 180 | *** | 923 | 61 | *** |
| | ***csf1rb*** | | 269 | | | 137 | |
| | *pdgfrβa* | 1956 | 235 | ** | 978 | 96 | *** |
| | ***pdgfrβb*** | | 308 | | | 166 | |
| *Tetraodon* | *csf1ra* | 1842 | 199 | *** | 919 | 55 | *** |
| | ***csf1rb*** | | 280 | | | 113 | |
| | *pdgfrβa* | 1969 | 217 | *** | 983 | 88 | *** |
| | ***pdgfrβb*** | | 299 | | | 151 | |

[a] Genes that show statistically significant increase in rate of molecular evolution are indicated in bold.
[b] $\chi^2$ tests: $*P < 0.05$, $**P < 0.01$, $*** P < 0.001$.

in coding as well as in noncoding regions (fig. 3). Furthermore, we identified several CNSs in the surrounding of the *pdgfrβ-csf1r* tandems with potential function in pigment cell transcription as they contain putative binding sites for transcription factors that are involved in pigment cell development (fig. 3).

## Functional Divergence Between *csf1r* and *pdgfrβ* Paralogs

One consequence of functional divergence after duplication is that some residues might be subject to altered functional constraints. This would lead to differences in evolutionary rates at these sites between paralogous proteins, i.e., type I functional divergence (Gu 1999).

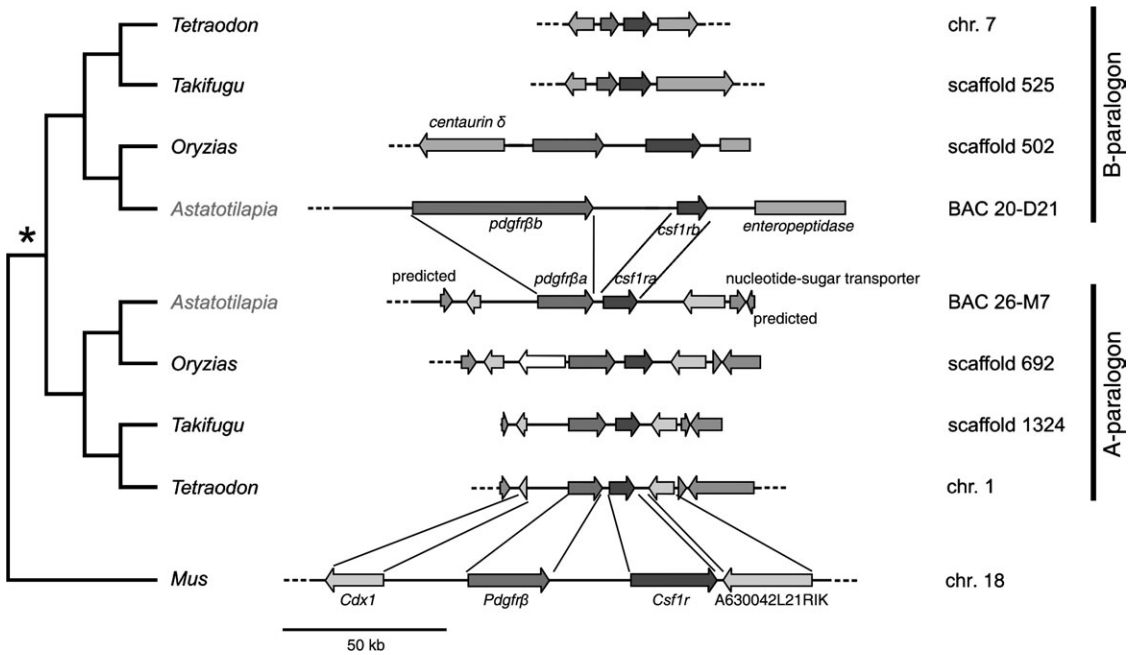


FIG. 2.—Genomic organization and synteny of *pdgfrβ-csf1r* paralogons in teleosts. The tree on the left illustrates the phylogenetic relationships, and a star indicates the FSGD leading to two paralogous loci in teleosts. The single locus of tetrapods is represented by the mouse. Genes that prolong synteny between teleost A-paralogons or B-paralogons are found up- and downstream of the *pdgfrβ-csf1r* tandems. The position of each locus within the respective genome assembly is given on the right. The teleost A-paralogon has retained a longer stretch of synteny with the tetrapod locus compared to the B-paralogon. The B-paralogon is more variable in size among teleosts than the A-paralogon and is expanded in cichlid (*Astatotilapia*) and medaka (*Oryzias*).

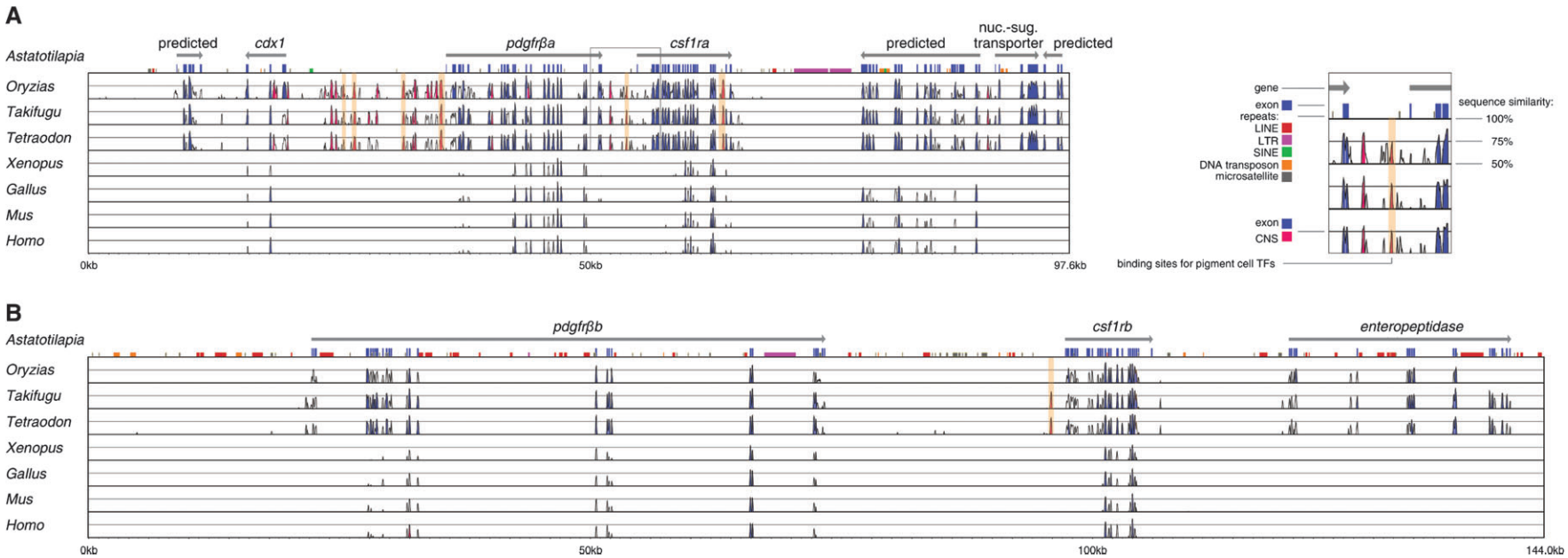


FIG. 3.—mVista plots of vertebrate *pdgfrβ-csf1r* paralogons. Cichlid paralogons were set as references. (*A*) A-paralogon (BAC 26-M7). (*B*) B-paralogon (BAC 20-D21). The intergenic region between *pdgfrβa* and *csf1ra* is used as figure legend. The B-paralogon is enlarged compared to the A-paralogon and has accumulated repetitive elements. Within teleosts, more CNSs are found in the A-paralogon compared to the B-paralogon, particularly upstream of *pdgfrβa*. Some of the CNSs might be responsible for the regulation of gene expression in pigment cells, as they represent putative binding sites for transcription factors (TFs) involved in pigment cell development.

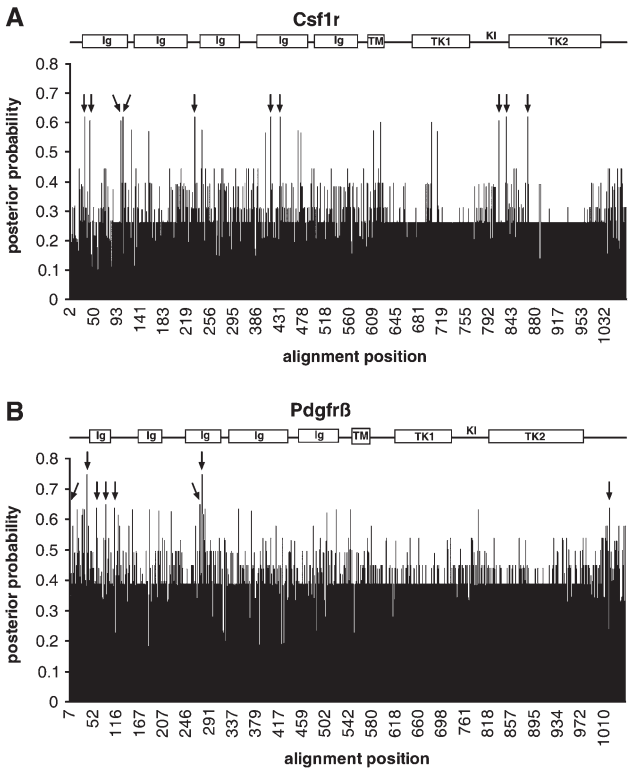


FIG. 4.—Site-specific profile of functional divergence between teleost type III RTK paralogons measured using DIVERGE. Arrows point to critical amino acid residues responsible for the functional divergence between Csf1r (*A*) and Pdgfrβ (*B*) paralogs after duplication in teleosts. When these amino acid residues are removed from the multiple alignment, the coefficient of type I functional divergence θ is no longer significantly greater than zero. These amino acid residues are mainly located in extracellular parts. IG, immunoglobulin-like domain; TM, transmembrane domain; TK1 and TK2, tyrosine kinase catalytic domains; KI, kinase insert domain.

Gu (1999) developed a method to detect site-specific altered functional constraints (or shifted evolutionary rates) by comparing the rate correlation between clades (e.g., A- vs. B-paralogs) under a given phylogeny. Using the DIVERGE software, we determined the coefficient of functional divergence (θ), which is defined as the probability that the rates in two clades are different enough between sites to be statistically independent (Gu and Vander Velden 2002). Rejection of the null hypothesis $\theta = 0$ suggests a significant change of evolutionary rates at some sites between the two clades (Gu and Vander Velden 2002). Our analyses revealed that functional divergence (i.e., $\theta > 0$) has occurred in both type III RTKs after duplication with values of $\theta_{Csf1r} = 0.298 \pm 0.086$ ($P < 0.01$) and $\theta_{Pdgfr\beta} = 0.426 \pm 0.110$ ($P < 0.001$). The posterior probabilities of being functional divergence related for each site in the alignments are depicted in figure 4*A* for Csf1r and in figure 4*B* for Pdgfrb. We determined individual cutoff values for significance for both type III RTKs by removing residues with the highest posterior probability until $\theta > 0$ was no longer significant (Csf1r: 0.607; Pdgfrβ: 0.636). Residues above this value are likely to contribute significantly to functional divergence between paralogs (Wang and Gu 2001). Ten residues were above the cutoff value in Csf1r and eight in Pdgfrβ. A comparison with the protein structure revealed that most of these residues are located in the extracellular part of the type III RTKs (fig. 4), which are defined by an extracellular ligand-binding domain consisting of five immunoglobulin-like domains and an intracellular part comprising the tyrosine kinase domain split by the kinase insert domain (Rousset et al. 1995).

## Discussion

### Genome Duplications and the Expansion of the RTK Multigene Family

The multigene family of type III RTKs has been expanded substantially by gene and especially genome duplications during vertebrate evolution (Rousset et al. 1995; J. Gu and X. Gu 2003; Leveugle et al. 2004), which is confirmed for teleost fishes by our study. We conclude that an ancestral type III RTK was duplicated in tandem to give rise to the proto-*pdgfr*–proto-*kit*/*csf1r* locus early in the vertebrate lineage (fig. 5*A*). This locus was then duplicated resulting in the *pdgfrα-kit* and the *pdgfrβ-csf1r* locus found in gnathostome genomes (see also Rousset et al. 1995; Leveugle et al. 2004), most likely during the second round of vertebrate genome duplication (2R). The FSGD (3R), finally, resulted in the presence of two *pdgfrβ-csf1r* paralogons in teleosts. However, we only found a single *csf1r* gene in the genome of zebrafish in the current genome assembly, as well as with our degenerate primer based approach, suggesting that the *csf1rb* gene was lost secondarily in the lineage leading to zebrafish. Recently, also a second *kit* gene locus resulting from the FSGD has been identified in zebrafish and pufferfishes (Mellgren and Johnson 2005) as would be predicted from our model.

### Asymmetric Coevolution in *pdgfrβ-csf1r* Paralogons

After the duplication in the fish lineage, the *pdgfrβ-csf1r* paralogons evolved at different rates of molecular evolution. The entire B-paralogon evolved faster than the A-paralogon, which becomes obvious at five levels of genomic organization:

*(i)* *pdgfrβb* and *csf1rb* coding sequences evolve significantly faster than *pdgfrβa* and *csf1ra*, respectively, as revealed by nonparametric relative rate tests in cichlids and the other three teleost genomes. The phylogenetic tree of *csf1r* genes (fig. 1*B*) suggests that the rate acceleration of molecular evolution in *csf1rb* appeared in the lineage leading to acanthopterygians (represented here by cichlid, pufferfishes, medaka, and swordtail) after its split from the lineage leading to protacanthopterygians (salmon and trout). However, discrimination between the mechanisms that might lead to accelerated evolution, relaxed functional constraints, or positive Darwinian selection, is rather difficult due to more than 200 Myr that have passed after the duplication and due to the episodic nature of duplicates' sequence evolution over several hundreds of millions of years (Raes and Van de Peer 2003).

*(ii)* Type III RTKs of the teleost B-paralogon gained additional introns after duplication, while A-paralogon type III RTKs maintained the ancestral exon-intron structure.

*(iii)* More genomic rearrangements have occurred in the chromosomal region of the B-paralogon type III

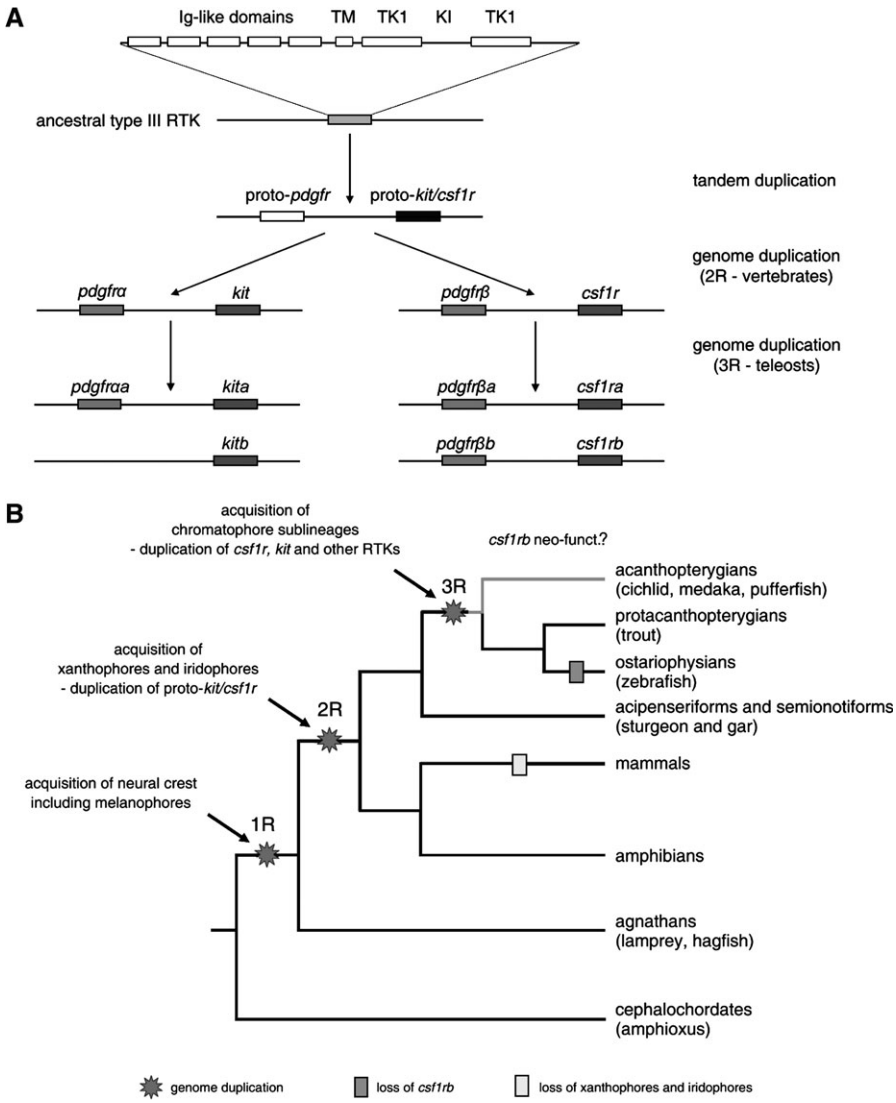


FIG. 5.—Type III RTK duplications and the evolution of coloration in vertebrates. (*A*) Model for the expansion of vertebrate type III RTK multigene family by gene and genome duplications. In teleost fishes, two *pdgfrβ-csf1r* paralogons are found as a result of the FSGD (3R). Similarly, two paralogous *kit* loci are found in teleosts (Mellgren and Johnson 2005). Abbreviations used for protein structure are as in figure 4. (*B*) Association of pigment cell innovations, gene and genome duplications, and the expansion of the type III RTK multigene family during vertebrate and teleost evolution (adapted and expanded from Mellgren and Johnson 2002). Teleost phylogeny is in congruence with recent analyses of our group (Steinke, Salzburger, and Meyer 2006). Two genome duplications (1R, 2R) are predicted to have occurred before the split of fish and tetrapods, a third genome duplication (3R) has occurred in the lineage leading to teleost fishes. These genome duplications were responsible for the expansion of the type III RTK multigene family. Neofunctionalization of *csf1rb* might have occurred in the lineage leading to acanthopterygians (grey branch), and *csf1rb* was probably lost in the lineage leading to zebrafish. Genome duplications and the associated duplication of type III RTKs are predicted to coincide phylogenetically with pigment cell innovations, such as the acquisition of melanophores and later on of xanthophores and iridophores. The latter pigment cell types were lost secondarily in the lineage leading to mammals. Chromatophore sublineages might have been present in the fish lineage ever since the FSGD.

RTKs, so that its synteny with the ancestral locus is restricted to the *pdgfrβb-csf1rb* gene tandem. In contrast, the synteny of the A-paralogon involves at least two additional flanking genes including *cdx1*, which links the A-paralogon to the ParaHox D cluster (Spring 2002; Ferrier et al. 2005).

*(iv)* The B-paralogons of cichlid and medaka have accumulated more repetitive elements than the A-paralogons leading to a pronounced genomic expansion of the B-paralogon. The absence of this situation in the pufferfish genomes might be explained by the general compaction of their genomes (Brenner et al. 1993).

*(v)* Noncoding regions and consequently promoter sequences are less conserved in the B-paralogon. The proportion of gene duplicates in eukaryotic genomes that evolve asymmetrically has been determined previously to be 5%–60%, depending on the age of the duplication event (Conant and Wagner 2003; Wagner et al. 2005). Van de Peer et al. (2001) found asymmetric evolution in approximately half of FSGD duplicates in zebrafish (Van de Peer et al. 2001). The situation of the *pdgfrβ-csf1r* paralogons, however, is exceptional because it involves tightly linked genes and also includes their genomic environment. The *kitb*

locus also underwent more genomic rearrangements than the *kita* locus in teleosts, including the loss of the *pdgfrαb* gene (Mellgren and Johnson 2005). Similar to our analyses, the duplicated teleost HoxA clusters were recently characterized as "coevolutionary units" that diverge concertedly in an asymmetric manner in coding and noncoding regions (Wagner et al. 2005). Hence, these results suggest that asymmetric divergence of fish-specific paralogons might be a rather common evolutionary pattern and that coevolutionary genomic units might not be restricted to the highly organized Hox clusters or to transcription factors in general.

### Functional Divergence of Type III RTK Paralogs in Teleosts After Duplication

All four type III RTKs originating from two rounds of genome duplication (2R and 3R/FSGD) are still present at least in the genomes of cichlid, medaka, and pufferfishes. Thus, nonfunctionalization, the loss of a gene duplicate, seems to be a rather uncommon mechanism for type III RTKs, although it might have occurred in the zebrafish *pdgfrβb-csf1rb* paralogon. Furthermore, all four type III RTKs are expressed in a variety of adult cichlid tissues (Supplementary Figure S2, Supplementary Material online).

An important mechanism that can preserve both paralogs after duplication in a genome is subfunctionalization, the subdivision of ancestral gene functions among paralogs (Postlethwait et al. 2004). However, all four type III RTKs were expressed nearly ubiquitously in adult cichlid (Supplementary Figure S2, Supplementary Material online), making expressional subfunctionalization in adult tissues rather unlikely. The similar expression of type III RTKs is surprising because of the pronounced sequence divergence in promoter regions between A- and B-paralogons within the cichlid genome. In pufferfish, small differences in gene expression between the type III RTK paralogs were found, but still similarities prevailed (Williams, Brenner, and Venkatesh 2002). Nevertheless, it is possible that differences in cellular, temporal, and quantitative gene expression might exist between the four type III RTK genes. Subfunctionalization can also be achieved at the protein level (Hughes 1994), but subdivision of protein functions would probably lead to more similar evolutionary rates of duplicates as observed here for the two type III RTK paralogons.

Several lines of evidence seem to suggest that neofunctionalization, the evolution of new gene functions, has occurred in B-paralogs of *csf1r* and *pdgfrβ*. First, the observed asymmetry of evolutionary rates between the two paralogons and particularly the amount of amino acid exchanges suggest neofunctionalization. Second, *csf1ra* is required by macrophages and osteoclasts in teleosts, a function that is shared with the single *Csf1r* in amniotes (Parichy et al. 2000; Roca et al. 2005) and, therefore, is likely to be ancestral. Furthermore, the *csf1ra* gene is essential to recruit pigment cell precursors to xanthophore fates in zebrafish (Parichy et al. 2000; Parichy and Turner 2003). Recently, we found that *csf1ra* is also expressed in xanthophores in cichlids (W. Salzburger, I. Braasch, and A. Meyer, unpublished data). A function of *Csf1r* in pigment cell or neural crest development has not been found in mammals, which is not surprising, because mammals lost xanthophores (and iridophores) secondarily. It would thus be interesting to analyze, whether or not *Csf1r* is also involved in the development of xanthophores in amphibians. Most likely, the pigmentation function of the teleost *csf1ra* gene might have already been present in the single ancestral *csf1r* gene. Consequently, because *csf1ra* fulfills ancestral functions in fish, *csf1rb* might have evolved new functions, which would explain the asymmetric rates of molecular evolution of the *csf1r* paralogs. In the light of these observations, we propose neofunctionalization to be the most likely mechanism for the evolution of present day's functions of the entire B-paralogon.

### Genome Duplication, Signaling Cascade Expansions, and the Evolution of Teleost Coloration

Type III RTKs function as membrane spanning cell-surface receptors that bind ligand homo- or heterodimers and are activated by homo- or heterodimerization, resulting in autophosphorylation of intracellular tyrosine residues and subsequent triggering of the signal transduction cascade (Schlessinger 2000). Amino acid residues that are most likely involved in functional divergence of *pdgfrβ* and *csf1r* paralogs are mainly found in the extracellular, ligand-binding, and receptor interacting domain, suggesting that new ligand-binding affinities and receptor interactions might have evolved in fish after the duplication of *pdgfrβ* and *csf1r*. Therefore, it seems plausible that the FSGD and the associated multiplication of RTKs improved the number of possible ligand-receptor and receptor-receptor interactions and, thus, might have increased the complexity of signal transduction by type III RTKs, probably by functional specialization of signaling pathway paralogs.

What might have been the impact of genome duplications and the expansion of the type III RTK family for the evolution of teleost coloration? According to the model presented in figure 5*B*, genome duplications and the resulting expansion of the RTK multigene family coincide phylogenetically with the appearance of pigment cell innovations. The first round of genome duplication in the vertebrate lineage (1R) took place after their split from cephalochordates. This event coincides with the acquisition of the neural crest in vertebrates and of the neural crest-derived melanophores (Shimeld and Holland 2000). The second round of genome duplication (2R), which probably took place in the gnathostome stem lineage after the divergence of agnathans (Holland et al. 1994), correlates with the duplication of the proto-*pdgfr*–proto-*kit*/*csf1r* locus (fig. 5*B*). The presence of xanthophores and iridophores in amphibians and reptiles (Bagnara 1998) and their absence in lampreys (Mellgren and Johnson 2002) suggests that these two pigment cell types arose only in the gnathostome lineage. The similar roles of *kit* and *csf1r* in the development of melanophores and xanthophores, respectively, seem to suggest that functional divergence of *kit* and *csf1r* was associated with the expansion of the pigment cell repertoire.

The FSGD (3R) occurred approximately 250–350 MYA, close to the most recent common ancestor of teleosts after their divergence from the basal lineages including

bichirs, sturgeons, gars, and bowfin (Hoegg et al. 2004; Crow et al. 2006). This genome duplication coincides with the emergence of fish-specific paralogs of *csf1r*, *pdgfrβ* (this study), *kit, pdgfrα* (Mellgren and Johnson 2005), and other pigmentation genes such as *mitf* (Altschmied et al. 2002) or *sox10* (Lang et al. 2005*b*). Thus, duplicates originating from the FSGD might have been retained particularly in the context of pigmentation. Ever since this event, chromatophore sublineages, i.e., subpopulations of chromatophores that develop under different genetic control (Mellgren and Johnson 2002), might have been present in teleosts. The *csf1rb* paralog was, most likely, lost secondarily in the lineage leading to zebrafish and probably underwent neofunctionalization in the lineage leading to acanthopterygians. In the light of the presented model, genome duplications, and the FSGD in particular are likely to have been responsible for the expansion of the type III RTK family and might have resulted in an increase in signal cascade complexity that could be causally linked to the complexity and diversity of coloration and color patterning in teleost fishes.

Finally, the present study underlines the suitability of cichlids as models in comparative genomic investigations—for studying their unique evolutionary success but also as genomic models of perciform fishes, the most species-rich group of teleosts (Kocher et al. 2004). Thus, we emphasize the importance of sequencing a cichlid genome in the near future.

## Supplementary Material

Supplementary Tables S1–S3 and Supplementary Figures S1–S2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

## Acknowledgments

We thank M. Lang, N. Siegel, and J. Luo for assistance in the lab and other members from the Meyer lab as well as J. S. Taylor and J.-N. Volff for valuable suggestions on the manuscript. This work was funded by a Marie Curie Fellowship of the European Union, grants from the Landesstiftung Baden-Württemberg GmbH, and from the Center for Junior Research Fellows, University of Konstanz (to W.S.) and by the Deutsche Forschungsgemeinschaft (DFG) and the University of Konstanz (to A.M.).

## Literature Cited

Altschmied, J., J. Delfgaauw, B. Wilde, J. Duschl, L. Bouneau, J. N. Volff, and M. Schartl. 2002. Subfunctionalization of duplicate mitf genes associated with differential degeneration of alternative exons in fish. Genetics **161**:259–267.

Amores, A., A. Force, Y. L. Yan et al. (13 co-authors). 1998. Zebrafish hox clusters and vertebrate genome evolution. Science **282**:1711–1714.

Bagnara, J. T. 1998. Comparative anatomy and physiology of pigment cells in nonmammalian tissues. Pp. 9–40 *in* J. J. Nordlund, R. E. Boissy, V. J. Hearing, R. A. King, and J.-P. Ortonne, eds. The pigmentary system: physiology and pathophysiology. Oxford University Press, New York.

Béjar, J., Y. Hong, and M. Schartl. 2003. Mitf expression is sufficient to direct differentiation of medaka blastula derived stem cells to melanocytes. Development **130**:6545–6553.

Brenner, S., G. Elgar, R. Sandford, A. Macrae, B. Venkatesh, and S. Aparicio. 1993. Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. Nature **366**: 265–268.

Christoffels, A., E. G. Koh, J. M. Chia, S. Brenner, S. Aparicio, and B. Venkatesh. 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Mol. Biol. Evol. **21**:1146–1151.

Conant, G. C., and A. Wagner. 2003. Asymmetric sequence divergence of duplicate genes. Genome Res. **13**:2052–2058.

Crow, K. D., P. F. Stadler, V. J. Lynch, C. Amemiya, and G. P. Wagner. 2006. The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. Mol. Biol. Evol. **23**: 121–136.

Dehal, P., and J. L. Boore. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. **3**:e314.

Dorsky, R. I., R. T. Moon, and D. W. Raible. 1998. Control of neural crest cell fate by the Wnt signalling pathway. Nature **396**:370–373.

Dutton, K. A., A. Pauliny, S. S. Lopes, S. Elworthy, T. J. Carney, J. Rauch, R. Geisler, P. Haffter, and R. N. Kelsh. 2001. Zebrafish colourless encodes sox10 and specifies non-ectomesenchymal neural crest fates. Development **128**:4113–4125.

Ferrier, D. E., K. Dewar, A. Cook, J. L. Chang, A. Hill-Force, and C. Amemiya. 2005. The chordate ParaHox cluster. Curr. Biol. **15**:R820–R822.

Furlong, R. F., and P. W. Holland. 2002. Were vertebrates octoploid? Philos. Trans. R. Soc. Lond. B Biol. Sci. **357**:531–544.

Gu, J., and X. Gu. 2003. Natural history and functional divergence of protein tyrosine kinases. Gene **317**:49–57.

Gu, X. 1999. Statistical methods for testing functional divergence after gene duplication. Mol. Biol. Evol. **16**:1664–1674.

Gu, X., and K. Vander Velden. 2002. DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. Bioinformatics **18**:500–501.

Hoegg, S., H. Brinkmann, J. S. Taylor, and A. Meyer. 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J. Mol. Evol. **59**:190–203.

Hoegg, S., and A. Meyer. 2005. Hox clusters as models for vertebrate genome evolution. Trends Genet. **21**:421–424.

Holland, P. W., J. Garcia-Fernandez, N. A. Williams, and A. Sidow. 1994. Gene duplications and the origins of vertebrate development. Dev. Suppl. 125–133.

Hughes, A. L. 1994. The evolution of functionally novel proteins after gene duplication. Proc. R. Soc. Lond. B Biol. Sci. **256**:119–124.

Hughes, A. L., and R. Friedman. 2003. 2R or not 2R: testing hypotheses of genome duplication in early vertebrates. J. Struct. Funct. Genomics **3**:85–93.

Jaillon, O., J. M. Aury, F. Brunet et al. (61 co-authors). 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature **431**:946–957.

Knight, R. D., S. Nair, S. S. Nelson, A. Afshar, Y. Javidan, R. Geisler, G. J. Rauch, and T. F. Schilling. 2003. Lockjaw encodes a zebrafish tfap2a required for early neural crest development. Development **130**:5755–5768.

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. **5**:288–298.

Kocher, T., R. D. Fernald, H. A. Hofmann, A. Meyer, N. Okada, and D. Penman. 2004. Genome sequence of a cichlid fish: the Tilapia (*Oreochromis niloticus*). http://hcqs.unh.edu/cichlid/.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. **5**:150–163.

Lang, D., M. M. Lu, L. Huang, K. A. Engleka, M. Zhang, E. Y. Chu, S. Lipner, A. Skoultchi, S. E. Millar, and J. A. Epstein.

2005*a*. Pax3 functions at a nodal point in melanocyte stem cell differentiation. Nature **433**:884–887.

Lang, M., T. Miyake, I. Braasch, D. Tinnemore, N. Siegel, W. Salzburger, C. T. Amemiya, and A. Meyer. 2005*b*. A BAC library of the East African haplochromine cichlid fish Astatotilapia burtoni. J. Exp. Zool. B Mol. Dev. Evol. **306B**:35–44.

Larhammar, D., L. G. Lundin, and F. Hallbook. 2002. The human Hox-bearing chromosome regions did arise by block or chromosome (or even genome) duplications. Genome Res. **12**:1910–1920.

Leveugle, M., K. Prat, C. Popovici, D. Birnbaum, and F. Coulier. 2004. Phylogenetic analysis of Ciona intestinalis gene superfamilies supports the hypothesis of successive gene expansions. J. Mol. Evol. **58**:168–181.

McLysaght, A., K. Hokamp, and K. H. Wolfe. 2002. Extensive genomic duplication during early chordate evolution. Nat. Genet. **31**:200–204.

Mellgren, E. M., and S. L. Johnson. 2002. The evolution of morphological complexity in zebrafish stripes. Trends Genet. **18**:128–134.

———. 2005. kitb, a second zebrafish ortholog of mouse Kit. Dev. Genes Evol. **215**:1–8.

Meyer, A., and M. Schartl. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. **11**:699–704.

Meyer, A., and Y. Van de Peer. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays **27**:937–945.

Naruse, K., M. Tanaka, K. Mita, A. Shima, J. Postlethwait, and H. Mitani. 2004. A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping. Genome Res. **14**:820–828.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, New York.

Parichy, D. M., D. G. Ransom, B. Paw, L. I. Zon, and S. L. Johnson. 2000. An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. Development **127**:3031–3044.

Parichy, D. M., J. F. Rawls, S. J. Pratt, T. T. Whitfield, and S. L. Johnson. 1999. Zebrafish sparse corresponds to an orthologue of c-kit and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. Development **126**: 3425–3436.

Parichy, D. M., and J. M. Turner. 2003. Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development. Development **130**:817–833.

Pond, S. L., S. D. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics **21**:676–679.

Postlethwait, J., A. Amores, W. Cresko, A. Singer, and Y. L. Yan. 2004. Subfunction partitioning, the teleost radiation and the annotation of the human genome. Trends Genet. **20**:481–490.

Postlethwait, J. H., I. G. Woods, P. Ngo-Hazelett, Y. L. Yan, P. D. Kelly, F. Chu, H. Huang, A. Hill-Force, and W. S. Talbot. 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res. **10**:1890–1902.

Raes, J., and Y. Van de Peer. 2003. Gene duplication, the evolution of novel gene functions, and detecting functional divergence of duplicates in silico. Appl. Bioinform. **2**:91–101.

Roca, F. J., M. A. Sepulcre, G. Lopez-Castejon, J. Meseguer, and V. Mulero. 2005. The colony-stimulating factor-1 receptor is a specific marker of macrophages from the bony fish gilthead seabream. Mol. Immunol. **43**:1418–1423.

Rousset, D., F. Agnes, P. Lachaume, C. Andre, and F. Galibert. 1995. Molecular evolution of the genes encoding receptor tyrosine kinase with immunoglobulinlike domains. J. Mol. Evol. **41**:421–429.

Salzburger, W., and A. Meyer. 2004. The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. Naturwissenschaften **91**: 277–290.

Salzburger, W., H. Niederstätter, A. Brandstätter, B. Berger, W. Parson, J. Snoeks, and C. Sturmbauer. 2006. Colour-assortative mating among populations of Tropheus moorii, a cichlid fish from Lake Tanganyika, East Africa. Proc. R. Soc. Lond. B Biol. Sci. **273**:257–266.

Santini, S., J. L. Boore, and A. Meyer. 2003. Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. Genome Res. **13**:1111–1122.

Schlessinger, J. 2000. Cell signaling by receptor tyrosine kinases. Cell **103**:211–225.

Shimeld, S. M., and P. W. Holland. 2000. Vertebrate innovations. Proc. Natl. Acad. Sci. USA **97**:4449–4452.

Spring, J. 1997. Vertebrate evolution by interspecific hybridisation—are we polyploid? FEBS Lett. **400**:2–8.

———. 2002. Genome duplication strikes back. Nat. Genet. **31**:128–129.

Steinke, D., W. Salzburger, and A. Meyer. 2006. Novel relationships among ten fish model species revealed based on a phylogenomic analysis using ESTs. J. Mol. Evol. (in press).

Swofford, D. 2000. PAUP*: phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland, Mass.

Taylor, J. S., I. Braasch, T. Frickey, A. Meyer, and Y. Van de Peer. 2003. Genome duplication, a trait shared by 22000 species of ray-finned fish. Genome Res. **13**:382–390.

Taylor, J. S., Y. Van de Peer, I. Braasch, and A. Meyer. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. Philos. Trans. R. Soc. Lond. B Biol. Sci. **356**:1661–1679.

Van de Peer, Y., J. S. Taylor, I. Braasch, and A. Meyer. 2001. The ghost of selection past: rates of evolution and functional divergence of anciently duplicated genes. J. Mol. Evol. **53**: 436–446.

Vandepoele, K., W. De Vos, J. S. Taylor, A. Meyer, and Y. Van de Peer. 2004. Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. Proc. Natl. Acad. Sci. USA **101**:1638–1643.

Volff, J. N. 2005. Genome evolution and biodiversity in teleost fish. Heredity **94**:280–294.

Wagner, G. P., K. Takahashi, V. Lynch, S. J. Prohaska, C. Fried, P. F. Stadler, and C. Amemiya. 2005. Molecular evolution of duplicated ray finned fish HoxA clusters: increased synonymous substitution rate and asymmetrical co-divergence of coding and non-coding sequences. J. Mol. Evol. **60**:665–676.

Wang, Y., and X. Gu. 2001. Functional divergence in the caspase gene family and altered functional constraints: statistical analysis and prediction. Genetics **158**:1311–1320.

Williams, H., S. Brenner, and B. Venkatesh. 2002. Identification and analysis of additional copies of the platelet-derived growth factor receptor and colony stimulating factor 1 receptor genes in fugu. Gene **295**:255–264.

Woods, I. G., C. Wilson, B. Friedlander, P. Chang, D. K. Reyes, R. Nix, P. D. Kelly, F. Chu, J. H. Postlethwait, and W. S. Talbot. 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. Genome Res. **15**:1307–1314.

Adriana Briscoe, Associate Editor

Accepted March 9, 2006