

RESEARCH ARTICLE

Pigmentation Pathway Evolution after Whole-Genome Duplication in Fish

Ingo Braasch,* Frédéric Brunet,† Jean-Nicolas Volff,† and Manfred Schartl*

*Physiological Chemistry I, University of Würzburg, Biozentrum, Würzburg, Germany; and †Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Institut Fédératif Biosciences Gerland Lyon Sud, Université Lyon 1, Centre National de la Recherche Scientifique, Institut National de Recherche Agronomique, École Normale Supérieure de Lyon, France

Whole-genome duplications (WGDs) have occurred repeatedly in the vertebrate lineage, but their evolutionary significance for phenotypic evolution remains elusive. Here, we have investigated the impact of the fish-specific genome duplication (FSGD) on the evolution of pigmentation pathways in teleost fishes. Pigmentation and color patterning are among the most diverse traits in teleosts, and their pigmentary system is the most complex of all vertebrate groups.

Using a comparative genomic approach including phylogenetic and synteny analyses, the evolution of 128 vertebrate pigmentation genes in five teleost genomes following the FSGD has been reconstructed. We show that pigmentation genes have been preferentially retained in duplicate after the FSGD, so that teleosts have 30% more pigmentation genes compared with tetrapods. This is significantly higher than genome-wide estimates of FSGD gene duplicate retention in teleosts. Large parts of the melanocyte regulatory network have been retained in two copies after the FSGD. Duplicated pigmentation genes follow general evolutionary patterns such as the preservation of protein complex stoichiometries and the overrepresentation of developmental genes among retained duplicates. These results suggest that the FSGD has made an important contribution to the evolution of teleost-specific features of pigmentation, which include novel pigment cell types or the division of existing pigment cell types into distinct subtypes. Furthermore, we have observed species-specific differences in duplicate retention and evolution that might contribute to pigmentary diversity among teleosts.

Our study therefore strongly supports the hypothesis that WGDs have promoted the increase of complexity and diversity during vertebrate phenotypic evolution.

Introduction

It is now generally accepted that several rounds of polyploidization and rediploidization have occurred in vertebrates, including two rounds of whole-genome duplication (WGD) at the base of the vertebrate lineage (1R/2R duplications; Ohno 1970; Furlong and Holland 2002; Dehal and Boore 2005; Putnam et al. 2008). An additional fish-specific genome duplication (FSGD) has taken place in ray-finned fishes before the radiation of teleost fishes (Amores et al. 1998; Taylor et al. 2003; Jaillon et al. 2004; Meyer and Van de Peer 2005; Kasahara et al. 2007). In addition, more recent, lineage-specific WGDs have been observed in various vertebrate lineages (Otto 2007).

Although mutational inactivation and loss of redundant copies (nonfunctionalization) is the most common fate of one copy of duplicated genes, many duplicate genes are retained after rediploidization. Hence, WGDs add substantial amounts of genetic material to a genome (Lynch and Conery 2000; Blomme et al. 2006). An important question is, whether specific classes of genes are preferentially retained in duplicate after a WGD event. Recent studies on WGDs in eukaryotes as divergent as yeasts (Scannell et al. 2007; Wapinski et al. 2007), ciliates (Aury et al. 2006), plants (Maere et al. 2005), early vertebrates (Putnam et al. 2008), teleosts (Brunet et al. 2006), and the frog *Xenopus laevis* (Semon and Wolfe 2008) revealed that slowly evolving as well as highly expressed genes are preferen-

tially retained, and gene dosage within protein complexes or metabolic pathways is generally kept. Furthermore, in multicellular organisms, genes involved in developmental processes, regulation of transcription and signal transduction are maintained at a high rate after WGDs. The retention of such genes is rather uncommon for paralogs generated through small-scale, more local duplications (Maere et al. 2005; Blomme et al. 2006; Brunet et al. 2006; Putnam et al. 2008).

The evolutionary consequences of WGDs and their potential contribution to the evolutionary success of anciently polyploid (paleopolyploid) species in the long term are not well understood (Otto 2007). Theoretical models predict that differential loss of gene duplicates in isolated populations may lead to genomic incompatibilities and ultimately speciation (Werth and Windham 1991; Lynch and Force 2000). Correlations of such reciprocal gene loss (RGL) and increased speciation rates after WGDs have been observed in yeasts (Scannell et al. 2006), ciliates (Aury et al. 2006), and teleost fish (Semon and Wolfe 2007). Furthermore, many authors have suggested that WGDs have provided the genetic raw material for important morphological transitions, key innovations, and increased phenotypic diversity and complexity (e.g., Ohno 1970; Holland et al. 1994; Aburomia et al. 2003; De Bodt et al. 2005), whereas this effect of WGDs was regarded as less important by others (Donoghue and Purnell 2005; Carroll 2008).

Here, we have analyzed the impact of the FSGD on the evolution of pigmentation pathways in teleost fish. Pigmentation and color patterning are among the most variable traits in vertebrates (Braasch et al. 2008; Protas and Patel 2008). Importantly, the pigmentary system of teleost fishes is the most diverse and complex of all vertebrates and

Key words: genome duplication, fish, conserved synteny, pigment cell, melanocyte, functional module.

E-mail: ingo.braasch@biozentrum.uni-wuerzburg.de.

Genome Biol. Evol. Vol. 2009:479–493.

doi:10.1093/gbe/evp050

Advance Access publication November 25, 2009

therefore offers an excellent opportunity to study the effect of WGD on phenotypic evolution (Braasch et al. 2007, 2008). Teleosts have more neural crest-derived pigment cell types than all other vertebrate groups. Although black melanophores, reflecting iridophores and yellow-red xanthophores/erythrophores must have already been present in the common ancestor of ray-finned fish and tetrapods, teleosts have evolved an additional pigment cell type, the whitish leucophores, as well as distinct subtypes of the aforementioned pigment cell types (Mellgren and Johnson 2002; Bagnara and Matsumoto 2006; Braasch et al. 2008). In addition, blue cyanophores are present in some teleost lineages (Bagnara et al. 2007), and there is recent evidence for red fluorescent pigmentation in several reef fish species (Michiels et al. 2008).

It has been suggested that the FSGD was of major importance for the evolution of teleost pigmentation (Mellgren and Johnson 2002; Braasch et al. 2008). In a previous study, we could show that teleost fishes have more genes encoding pigment synthesis enzymes than any other vertebrate group as result of the FSGD (Braasch et al. 2007). Here, we have expanded our analysis to a list of all 128 known “vertebrate pigmentation genes.” We define vertebrate pigmentation genes as those genes that were shown in at least one vertebrate species to be involved in the development and/or differentiation of neural crest-derived pigment cells. The fact that almost all pigmentation genes found in mammals are also found to be involved in pigmentation in teleosts and vice versa (reviewed, e.g., in Rawls et al. 2001; Braasch et al. 2008) points to a general evolutionary conservation of vertebrate pigment cell development and justifies our assumption that a pigmentation gene of one vertebrate group is highly likely to have pigmentation functions in other lineages.

Importantly, many of the genes involved in pigment cell development have other functions not related to pigmentation. According to the duplication-degeneration-complementation model (Force et al. 1999), ancestral genes might be distributed among duplicated genes (subfunctionalization). Thus, in some cases, pigmentation functions could become separated from other functions after duplication. In this case, the duplication would not increase the number of pigmentation genes per se because only one paralog retains the pigmentation function. However, because the paralog that keeps the pigmentation function will be released from functional constraints imposed by other essential functions present in the ancestral gene, the duplication still might facilitate evolution of pigmentation by specializing one of the paralogs for the pigmentation function (Braasch et al. 2008). Thus, for the sake of the present study, the term “pigmentation gene” corresponds to the presumed ancestral gene’s pigmentation function at the time of its duplication.

By reconstructing the evolutionary history of pigmentation gene families in the vertebrate lineage and by carefully distinguishing gene duplicates generated in the FSGD from those originating in the earlier rounds of WGD (1R, 2R) and in small-scale duplication events, we show that extant teleost genomes (zebrafish, medaka, stickleback, Tetraodon, fugu) contain $\approx 30\%$ more pigmentation genes compared with tetrapods as result of FSGD. This is signif-

icantly higher than available estimates for genome-wide retention rates ($\approx 12\text{--}24\%$). We found important differences in FSGD duplicate retention rates among functional categories of pigmentation genes, and most of the known major regulators of pigment cell development have been retained in duplicate. Our results therefore point to a major role of WGDs in the evolution of vertebrate phenotypes.

Materials and Methods

Sequence Database Surveys

Nucleotide sequences of pigmentation genes were identified using Blast searches against GenBank (<http://www.ncbi.nlm.nih.gov/>) and Ensembl genome assemblies (<http://www.ensembl.org/index.html>; version 50, July 2008) of human (*Homo sapiens*; Hsa), mouse (*Mus musculus*; Mmu), chicken (*Gallus gallus*; Gga), frog (*Xenopus tropicalis*; Xtr), zebrafish (*Danio rerio*; Dre), medaka (*Oryzias latipes*; Ola), spotted green pufferfish (*Tetraodon nigroviridis*; Tni), torafugu (*Takifugu rubripes*; Tru), and stickleback (*Gasterosteus aculeatus*; Gac). Usually, human sequences were used as initial queries. The results of the Blast searches were double-checked against the Ensembl gene families (if available) to collect missing sequences and to get an initial overview of the gene family tree topology. Pigmentation genes and their accession numbers are provided in supplementary tables S1–S2 and supplementary figures S1–S24 (Supplementary Material online).

Sequence Alignments and Phylogenetic Reconstructions

Nucleotide sequences obtained from Blast searches were loaded into BioEdit (Hall 1999), translated into proteins, and aligned using ClustalW (Thompson et al. 1994). Alignments were carefully checked and ambiguously aligned regions were removed before phylogenetic analyses. Maximum likelihood phylogenies based on protein data were computed with PHYML (Guindon et al. 2005) with 100 bootstrap replicates. Models of protein evolution (mainly JTT + I + G) and parameter values were determined with ProtTest (Abascal et al. 2005). MEGA4 (Tamura et al. 2007) was used to obtain Neighbor-Joining bootstrap values of 10,000 replicates and to draw phylogenies. Trees were rooted either with the closest vertebrate paralog in case of larger gene families or with an invertebrate ortholog. For some phylogenies, no suitable outgroup sequence was available (*Dtmbp1*, *Atp6v0c*, *Nsf*, *Itgb1*). Although Blast hits for potential outgroup sequences were available, these were too divergent for a reliable alignment and were thus excluded from the phylogenetic analyses.

Synteny Analyses

For microsynteny analyses, genes surrounding the human ortholog were used as initial queries for Blast searches against the teleost genome assemblies at Ensembl, followed by reciprocal Blast searches of the best hits against human and other teleost genomes.

For macrosynteny analysis, the last release of Tetraodon genome assembly (v8) was used to generate a

rose-window representation as in Jaillon et al. (2004) and Brunet et al. (2006). In this meta-analysis, we took all Tetraodon proteins (available from the ftp site of Ensembl v50), did a Blast all against all of these proteins under default parameters, and selected the first hits for which genes are on different chromosomes. A cutoff P value $\leq 1 \times 10^{-3}$ was applied, and the background was reduced by removing intrachromosomal links and links between chromosomes for which the total number is below eight. Of note, with this last release and this method, we increased the number of strong links between the chromosomes, corroborating the evidence for the FSGD obtained previously (Jaillon et al. 2004; Brunet et al. 2006). This whole-genome analysis of gene duplication was superimposed with the links collected for the duplicated pigmentation genes, which confirmed the FSGD origin for most of the duplicated pigmentation genes.

Functional Annotation Analysis

The Functional Annotation Tool of the DAVID Bioinformatics Resources (<http://david.abcc.ncifcrf.gov/>) was used to test for enrichment in gene ontology (GO) functional terms (Huang et al. 2009). From the lists of pigmentation and liver genes (supplementary tables S2–S3, Supplementary Material online), the Ensembl gene IDs of human orthologs were uploaded and compared with the human reference genome as background. Enrichments were considered significant if the modified Fisher Exact P value was < 0.05 .

Results

Vertebrate Pigmentation Genes

As a first step, we undertook database and literature surveys and generated a list of vertebrate pigmentation genes. From mammals, the list included over 100 known genes involved in coat color formation in mice (Bennett and Lamoreux 2003) from the Color Gene Database (<http://www.espcr.org/micemut/>) extended by genes important for skin and hair color differences between human populations, which were recently identified through genome-wide association studies (Bastian and Pinkel 2008). Many of these genes have also been identified as coat color genes in other mammals (e.g., rat, cat, dog, cow, sheep, pig, horse). Additional genes in the list affect plumage coloration in birds (e.g., Miwa et al. 2007). Pigmentation genes identified in teleosts were obtained from pigmentation mutant collections in zebrafish (Haffter et al. 1996; Kelsh et al. 1996; Odenthal et al. 1996) available from the Zebrafish Model Organism Database (ZFIN; <http://zfin.org>) and in medaka (Kelsh et al. 2004). Finally, we included information given by the Gene Ontology Database (GO term “pigmentation” GO:0043473 and children terms; <http://www.geneontology.org/>). From the initial pigmentation gene list, we omitted genes affecting pigmentation solely by their function in the development of hair or the retinal pigment epithelium (RPE). Genes required for the production of pheomelanin, a lighter type of melanin pigment not found in teleosts and other poikilotherm vertebrates (Fujii 1993b), were also not included. These genes are generally present in

vertebrates including teleosts but have acquired their role in pheomelanin synthesis in the lineages of mammals and birds.

The final list of vertebrate pigmentation genes included in this study consists of 128 genes and is presented in supplementary table S1 (Supplementary Material online). Genes were grouped into ten functional categories adopting the classification for mouse coat color mutants (<http://www.espcr.org/micemut/>) and using additional information given by ZFIN (<http://zfin.org>) and the Gene Ontology Database (<http://www.geneontology.org/>): 1) melanophore development, 2) components of melanosomes, 3) melanosome biogenesis, 4) melanosome transport, 5) regulation of melanogenesis, 6) systemic effects, 7) xanthophore development, 8) pteridine synthesis, 9) iridophore development, and 10) uncategorized function (supplementary table S1, Supplementary Material online). Other categories may exist but are not represented by a known gene so far (e.g., pterinosome biogenesis, transport, etc.). Some genes were included into multiple categories because they are involved in the formation of different pigment cell types (e.g., *Sox10*, Dutton et al. 2001).

Furthermore, we chose “liver genes” to contrast our data on pigmentation genes with another category of functionally related genes that is not expected to be under high selection for duplicate gene retention. To this end, 187 liver genes obtained from the Gene Ontology Database (GO term “liver development” GO:0001889; <http://www.geneontology.org/>) and the collection of zebrafish liver mutants from ZFIN (<http://zfin.org>) were analyzed similarly as the pigmentation genes.

Identification of 46 Fish-Specific Pairs of Pigmentation Gene Duplicates

Fish-specific gene duplicates were identified by complementary approaches combining phylogenetic reconstructions and synteny analyses. Sequence information from the genome assemblies of five teleosts (zebrafish, medaka, stickleback, Tetraodon, fugu) and four tetrapods (frog, chicken, mouse, human) was included. This comparative genomic approach is particularly powerful to infer the duplication history of gene families in case of ambiguous tree topologies as well as species-specific gene order rearrangements or genome assembly problems (Braasch et al. 2008). Distinguishing FSGD duplicates from other types of WGD (1R/2R) or local gene duplications is essential because duplicates of varying age and type show different patterns of duplicate retention and functional evolution (Maere et al. 2005; Aury et al. 2006; Wapinski et al. 2007). Criteria that were used to conclude that pairs of gene duplicates were derived from the FSGD are illustrated by the example of *Sox10* in figure 1. According to the tree topology, the duplication of the teleosts paralogs should date back to one common ancestor of the five teleost species but within the fish lineage after split from tetrapods (Fig. 1A). Furthermore, the gene duplicates should be part of paralogons (i.e., large chromosomal blocks of duplicated gene pairs) that show conserved synteny with each other and with an outgroup species, for example, human (double conserved

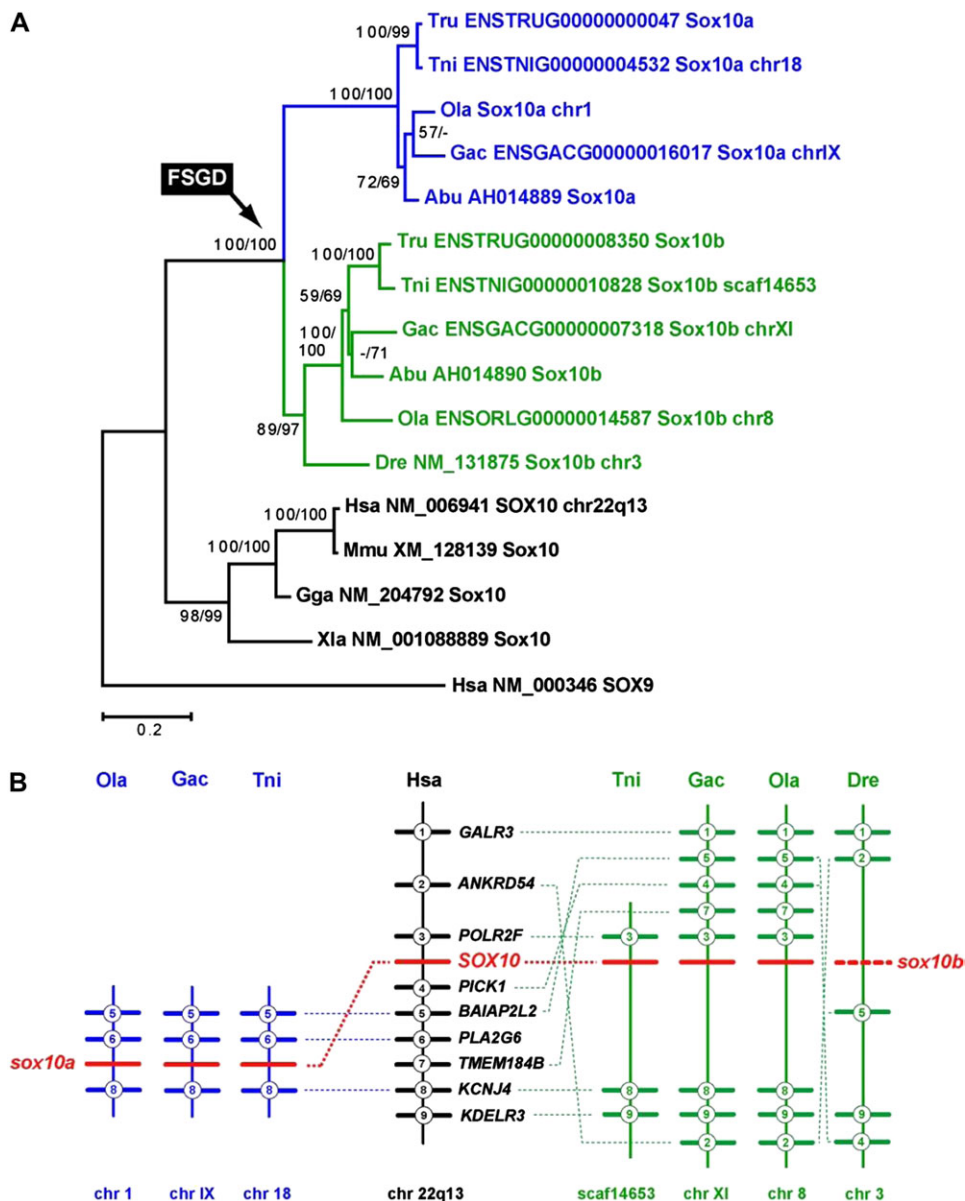


FIG. 1.—Evolution of the *Sox10* transcription factor gene. (A) Maximum likelihood (ML) phylogeny of vertebrate *Sox10* proteins based on 563 AA positions. The tree is rooted with human SOX9. Bootstrap values (ML/Neighbor-Joining [NJ]) above 50% are shown. The FSGD generated two *Sox10* in teleosts, *Sox10a* (blue), and *Sox10b* (green). *Sox10a* is missing in zebrafish. (B) The microsynteny pattern of *Sox10* regions in vertebrate genome shows double conserved synteny between the two teleost *sox10* paralogs and human chr22q13. A *sox10a* paralog could not be identified in zebrafish, and the zebrafish *sox10b* (dashed red bar) is not included in the Zv7 genome assembly but has been mapped to chr3 (Dutton et al. 2001). Numbered bars represent genes contributing to conserved synteny, and genes that do not contribute to conserved synteny are not shown. Dotted lines connect orthologous genes.

synteny). On a fine scale, genes in the vicinity of the paralogs are connected in a pattern of “microsynteny” (fig. 1B). On the genome-wide scale, FSGD-duplicated pigmentation genes take part in a higher level pattern of “macrosynteny” connecting entire paralogous chromosomes (fig. 2). Macrosynteny patterns have been previously used to reconstruct the protochromosomes of the pre-FSGD karyotype of the ancestral euteleost (Jaillon et al. 2004; Woods et al. 2005; Kasahara et al. 2007). Here, we used the most recent analysis of Kasahara et al. (2007) to assign pigmentation gene duplicates to the 13 protochromosomes (*a–m*) from which they are derived.

Of the 128 pigmentation genes analyzed, 46 genes (35.9%) were retained in two copies after the FSGD in at least 1 of the 5 teleost genomes and 65.2% (30/46) of the duplicated genes were present in two copies in all five teleost genomes (table 1; supplementary table S2, supplementary figs. S1–S24, Supplementary Material online). For 82 genes (64.1%), one of the paralogs has been lost in all five teleosts (supplementary table S2, Supplementary Material online). Twenty-two pairs of duplicates were previously shown to be derived from the FSGD, and 24 new pairs were identified in the present study (table 1). The 46 pigmentation gene paralogs are derived from the pre-FSGD

Table 1
FSGD-Duplicated Teleost Pigmentation Genes

Genes	Function	Homo	Coorthologs	<i>Danio</i>	<i>Oryzias</i>	<i>Gasterosteus</i>	<i>Tetraodon</i>	Protochromosomes	Figures/ref.
<i>Adam17</i>	a	Hsa2	<i>adam17a</i> <i>adam17b</i>	Dre17 Dre20	Ola22 Ola24	GacXV +	Tni10 Tni14	<i>a</i>	Figure S1
<i>Pomc</i>	e	Hsa2	<i>pomca</i> <i>pomcb</i>	Dre17 Dre20	Ola2 Ola24	GacX GacXVIII	+ Tni14	<i>a</i>	Ref. 1
<i>Rab32</i>	b	Hsa6	<i>rab32a</i> <i>rab32b</i>	Dre20 Dre8	Ola24 —	GacXVIII GacXV	Tni14 Tni10	<i>a</i>	Figure S2
<i>Slc24a4</i>	b	Hsa14	<i>slc24a4a</i> <i>slc24ab</i>	Dre17 +	— Ola24	— GacXVIII	— +	<i>a</i>	Figure S3
<i>Dtnbp1</i>	c	Hsa6	<i>dtnbp1a</i> <i>dtnbp1b</i>	Dre19 Dre16	Ola16 Ola11	GacXX GacX	+ +	<i>b</i>	Figure S4
<i>Hdac1</i>	a	Hsa1	<i>hdac1a</i> <i>hdac1b</i>	Dre19 —	Ola11 Ola22	GacX GacXV	Tni21 Tni10	<i>b</i>	Figure S5
<i>Pabpc1</i>	j	Hsa8	<i>pabpc1a</i> <i>pabpc1b</i>	Dre16 Dre19	Ola16 Ola11	GacXX GacX	Tni8 Tni21	<i>b</i>	Figure S6
<i>Creb1</i>	a, e	Hsa2	<i>creb1a</i> <i>creb1b</i>	Dre1 Dre20	Ola2 Ola21	GacI GacXVI	Tni3 Tni2	<i>c</i>	Figure S7
<i>Ednrb1</i>	a, g, i	Hsa13	<i>ednrb1a</i> <i>ednrb1b</i>	Dre1 Dre9	Ola17 Ola21	GacIII GacXVI	Tni15 Tni3	<i>c</i>	Ref. 2
<i>En1</i>	a	Hsa2	<i>eng1a</i> <i>eng1b</i>	Dre9 Dre1	+ Ola2	GacXVI GacI	+ +	<i>c</i>	Ref. 3
<i>Gja5</i>	j	Hsa1	<i>gja5a</i> <i>gja5b</i>	Dre1 Dre9	+ Ola21	GacVI GacXVI	Tni17 Tni2	<i>c?</i>	Figure S8
<i>Mlph</i>	d	Hsa2	<i>mlpha</i> <i>mlphb</i>	Dre6 Dre9	+ Ola21	+ GacXVI	Tni3 Tni2	<i>c</i>	Figure S9
<i>Zic2</i>	a	Hsa13	<i>zic2a</i> <i>zic2b</i>	Dre9 Dre1	Ola21 —	GacXVI —	Tni2 —	<i>c</i>	Figure S10
<i>Atp6v0c</i>	f	Hsa16	<i>atp6v0ca</i> <i>atp6v0cb</i>	Dre3 Dre24	Ola8 Ola19	GacXI GacV	+ Tni2	<i>e</i>	Figure S11
<i>Mgrn1</i>	e	Hsa16	<i>mgrn1a</i> <i>mgrn1b</i>	Dre12 Dre3	— Ola8	— GacXI	— Tni3	<i>e</i>	Figure S12
<i>Nsf</i>	c	Hsa17	<i>nsfa</i> <i>nsfb</i>	Dre3 Dre12	+ Ola19	GacXI GacV	+ Tni2	<i>e</i>	Figure S13
<i>Sox9</i>	a, i	Hsa17	<i>sox9a</i> <i>sox9b</i>	Dre12 Dre3	+ Ola8	GacV GacXI	Tni2 Tni3	<i>e</i>	Ref. 4
<i>Sox10</i>	a, g, i	Hsa22	<i>sox10a</i> <i>sox10b</i>	— Dre3	Ola1 Ola8	GacIX GacXI	Tni18 +	<i>e</i>	Figure1
<i>Kit</i>	a	Hsa4	<i>kiia</i> <i>kiib</i>	Dre20 Dre1	Ola4 Ola1	GacVIII GacIX	Tni1 Tni18	<i>f</i>	Ref. 5
<i>Qdpr</i>	h	Hsa4	<i>qdpra</i> <i>qdprb</i>	Dre20 Dre1	+ —	GacIV —	Tni20 —	<i>f</i>	Ref. 6
<i>Tyrp1</i>	b	Hsa9	<i>tyrp1a</i> <i>tyrp1b</i>	Dre7 Dre1	Ola18 Ola1	GacVII GacIX	+ —	<i>f</i>	Ref. 6
<i>Csflr</i>	g	Hsa5	<i>csflra</i> <i>csflrb</i>	+ —	Ola10 Ola14	GacIV GacVII	Tni1 Tni7	<i>g</i>	Ref. 7
<i>Smtl</i>	g	—	<i>smtla</i> <i>smtlb</i>	Dre14 Dre10	Ola13 —	GacI —	+ —	<i>g</i>	Refs. 8 and 9
<i>Frem2</i>	a	Hsa13	<i>frem2a</i> <i>frem2b</i>	Dre15 Dre10	Ola13 Ola14	GacI GacVII	Tni16 Tni7	<i>h</i>	Figure S14
<i>Myo7A</i>	d	Hsa11	<i>myo7Aa</i> <i>myo7Ab</i>	Dre18 Dre21	Ola13 Ola14	GacI GacVII	+ +	<i>h</i>	Figure S15
<i>Rab38</i>	b	Hsa11	<i>rab38a</i> <i>rab38b</i>	Dre15 Dre10	Ola13 Ola14	GacI GacVII	+ Tni7	<i>h</i>	Figure S16
<i>Tyr</i>	b	Hsa11	<i>tyra</i> <i>tyrb</i>	Dre15 —	Ola13 Ola14	GacI GacVII	+ Tni7	<i>h</i>	Ref. 6
<i>Ghr</i>	g	Hsa5	<i>ghra (slr)</i> <i>ghrb</i>	Dre8 Dre21	Ola9 Ola12	GacXIII GacXIV	Tni12 +	<i>i</i>	Ref. 9
<i>Gnaq</i>	a	Hsa9	<i>gnaqa</i> <i>gnaqb</i>	Dre5 —	Ola9 Ola12	GacXIII GacVII/XIV	Tni12 Tni4	<i>i</i>	Figure S17
<i>Spr</i>	h	Hsa2	<i>spra</i> <i>sprb</i>	Dre5 Dre8	Ola10 —	GacXIV GacXIII	+ —	<i>i</i>	Ref. 6
<i>Myo5A</i>	d	Hsa15	<i>myo5Aa</i> <i>myo5Ab</i>	Dre18 Dre25	Ola3 Ola6	GacII GacXIX	Tni5 Tni13	<i>j</i>	Figure S18
<i>Trpm1</i>	a	Hsa15	<i>trpm1a</i> <i>trpm1b</i>	Dre7 Dre25	Ola3 Ola6	GacII GacXIX	Tni5 Tni13	<i>j</i>	Figure S19
<i>Atp6v1e1</i>	f	Hsa22	<i>atp6v1e1a</i> <i>atp6v1e1b</i>	Dre4 Dre25	Ola23 Ola6	GacIV GacXIX	Tni19 Tni13	<i>k</i>	Figure S20
<i>Kitl</i>	a	Hsa12	<i>kiila</i> <i>kiilb</i>	Dre25 Dre4	+ Ola23	GacXIX GacIV	Tni13 Tni19	<i>k</i>	Ref. 10

Table 1
Continued

Genes	Function	Homo	Coorthologs	<i>Danio</i>	<i>Oryzias</i>	<i>Gasterosteus</i>	<i>Tetraodon</i>	Protochromosomes	Figures/ref.
<i>Edn3</i>	a, g, i	Hsa20	<i>edn3a</i> <i>edn3b</i>	Dre11 Dre23	Ola5 Ola7	GacXVII GacXII	— Tni9	<i>l</i>	Ref. 2
<i>Erb3</i>	a	Hsa12	<i>erbb3a</i> <i>erbb3b</i>	Dre6 Dre23	Ola5 Ola7	+ GacXII	Tni11 Tni9	<i>l</i>	Refs. 11 and 12
<i>Mitf</i>	a	Hsa3	<i>mitfa</i> <i>mitfb</i>	Dre6 Dre23	Ola5 Ola7	GacXVII GacXII	Tni11 Tni9	<i>l</i>	Ref. 13
<i>Silv</i>	b	Hsa12	<i>silva</i> <i>silvb</i>	Dre11 Dre23	Ola5 Ola7	+ GacXII	Tni11 Tni9	<i>l</i>	Ref. 6
<i>Pax7</i>	g	Hsa1	<i>pax7a</i> <i>pax7b</i>	Dre11 Dre23	Ola5 +	+ GacXII	Tni11 +	<i>l</i>	Ref. 14
<i>Egfr</i>	a	Hsa7	<i>egfra</i> <i>egfrb</i>	Dre2 Dre24	Ola17 Ola20	GacIII +	Tni15 Tni4	<i>m</i>	Refs. 11 and 12
<i>Gna11</i>	a	Hsa19	<i>gna11a</i> <i>gna11b</i>	Dre22 Dre2	Ola4 Ola22	GacVIII GacXV	Tni1 Tni10	<i>m</i>	Figure S17
<i>Irf4</i>	j	Hsa6	<i>irf4a</i> <i>irf4b</i>	Dre2 Dre20	Ola17 Ola4	GacIII GacVIII	+ Tni1	<i>m</i>	Figure S21
<i>Itgb1</i>	a	Hsa10	<i>itgb1a</i> <i>Itgb1b</i>	Dre24 Dre2	Ola20 Ola17	GacXXI +	+ Tni15	<i>m</i>	Figure S22
<i>Mcoln3</i>	a	Hsa1	<i>mcoln3a</i> <i>mcoln3b</i>	— Dre23	Ola17 Ola4	GacIII GacVIII	Tni15 Tni1	<i>m</i>	Figure S23
<i>Myc</i>	a	Hsa8	<i>myca</i> <i>mycb</i>	Dre24 Dre2	Ola20 Ola17	GacXXI —	Tni6 —	<i>m</i>	Figure S24
<i>Pax3</i>	a, g	Hsa2	<i>pax3a</i> <i>pax3b</i>	Dre2 Dre15	Ola17 Ola13	GacIII +	Tni15 Tni16	<i>m</i>	Ref. 14

NOTE.—Genomic location of the human gene and teleost coorthologs is indicated. Functions: a, melanophore development; b, components of melanosomes; c, melanosome construction; d, melanosome transport; e, regulation of melanogenesis; f, systemic effects; g, xanthophore development; h, pteridine synthesis; i, iridophore development; j, uncategorized function. Assignment to the pre-FSGD protochromosomes is according to Kasahara et al. (2007). Supplementary figures S1–S24 (phylogenetic trees) are found as Supplementary Material online. Key references (ref.) show the FSGD origin of teleost duplicates: 1, de Souza et al. (2005); 2, Braasch, Volf, and Schartl (2009); 3, Taylor et al. (2003); 4, Yan et al. (2005); 5, Siegel et al. (2007); 6, Braasch et al. (2007); 7, Braasch et al. (2006); 8, Zhu et al. (2004); 9, Fukamachi and Meyer (2007); 10, Hultman et al. (2007); 11, Gomez et al. (2004); 12, Froschauer et al. (2006); 13, Altschmid et al. (2002); 14, Minchin and Hughes (2008).

Material online). This gives a RGL rate of 3% (4/128 genes) between zebrafish and acanthomorphs for pigmentation genes. Adding the four RGL genes to the 46 retained duplicates at least 50 (39.1%) pigmentation gene duplicates had been retained in the last common euteleost ancestor of the five analyzed teleosts (fig. 3). This might still be an underestimation, as possible independent loss of orthologous gene copies in divergent lineages is not taken into account. Convergent duplicate loss becomes progressively more likely compared with RGL with advancing time since the duplication event and increasing sequence divergence of paralogs (Scannell et al. 2007). A similar analysis of liver genes revealed 4/187 instances (2%) of RGL (supplementary table S3, Supplementary Material online). Thus, the estimate for FSGD duplicate retention in the euteleost ancestor for liver genes is 21.4% (40/187 genes; fig. 3; supplementary fig. S25, Supplementary Material online). Within the ~50 Myr between the FSGD and the divergence of zebrafish from the acanthomorphs (medaka, stickleback, pufferfishes; Kasahara et al. 2007), both, pigmentation and liver genes, have lost the majority of FSGD duplicates. However, the loss rate has been significantly lower (χ^2 test, $P = 0.00065$) for pigmentation genes (60.9%) compared with liver genes (78.6%).

During the divergence of euteleosts, further lineage-specific duplicate losses have occurred that can be mapped onto the phylogeny (fig. 3). Similarly, independent loss of several pigmentation gene duplicates has occurred in the

lineage leading to zebrafish (ten losses) as well as in the lineage leading to acanthomorphs (nine losses). Within acanthomorphs, most gene losses have occurred shortly after the split from the zebrafish lineage and later lineage-specific loss of duplicates is rare compared with the liver genes (fig. 3).

In the extant five teleost genomes, finally, a similar FSGD-duplicated retention rate of ~30% is observed for pigmentation genes across the euteleost phylogeny. Pufferfishes do not deviate from this trend despite the massive genome compaction that has occurred in their lineage (Brenner et al. 1993). Liver genes, in contrast, have retained only around 12–16% of FSGD duplicates in the extant teleosts due to their higher loss rates before and after the euteleost divergence (fig. 3).

Duplicate Retention Rates Differ among Functional Categories

Next, we asked about the functional distribution of duplicated pigmentation genes. A comparison of the different pigment cell types shows (table 2) that the retention rate for melanophores (37.0%; 37/100 genes; categories a–f) and iridophores (36.4%; 4/11 genes; category i) are similar to the rate for all pigmentation genes (35.9%; 46/128 genes). For xanthophores, the retention rate seems to be elevated (50%; 11/22 genes; categories g and h) but this is not statistically significant (χ^2 test, $P = 0.20937$). Noticeably,

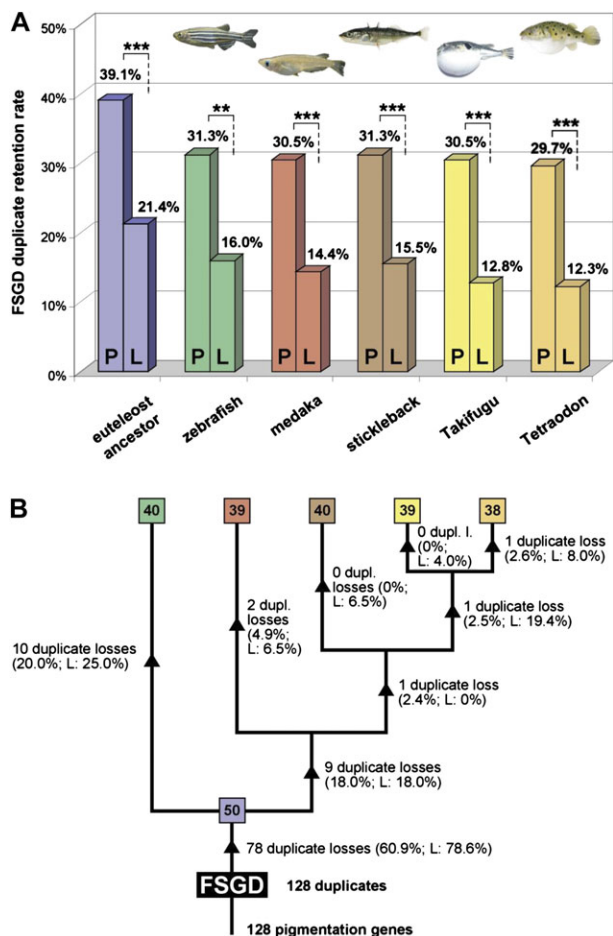


FIG. 3.—Evolution of teleost pigmentation gene repertoires following the FSGD. (A) In the five extant teleost genomes, around 30% of FSGD-duplicated pigmentation genes (P; $n = 128$) have been retained. Duplicated liver genes (L; $n = 187$) have been retained by only 14% in extant teleosts. In the hypothetical euteleost ancestor, almost twice as many duplicates had been retained for pigmentation genes compared with liver genes (39.1% vs. 21.4%). For all five teleost genomes as well as for the reconstructed euteleost ancestor, the retention rates of pigmentation versus liver genes are significantly different (χ^2 tests; $**P < 0.01$, $***P < 0.001$). (B) Losses of pigmentation gene duplicates mapped onto the teleost phylogeny (Setiamarga et al. 2008). Numbers of pigmentation genes retained in duplicate are given in boxes. Gene loss rates in percent are based on the number of retained duplicates at the previous node in the phylogeny. The majority of duplicate losses have occurred before the divergence of the five euteleost species (60.9%). Within acanthomorphs (medaka, stickleback, pufferfishes), further losses have occurred shortly after the split from the zebrafish lineage (18.0%) and lineage-specific gene losses are relatively rare. Gene loss rates of pigmentation genes are generally lower than of liver genes (L; see also supplementary fig. S25, Supplementary Material online).

genes that have been assigned to more than one pigment cell type are retained at a rate of 63.6% (7/11 genes; supplementary table S1, Supplementary Material online).

There are important differences among the different functional classes of pigmentation genes (table 2). Within the six melanophore categories, four of them have a higher retention rate than the overall average for pigmentation genes: melanophore development (48.9%; 22/45 genes), melanosome components (60.0%; 6/10 genes), melanosome transport (75.0%; 3/4 genes), and regulation of melanogenesis (42.9%; 3/7 genes). However, none of them

Table 2
FSGD Duplicate Retention Rates According to Pigmentation Gene Functions

Category	Number of genes	FSGD duplicates	Retention rate (%)
All pigmentation genes	128	46	35.9
Melanophore genes	100	37	37.0
Melanophore development	45	22	48.9
Melanosome components	10	6	60.0
Melanosome biogenesis*	22	2	9.1
Melanosome transport	4	3	75.0
Regulation of melanogenesis	7	3	42.0
Systemic effects	13	2	15.4
xanthophore genes	22	11	50.0
Xanthophore development**	11	9	81.8
Pteridine synthesis	11	2	18.2
Iridophore genes	11	4	36.4
Uncategorized	10	3	30.0
Genes in multiple categories	11	7	63.6

NOTE.—* $P < 0.05$, ** $P < 0.01$ in χ^2 tests compared with the category “all pigmentation genes.”

is significantly different from the pigmentation gene average. In contrast, genes involved in melanosome biogenesis have a significantly reduced FSGD duplicate retention rate of 9.1% compared with the pigmentation gene average (2/22 genes; χ^2 test, $P = 0.01264$). The category of systemic effect genes also has a comparatively low retention rate (15.4%; 2/13 genes).

The two xanthophore categories deviate in their FSGD duplicate retention rate. The rate for pteridine synthesis genes, which are involved in the formation of the yellow pigment, is comparatively low (18.2%; 2/11 genes). Interestingly, sepiapterin reductase, an enzyme that is involved in multiple steps of this metabolic pathway (Ziegler 2003), is encoded by two FSGD duplicates in teleosts (Braasch et al. 2007).

Xanthophore development genes, finally, have the highest retention rate of all categories (81.8%), which is significantly higher than the pigment cell average rate (χ^2 test, $P = 0.00282$). Due to this high retention rate in xanthophore development, xanthophore genes have a higher FSGD duplicate retention rate than melanophores and iridophores.

Duplications in the Pigment Cell Regulatory Networks

Crucial questions are how the duplicated teleost pigmentation genes are connected to each other and at which positions FSGD duplicates have been retained in the regulatory network of pigment cell development. Figure 4 illustrates our current understanding of the signaling network in vertebrate melanocyte/melanophore development with the *Mitf* transcription factor as master regulator in the central position (Béjar et al. 2003; Tachibana et al. 2003; Levy et al. 2006; Lin and Fisher 2007). Nearly all known key players of this network, that is, *Sox10*, *Pax3*, *Kit*, *Kitl*, *Ednrb1*, *Edn3*, and *Mitf*, have been retained in duplicate after FSGD. Thus, it appears that large parts of this pathway are present in two copies in teleosts. However, there are also subtle differences between teleost species in the retention of pathway component duplicates (fig. 4).

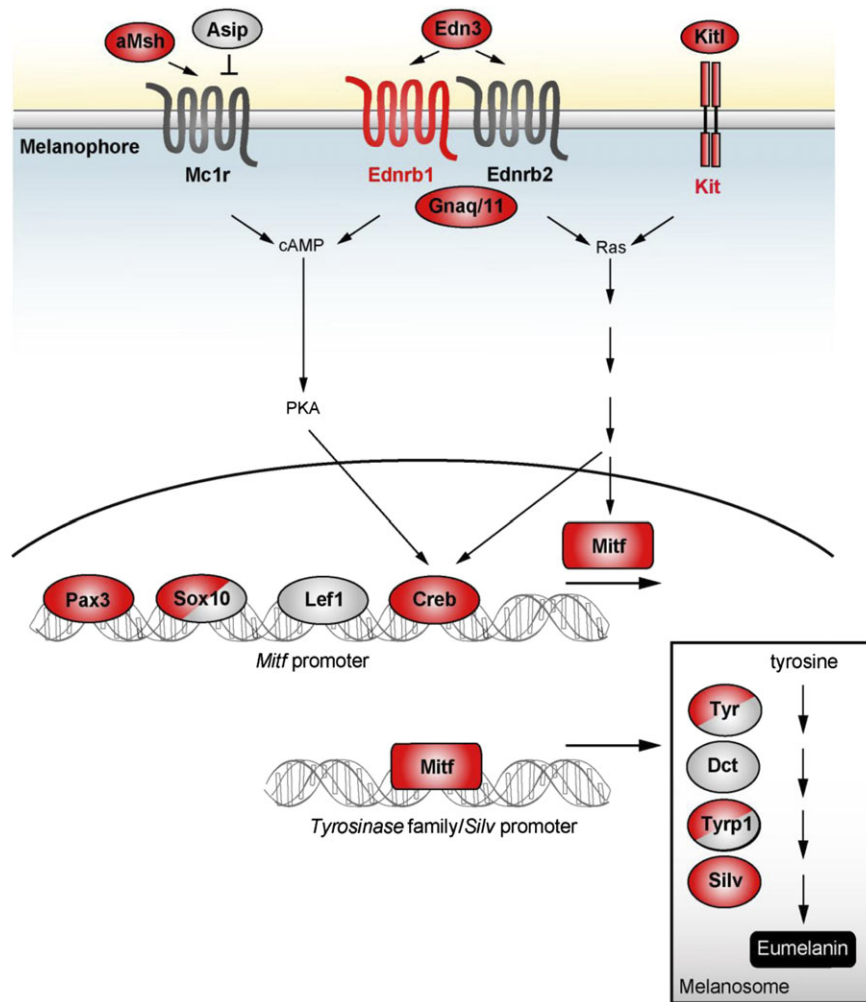


FIG. 4.—Impact of the FSGD on the melanocyte/phore signaling network. Many components have been identified by pigmentation mutants in both, mammals and teleosts. Diverse external signals are integrated on the promoter of the *Mitf* transcription factor gene, the master regulator of melanophore development (Béjar et al. 2003; Levy et al. 2006). *Mitf* regulates the expression of melanogenic enzymes, which catalyze the biosynthesis of melanin from tyrosine in melanosomes. The α MSH peptide is encoded by the *Pomc* gene. *Lef1* is a downstream target of the Wnt signaling pathway. Red indicates duplications as result of the FSGD present in all five teleost genomes analyzed and gray indicates singleton genes. Molecules with divided red/gray shading indicate genes retained in duplicate in some teleost lineages but singleton in others (see table 1). Most of the key regulators of the melanophore signaling network have been retained in two copies in teleosts. The signaling network is adapted from Tachibana et al. (2003), Levy et al. (2006), and Lin and Fisher (2007).

Almost nothing is known about the composition of the regulatory networks in other pigment cell types. The current knowledge of the xanthophore network is summarized in supplementary figure S26 (Supplementary Material online). Also here, large parts have been retained in two copies after the FSGD and differences between teleost lineages are apparent.

Discussion

Expansion of the Pigmentation Gene Repertoire in Teleosts as Result of the FSGD

The listing of vertebrate pigmentation genes assembled in the present study (supplementary table S1, Supplementary Material online) shows that in many cases orthologous genes are affected in pigmentation mutants in mammals, in birds, and also in teleosts. This points, once again, to the evolutionary conservation of the pigmentary

system, at least with respect to the development of melanophores/melanocytes, the only pigment cell type present in all vertebrate groups.

However, an important difference between tetrapods and teleosts is that the latter, as result of the FSGD, have an increased repertoire of genes potentially involved in pigmentation. Extant teleost genomes have retained around 30% of duplicated pigmentation genes from the FSGD. This FSGD duplicate retention rate for teleost pigmentation genes is higher than several genome-wide estimates for teleosts that range between 12% and 24% based on different approaches and data sets (Postlethwait et al. 2000; Jaillon et al. 2004; Woods et al. 2005; Brunet et al. 2006; Kassahn et al. 2009). Importantly, genome-wide FSGD duplicate retention does not differ significantly between the five teleost genomes (Kassahn et al. 2009).

We compared our data on pigmentation and liver genes to the genome-wide FSGD-duplicated retention rates

from Brunet et al. (2006) and Kassahn et al. (2009) that have used synteny and phylogenetic methods as well. This revealed that the FSGD duplicate retention rate is significantly higher than for the genome-wide estimates (supplementary table S4, Supplementary Material online). For example, 29.7% (38/128) of FSGD pigmentation gene duplicates have been retained in Tetraodon (fig. 2). In contrast, Brunet et al. (2006), which used the most comparable approach to ours to detect FSGD duplicates, reported a significantly lower (χ^2 test, $P = 1.7 \times 10^{-5}$) genome-wide FSGD duplicate retention rate in Tetraodon of 15.4% (364/2,371 analyzed genes). The genome-wide estimates for duplicate retention from Kassahn et al. (2009) are generally also significantly lower than for pigmentation genes (supplementary table S4, Supplementary Material online). Liver gene duplicates, in contrast, only have been retained in 12–16% of cases, which is similar or even lower than the genome-wide retention estimates (supplementary table S4, Supplementary Material online). Hence, pigmentation genes seem to have been preferentially retained in two copies following the FSGD. The FSGD therefore has generated a significant proportion of additional genes potentially involved in pigmentation.

Sato et al. (2009) recently reported FSGD duplicate retention rates for signal transduction and metabolic pathways in zebrafish, medaka, stickleback, and Tetraodon. Particularly, the retention rates for the signal transductions pathways involved in long-term potentiation of synaptic transmission, olfactory transduction, and taste transduction (25–35%) were higher than genome-wide estimates. These retention rates for signal transduction are in the range of pigmentation genes and do not significantly differ from them (supplementary table S4, Supplementary Material online). Thus, besides pigmentation, other pathways also show high rates of FSGD duplicate retention.

An important question is whether the high retention rate of duplicated pigmentation genes is due to the overrepresentation of certain functional categories among pigmentation genes. For example, enrichment for transcription factors, developmental genes, and cell communication proteins has been found for retained duplicates after the FSGD (Brunet et al. 2006; Kassahn et al. 2009) and the earlier vertebrate genome duplications (Putnam et al. 2008). Pigmentation genes are not enriched for transcription factor activity or cell communication. More general, developmental terms are enriched in pigmentation genes but likewise this also applies, for example, for liver genes (supplementary table S5, Supplementary Material online). This shows that although the enrichment of developmental terms among pigmentation genes may contribute to their retention, it is unlikely that this functional category alone is responsible for the expansion of the pigmentation gene repertoire after the FSGD. In fact, some specific groups of pigmentation genes, which are not enriched for developmental terms, have an exceptionally high FSGD duplicate retention rate (see below). It remains to be determined, if other properties like the number of gene functions or network connections could explain the differences in retention rate between pigmentation genes and other functional groups such as the liver genes.

Previous studies showed that there is broad variation in the enrichment/depletion of GO terms among FSGD du-

plicates (Brunet et al. 2006; Kassahn et al. 2009). Our data on functional categories of pigmentation genes (table 2) further suggest that FSGD duplicate retention rates can deviate considerably from the genome-wide estimates. In future studies, it will be thus necessary to further study in more detail FSGD gene duplicate retention and loss to identify functional groups, pathways, and networks with particularly high or low retention rates. Those pathways will then be interesting candidates to further study the impact of the FSGD for teleost evolution on the functional level.

Patterns of Pigmentation Gene Duplicate Loss and Retention

The reconstruction of pigmentation gene repertoire evolution after the FSGD revealed that, taking RGL into account, at least around 40% of FSGD-duplicated pigmentation genes have been retained in the last common euteleost ancestor of zebrafish, medaka, stickleback, and pufferfishes. Unfortunately, comparable genome-wide estimates are not available at present. Sato et al. (2009) estimated the FSGD duplicate retention rate for signal transduction pathway genes in the last euteleost ancestor to be 42.7%. For liver genes (fig. 3) and the metabolic tricarboxylic acid cycle (Sato et al. 2009), in contrast, the FSGD retention rate in the euteleost ancestor had been around 20% and 26%, respectively. Taken together, these data point to a comparatively low rate of duplicate pigmentation gene loss before the euteleost divergence.

But also euteleosts appear to have lost comparatively few pigmentation duplicates. According to previous estimates, only around 50% of genes retained in duplicate in zebrafish have also been retained in two copies in pufferfishes (Taylor et al. 2003; Woods et al. 2005). Similarly, for liver genes, this value is estimated around 57% (17/30). For pigmentation genes, in contrast, 80% (32/40 genes) of retained duplicates in zebrafish are also retained in pufferfishes. Thus, lineage-specific loss of gene duplicates seems to have been comparatively low for pigmentation genes. This pattern would be expected if the duplicated genes have kept only low redundancy levels and both have become essential before the divergence of euteleosts, precluding later gene loss. As the pigmentary innovations of the teleosts lineage are thought to predate the split of zebrafish and acanthomorphs (Braasch et al. 2008), this observation points to a possible phylogenetic correlation of the refinement of the teleost pigmentary system and the acquisition of important functions by duplicated pigmentation genes.

Pigmentation Genes Follow General Patterns of WGD Duplicate Retention

Our analyses suggest that there are important differences in the rate of FSGD duplicate retention among pigment cell types as well as among the different functional categories for a given cell type. Although many of the comparisons are statistically not significant due to the relative small set of genes analyzed for most categories, several interesting trends become apparent.

For example, almost all genes involved in melanosome biogenesis have been reverted to the single-gene status after the FSGD. Genes of this category encode proteins with general functions in the formation of melanosomes and other lysosome-related organelles. These proteins build large complexes that route other proteins to organelles including melanosomes (Bennett and Lamoreux 2003; Raposo and Marks 2007). Genes of this category seem to have been lost concertedly after the FSGD to avoid detrimental imbalances and to keep stoichiometry of complex units, following the predictions of the “gene balance hypothesis” (Papp et al. 2003). Furthermore, this group of genes is usually prone to the loss of duplicated genes because they are generally singleton genes in vertebrates, that is, reduction to the single-gene status has also occurred after the 1R/2R WGDs (data not shown). Theoretically, stoichiometry would also be maintained if all subunit duplicates had been retained after the FSGD. However, it has been previously shown for *Paramecium*, which has also undergone multiple rounds of WGD, that this is only a transient phase and complex constituents tend to return to the single-gene status after old WGDs (Aury et al. 2006). Here, this also is the case for 1R/2R/FSGD duplicates of melanosome biogenesis protein complex genes.

Sixty percent of genes encoding the components of melanosomes, that is, proteins present in mature melanosomes (Bennett and Lamoreux 2003), have retained both copies after the duplication in the FSGD. To this category belong mainly the enzymes involved in melanin synthesis. The components of melanosomes are specialized for pigimentary functions and are not enriched for developmental genes. The high retention rate of this group has been also observed in our previous analysis of teleost pigment synthesis pathways (Braasch et al. 2007). It might be related to subfunctionalization of paralogs for melanogenesis within the melanosomes of melanophores versus those of the RPE, for which there is some evidence from teleost pigmentation mutants (reviewed in Braasch et al. 2008). Furthermore, our functional analysis of *tyrp1* gene duplicates revealed that FSGD paralogs at least from this functional group diverge in a lineage-specific manner (Braasch, Liedtke, et al. 2009).

Genes involved in melanosome transport within the pigment cell also might be preferentially retained after the FSGD. Although this observation is based on four genes only, it might reveal important innovations in physiological color change. Physiological color change by moving melanosomes from the cell center to the periphery (darkening) or vice versa (lightening) is very important for many behavioral aspects and is highly sophisticated in teleosts compared with other vertebrates (Fujii 1993a; Fujii 2000).

Genes involved in the development of pigment cells are retained in duplicate at a high rate. This is the case for genes required for the development of melanophores as well as for xanthophore development. These two categories contain many transcription factors and signaling pathway components (supplementary table S1, Supplementary Material online). Retention and functional divergence of duplicated developmental regulatory genes would be expected if the FSGD has made an important contribution to teleost-specific features of pigment cell evolution such

as the partitioning of melanophores and xanthophores into distinct subpopulations that are under different genetic controls (Odenthal et al. 1996; Mellgren and Johnson 2002).

Most striking is the high retention rate for genes involved in xanthophore development with around 82%. This category consists mainly of genes for signaling and transcriptional control, reflecting the general high retention rate of development genes after the FSGD (Brunet et al. 2006). Furthermore, most of them (*Edn3*, *Ednrb1*, *Sox10*, *Pax3*, etc.) are involved in the development of more than one pigment cell type. Thus, it remains open whether the high FSGD retention rate for xanthophore developmental genes is indeed founded in xanthophore-specific characteristics. If so, one would expect also an exceptional high FSGD retention rate for xanthophore-specific genes, of which, unfortunately, only few have been identified so far. Interestingly, xanthophores are the evolutionary youngest pigment cell type that seems to have emerged in the gnathostome lineage after the divergence of agnathans. Melanophores and iridophores, in contrast, were already present in the vertebrate ancestor (Mellgren and Johnson 2002; Braasch et al. 2008).

In conclusion, the comparison of functional categories of pigmentation genes reveals that the retention of FSGD duplicates, being noticeably high, follows more general WGD duplicate retention patterns, that is, preservation of protein complex stoichiometries and overrepresentation of developmental genes among retained duplicates. Genes with multiple functions for pigment cell development also seem to be preferentially retained. Other studies also found multifunctionality to be an important determinant for duplicate retention after the FSGD (Sato et al. 2009) and the early vertebrate WGDs (Hufton et al. 2009).

Evolution of Teleost Pigmentation Pathways by Genome Duplication

Looking at duplicated pigmentation genes in the context of regulatory networks revealed that large parts of the melanophore and xanthophore pathways (fig. 4; supplementary fig. S26, Supplementary Material online) are present in two copies in teleosts. As all pigment cell types develop from a common precursor cell (Bagnara et al. 1979), a similar situation can be expected for the developmental pathways of other pigment cells.

According to a model recently developed by Freeling and Thomas (2006), increasing morphological complexity in the evolution of plants and animals is passively driven by recurrent WGDs. As result of WGDs, duplicated functional modules would accumulate in the gene pool and provide the genetic fundament for morphological innovations. For example, this may be the case when duplicated pathway components become specialized for sister cell types derived from a common evolutionary precursor (Arendt 2008). Despite these and other hypotheses based on theoretical models, the understanding of duplicated functional modules, pathways, and networks in multicellular organisms remains rudimentary. There are so far only few examples for WGD-derived functional module or pathway duplications: the starch biosynthesis pathway in rice (Wu et al. 2008), the C4 photosynthetic pathway in grasses (Wang et al. 2009),

the egesterol pathway in the fungus *Rhizopus oryzae* (Ma et al. 2009), and presumably the core network of cartilage development in vertebrates (Hecht et al. 2008). The present study suggests that the teleost pigmentation pathways represent further examples of WGD-duplicated functional modules. In contrast to the aforementioned studies, the pigmentation pathways analyzed here include in addition to the enzymes implicated in the biosynthesis of defined products such diverse gene categories as developmental genes for lineage specification and differentiation, subcellular organelle formation, intracellular transport, and homeostasis.

Duplications at different levels of organismal organization (e.g., gene network, cell type, tissue, body segment, behavioral program) are considered major sources of evolutionary novelties (West-Eberhard 2003). Oakley et al. (2007) proposed that “coduplication”, that is, the phylogenetic coincidence of gene duplications with the appearance of new structures should be considered as a null model for the origin of evolutionary novelties. In the present case, there seems to be a phylogenetic correlation between the pigmentary innovations such as leucophores that appear to predate the euteleost lineage (Braasch et al. 2008) and the generation of additional pigmentation genes by the FSGD, which is in line with the coduplication null model.

Relatively little is known about how the components of the pigment cell pathways have evolved after duplication with respect to functional divergence. For example, they could have evolved new functions (neofunctionalization) involved in teleost pigmentary novelties. Another evolutionary fate of functional module components could have been subfunctionalization of ancestral functions, as observed for the *mitf* paralogs in zebrafish (Lister et al. 2001; Altschmied et al. 2002). The ancestral functions in RPE and in melanophore development have been distributed with *mitfa* being expressed in melanophores and *mitfb* in the RPE (Lister et al. 2001). This functional specialization is also seen for the *silv* paralogs in zebrafish (Schonthaler et al. 2005) and the *tyrpl* duplicates in medaka (Braasch, Liedtke, et al. 2009). The expression of *silv* and *tyrpl* genes is under the transcriptional control of *Mitf* (fig. 4). Thus, it appears that at least parts of the duplicated functional module shown in figure 4 have become specialized for the melanophore or the RPE (for a review, see also Braasch et al. 2008).

More generally, however, the function of both paralogs as pigmentation genes remains to be demonstrated. In some cases, one of the two duplicates might have become specialized for the pigmentation function. This is seen for the Kit ligand–receptor pair. After FSGD, only *kitla* and *kita* are still involved in melanophore development, whereas *kitlb* and *kitb* have different, yet unknown functions not related to pigmentation (Mellgren and Johnson 2005; Hultman et al. 2007). Nevertheless, such cases can increase the evolvability of the pigmentary system because it releases it from functional constraints imposed by other essential functions of the ancestral gene.

Finally, our analysis also gives some insights into the potential contribution of the FSGD to the diversity of pigmentation among teleosts. There are subtle differences between teleost species in the retention of pathway component

duplicates (table 1 and fig. 4). Such modifications (e.g., loss of *sox10a* in zebrafish) might be the genomic basis of species-specific differences in pigment cell repertoires like the absence of leucophores from larval pigment patterns in the zebrafish. On top of that, lineage-specific divergence of retained pigmentation paralogs provides potential for further pigmentary divergence among teleosts.

Conclusions

Our study shows that in teleosts the increase of complexity in the pigmentary system is correlated with an increase in the repertoire of pigmentation genes. Modern teleosts have around 30% more pigmentation genes than tetrapods as result of FSGD. Large parts of the functional modules of pigment cell development and differentiation have been retained after the FSGD, presumably providing the genetic raw material for teleost-specific pigment cell innovations. Importantly, WGDs enable at the same time the duplication of all genes found in a given, fine-tuned pathway, generating a complete copy of the ancestral network upon which selection can act. Such duplication of entire pathways would hardly be possible by succeeding simple gene duplications. It has been proposed that genetic novelties arising from 1R/2R in early vertebrates might have enabled the evolution of major vertebrate innovations such as the endoskeleton or the neural crest and its derivatives (Ohno 1970; Shimeld and Holland 2000; Wada and Makabe 2006; Zhang and Cohn 2008). However, empirical data supporting these hypotheses are mainly restricted to few genes. Despite the fact that protein interaction is an important parameter for paralog evolution (Bridgham et al. 2008) and that network duplication might substantially contribute to cell type diversification (Arendt 2008), information on how developmental signaling cascades and regulatory networks have been influenced by WGDs is sparse. By the example of pigmentation pathways, the present study provides evidence for existence of duplicates of regulatory networks as result of the FSGD in teleosts, supporting an important role of WGDs for the phenotypic evolution in vertebrates.

The present study and the recent study by Sato et al. (2009) have presented a framework for future evolutionary investigation of other networks following WGD. For example, visual perception networks might also have been particularly prone for duplicate retention (Bowmaker 2008; Gojobori and Innan 2009) and could have evolved concomitantly to pigmentary innovations. Finally, our survey for pigmentation genes will be a helpful resource for ongoing studies using teleosts as models for pigmentary diseases and melanoma formation and investigating the evolution and development of pigmentation in colorful teleosts such as cichlids and guppies.

Supplementary Material

Supplementary tables S1–S5 and figures S1–S26 are available at *Genome Biology and Evolution* online (http://www.oxfordjournals.org/our_journals/gbe/).

Acknowledgments

We would like to thank the members of our laboratories for helpful discussions. This work was supported by grants from the German Science Foundation (DFG, to M.S. and J.N.V.) and by the Biofuture program of the Bundesministerium für Bildung und Forschung (BMBF) (to J.N.V.), the Association pour la Recherche contre le Cancer (ARC) (to J.N.V.), the French Institute for Agronomy Research (Institut National de Recherche Agronomique PHASE) (to J.N.V.), the French Research Agency (ANR) (to J.N.V.), and the Foundation pour la Recherche Médicale (FRM) (to J.N.V.).

Literature Cited

- Abascal F, Zardoya R, Posada D. 2005. ProfTest: selection of best-fit models of protein evolution. *Bioinformatics*. 21:2104–2105.
- Aburomia R, Khaner O, Sidow A. 2003. Functional evolution in the ancestral lineage of vertebrates or when genomic complexity was wagging its morphological tail. *J Struct Funct Genomics*. 3:45–52.
- Altschmied J, et al. 2002. Subfunctionalization of duplicate *mitf* genes associated with differential degeneration of alternative exons in fish. *Genetics*. 161:259–267.
- Amores A, et al. 1998. Zebrafish *hox* clusters and vertebrate genome evolution. *Science*. 282:1711–1714.
- Arendt D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet*. 9:868–882.
- Aury JM, et al. 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature*. 444:171–178.
- Bagnara JT, Fernandez PJ, Fujii R. 2007. On the blue coloration of vertebrates. *Pigment Cell Res*. 20:14–26.
- Bagnara JT, Matsumoto J. 2006. Comparative anatomy and physiology of pigment cells in nonmammalian tissues. In: Nordlund JJ, Boissy RE, Hearing VJ, King RA, Ortonne J-P, editors. *The pigimentary system: physiology and pathophysiology*. New York: Oxford University Press. p. 9–40.
- Bagnara JT, et al. 1979. Common origin of pigment cells. *Science*. 203:410–415.
- Bastian BC, Pinkel D. 2008. Expanding the genetic spectrum of pigmentation. *Pigment Cell Melanoma Res*. 21:507–508.
- Béjar J, Hong Y, Scharlt M. 2003. *Mitf* expression is sufficient to direct differentiation of medaka blastula derived stem cells to melanocytes. *Development*. 130:6545–6553.
- Bennett DC, Lamoreux ML. 2003. The color loci of mice—a genetic century. *Pigment Cell Res*. 16:333–344.
- Blomme T, et al. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. *Genome Biol*. 7:R43.
- Bowmaker JK. 2008. Evolution of vertebrate visual pigments. *Vision Res*. 48:2022–2041.
- Braasch I, Liedtke D, Volff JN, Scharlt M. 2009. Pigmentary function and evolution of *tyrp1* gene duplicate in fish. *Pigment Cell Melanoma Res*. 22:839–850.
- Braasch I, Salzburger W, Meyer A. 2006. Asymmetric evolution in two fish-specifically duplicated receptor tyrosine kinase paralogs involved in teleost coloration. *Mol Biol Evol*. 23:1192–1202.
- Braasch I, Scharlt M, Volff JN. 2007. Evolution of pigment synthesis pathways by gene and genome duplication in fish. *BMC Evol Biol*. 7:74.
- Braasch I, Volff JN, Scharlt M. 2008. The evolution of teleost pigmentation and the fish-specific genome duplication. *J Fish Biol*. 73:1891–1918.
- Braasch I, Volff JN, Scharlt M. 2009. The endothelin system: evolution of vertebrate-specific ligand-receptor interactions by three rounds of genome duplication. *Mol Biol Evol*. 26:783–799.
- Brenner S, et al. 1993. Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature*. 366:265–268.
- Bridgham JT, Brown JE, Rodriguez-Mari A, Catchen JM, Thornton JW. 2008. Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. *PLoS Genet*. 4:e1000191.
- Brunet FG, et al. 2006. Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol Biol Evol*. 23:1808–1816.
- Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*. 134:25–36.
- De Bodt S, Maere S, Van de Peer Y. 2005. Genome duplication and the origin of angiosperms. *Trends Ecol Evol*. 20:591–597.
- Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*. 3:e314.
- de Souza FS, Bumashny VF, Low MJ, Rubinstein M. 2005. Subfunctionalization of expression and peptide domains following the ancient duplication of the proopiomelanocortin gene in teleost fishes. *Mol Biol Evol*. 22:2417–2427.
- Donoghue PCJ, Purnell MA. 2005. Genome duplication, extinction and vertebrate evolution. *Trends Ecol Evol*. 20:312–319.
- Dutton KA, et al. 2001. Zebrafish *colourless* encodes *sox10* and specifies non-ectomesenchymal neural crest fates. *Development*. 128:4113–4125.
- Force A, et al. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 151:1531–1545.
- Freeling M, Thomas BC. 2006. Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Res*. 16:805–814.
- Froschauer A, Braasch I, Volff JN. 2006. Fish genomes, comparative genomics and vertebrate evolution. *Curr Genomics*. 7:43–57.
- Fujii R. 1993a. Cytophysiology of fish chromatophores. *Int Rev Cytol*. 143:191–255.
- Fujii R. 1993b. Coloration and chromatophores. In: Evans DH, editor. *The physiology of fishes*. Boca Raton (FL): CRC Press. p. 535–562.
- Fujii R. 2000. The regulation of motile activity in fish chromatophores. *Pigment Cell Res*. 13:300–319.
- Fukamachi S, Meyer A. 2007. Evolution of receptors for growth hormone and somatolactin in fish and land vertebrates: lessons from the lungfish and sturgeon orthologues. *J Mol Evol*. 65:359–372.
- Furlong RF, Holland PW. 2002. Were vertebrates octoploid? *Philos Trans R Soc Lond B Biol Sci*. 357:531–544.
- Gojobori J, Innan H. 2009. Potential of fish opsin gene duplications to evolve new adaptive functions. *Trends Genet*. 25:198–202.
- Gomez A, Volff JN, Hornung U, Scharlt M, Wellbrock C. 2004. Identification of a second *egfr* gene in *Xiphophorus* uncovers an expansion of the epidermal growth factor receptor family in fish. *Mol Biol Evol*. 21:266–275.
- Guindon S, Lethiec F, Duroux P, Gascuel O. 2005. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res*. 33:W557–W559.
- Haffter P, et al. 1996. Mutations affecting pigmentation and shape of the adult zebrafish. *Dev Genes Evol*. 206:260–276.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 41:95–98.

- Hecht J, et al. 2008. Evolution of a core gene network for skeletogenesis in chordates. *PLoS Genet.* 4:e1000025.
- Holland PW, Garcia-Fernandez J, Williams NA, Sidow A. 1994. Gene duplications and the origins of vertebrate development. *Dev Suppl.* 125–133.
- Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 4:44–57.
- Hufton AL, et al. 2009. Deeply conserved chordate non-coding sequences preserve genome synteny but do not drive gene duplicate retention. *Genome Res.* 19:2036–2051.
- Hultman KA, Bahary N, Zon LI, Johnson SL. 2007. Gene duplication of the zebrafish *kit ligand* and partitioning of melanocyte development functions to *kit ligand a*. *PLoS Genet.* 3:e17.
- Jaillon O, et al. 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature.* 431:946–957.
- Kasahara M, et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature.* 447:714–719.
- Kassahn KS, Dang VT, Wilkins SJ, Perkins AC, Ragan MA. 2009. Evolution of gene function and regulatory control after whole-genome duplication: comparative analyses in vertebrates. *Genome Res.* 19:404–418.
- Kelsh RN, et al. 1996. Zebrafish pigmentation mutations and the processes of neural crest development. *Development.* 123:369–389.
- Kelsh RN, et al. 2004. The Tomita collection of medaka pigmentation mutants as a resource for understanding neural crest cell development. *Mech Dev.* 121:841–859.
- Levy C, Khaled M, Fisher DE. 2006. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med.* 12:406–414.
- Lin JY, Fisher DE. 2007. Melanocyte biology and skin pigmentation. *Nature.* 445:843–850.
- Lister JA, Close J, Raible DW. 2001. Duplicate *mitf* genes in zebrafish: complementary expression and conservation of melanogenic potential. *Dev Biol.* 237:333–344.
- Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science.* 290:1151–1155.
- Lynch M, Force A. 2000. The origin of interspecific genomic incompatibility via gene duplication. *Am Nat.* 156:590–605.
- Ma LJ, et al. 2009. Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. *PLoS Genet.* 5:e1000549.
- Maere S, et al. 2005. Modeling gene and genome duplications in eukaryotes. *Proc Natl Acad Sci USA.* 102:5454–5459.
- Mellgren EM, Johnson SL. 2002. The evolution of morphological complexity in zebrafish stripes. *Trends Genet.* 18:128–134.
- Mellgren EM, Johnson SL. 2005. *kitb*, a second zebrafish ortholog of mouse *Kit*. *Dev Genes Evol.* 215:470–477.
- Meyer A, Van de Peer Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays.* 27:937–945.
- Michiels NK, et al. 2008. Red fluorescence in reef fish: a novel signalling mechanism? *BMC Ecol.* 8:16.
- Minchin JE, Hughes SM. 2008. Sequential actions of Pax3 and Pax7 drive xanthophore development in zebrafish neural crest. *Dev Biol.* 317:508–522.
- Miwa M, et al. 2007. *Endothelin receptor B2 (EDNRB2)* is associated with the panda plumage colour mutation in Japanese quail. *Anim Genet.* 38:103–108.
- Nelson JS. 2006. *Fishes of the world*. New York: John Wiley & Sons.
- Oakley TH, Plachetzki DC, Rivera AS. 2007. Furcation, field-splitting, and the evolutionary origins of novelty in arthropod photoreceptors. *Arthropod Struct Dev.* 36:386–400.
- Odenthal J, et al. 1996. Mutations affecting xanthophore pigmentation in the zebrafish, *Danio rerio*. *Development.* 123:391–398.
- Ohno S. 1970. *Evolution by gene duplication*. New York: Springer-Verlag.
- Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell.* 131:452–462.
- Papp B, Pal C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. *Nature.* 424:194–197.
- Postlethwait JH, et al. 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res.* 10:1890–1902.
- Protas ME, Patel NH. 2008. Evolution of coloration patterns. *Annu Rev Cell Dev Biol.* 24:425–446.
- Putnam NH, et al. 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature.* 453:1064–1071.
- Raposo G, Marks MS. 2007. Melanosomes—dark organelles enlighten endosomal membrane transport. *Nat Rev Mol Cell Biol.* 8:786–797.
- Rawls JF, Mellgren EM, Johnson SL. 2001. How the zebrafish gets its stripes. *Dev Biol.* 240:301–314.
- Sato Y, Hashiguchi Y, Nishida M. 2009. Temporal pattern of loss/persistence of duplicate genes involved in signal transduction and metabolic pathways after teleost-specific genome duplication. *BMC Evol Biol.* 9:127.
- Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature.* 440:341–345.
- Scannell DR, et al. 2007. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. *Proc Natl Acad Sci USA.* 104:8397–8402.
- Schonthaler HB, et al. 2005. A mutation in the silver gene leads to defects in melanosome biogenesis and alterations in the visual system in the zebrafish mutant fading vision. *Dev Biol.* 284:421–436.
- Semon M, Wolfe KH. 2007. Reciprocal gene loss between *Tetraodon* and zebrafish after whole genome duplication in their ancestor. *Trends Genet.* 23:108–112.
- Semon M, Wolfe KH. 2008. Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. *Proc Natl Acad Sci USA.* 105:8333–8338.
- Setiamarga DH, et al. 2008. Interrelationships of Atherinomorpha (medakas, flyingfishes, killifishes, silversides, and their relatives): the first evidence based on whole mitogenome sequences. *Mol Phylogenet Evol.* 49:598–605.
- Shimeld SM, Holland PW. 2000. Vertebrate innovations. *Proc Natl Acad Sci USA.* 97:4449–4452.
- Siegel N, Hoegg S, Salzburger W, Braasch I, Meyer A. 2007. Comparative genomics of ParaHox clusters of teleost fishes: gene cluster breakup and the retention of gene sets following whole genome duplications. *BMC Genomics.* 8:312.
- Tachibana M, Kobayashi Y, Matsushima Y. 2003. Mouse models for four types of Waardenburg syndrome. *Pigment Cell Res.* 16:448–454.
- Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol.* 24:1596–1599.
- Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y. 2003. Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res.* 13:382–390.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

- Wada H, Makabe K. 2006. Genome duplications of early vertebrates as a possible chronicle of the evolutionary history of the neural crest. *Int J Biol Sci.* 2:133–141.
- Wang X, et al. 2009. Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biol.* 10:R68.
- Wapinski I, Pfeffer A, Friedman N, Regev A. 2007. Natural history and evolutionary principles of gene duplication in fungi. *Nature.* 449:54–61.
- Werth CR, Windham MD. 1991. A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. *Am Nat.* 137:515–526.
- West-Eberhard MJ. 2003. *Developmental plasticity and evolution.* New York: Oxford University Press.
- Woods IG, et al. 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res.* 15:1307–1314.
- Wu Y, Zhu Z, Ma L, Chen M. 2008. The preferential retention of starch synthesis genes reveals the impact of whole-genome duplication on grass evolution. *Mol Biol Evol.* 25:1003–1006.
- Yan YL, et al. 2005. A pair of Sox: distinct and overlapping functions of zebrafish *sox9* co-orthologs in craniofacial and pectoral fin development. *Development.* 132:1069–1083.
- Zhang G, Cohn MJ. 2008. Genome duplication and the origin of the vertebrate skeleton. *Curr Opin Genet Dev.* 18:387–393.
- Zhu Y, et al. 2004. Cloning of *somatolactin alpha* and *beta* cDNAs in zebrafish and phylogenetic analysis of two distinct somatolactin subtypes in fish. *J Endocrinol.* 182:509–518.
- Ziegler I. 2003. The pteridine pathway in zebrafish: regulation and specification during the determination of neural crest cell fate. *Pigment Cell Res.* 16:172–182.

Kenneth Wolfe, Associate Editor

Accepted November 23, 2009