Systematic Variation in the Pattern of Gene Paralog Retention between the Teleost Superorders Ostariophysi and Acanthopterygii

Daniel Garcia de la serrana1,*, Edson A. Mareco2, and Ian A. Johnston1

1School of Biology, Scottish Oceans Institute, University of St Andrews, Fife, United Kingdom
2Department of Morphology, Institute of Biosciences of Botucatu, UNESP—Univ Estadual Paulista, São Paulo, Brazil

*Corresponding author: E-mail: dgdlsc@st-andrews.ac.uk.

Accepted: April 7, 2014

Abstract

Teleost fish underwent whole-genome duplication around 450 Ma followed by diploidization and loss of 80–85% of the duplicated genes. To identify a deep signature of this teleost-specific whole-genome duplication (TSGD), we searched for duplicated genes that were systematically and uniquely retained in one or other of the superorders Ostariophysi and Acanthopterygii. TSGD paralogs comprised 17–21% of total gene content. Some 2.6% (510) of TSGD paralogs were present as pairs in the Ostariophysi genomes of *Danio rerio* (Cypriniformes) and *Astyanax mexicanus* (Characiformes) but not in species from four orders of Acanthopterygii (Gasterosteiformes, *Gasterosteus aculeatus*; Tetraodontiformes, *Tetraodon nigroviridis*; Perciformes, *Oreochromis niloticus*; and Beloniformes, *Oryzias latipes*) where a single copy was identified. Similarly, 1.3% (418) of total gene number represented cases where TSGD paralogs pairs were systematically retained in the Acanthopterygian but conserved as a single copy in Ostariophysi genomes. We confirmed the generality of these results by phylogenetic and synteny analysis of 40 randomly selected lineage-specific paralogs (LSPs) from each superorder and completed with the transcriptomes of three additional Ostariophysi species (*Ictalurus punctatus* [Siluriformes], *Sinocyclocheilus* species [Cypriniformes], and *Piaractus mesopotamicus* [Characiformes]). No chromosome bias was detected in TSGD paralog retention. Gene ontology (GO) analysis revealed significant enrichment of GO terms relative to the human GO SLIM database for “growth,” “Cell differentiation,” and “Embryo development” in Ostariophysi and for “Transport,” “Signal Transduction,” and “Vesicle mediated transport” in Acanthopterygii. The observed patterns of paralog retention are consistent with different diploidization outcomes having contributed to the evolution/diversification of each superorder.

Key words: fish evolution, ploidy, gene loss, whole-genome duplication.

Introduction

Polyploidy, involving whole-genome duplication (WGD) and the doubling of gene content, is considered a major feature of the evolution of eukaryotic genomes (Taylor et al. 2003). WGD is usually followed by diploidization and the loss of gene paralogs, a process that may occur over a protracted period (Brunet et al. 2006; Kasahara et al. 2007). Signatures of ancient polyploidy events are evident in many eukaryotic genomes (Jiao et al. 2011; Zhan et al. 2014). For example, the ancestral genome of vertebrates is thought to have undergone two consecutive rounds (1R/2R) of WGD (Dehal and Boore 2005), with a third round (3R) in the lineage leading to teleost fish (Taylor et al. 2003; Jaillon et al. 2004). WGD at the base of the teleost fish radiation (teleost-specific whole-genome duplication [TSGD]) was estimated at 450–320 Ma (3R) (Kuraku and Meyer 2009; Sato and Nishida 2010). It is thought that around 15% of TSGD paralogs have been retained in the diploid genome of modern species (Braasch and Postlethwaite 2012). Several mechanisms have been suggested to explain the retention of paralogs after WGD or small-scale duplications including the appearance of mutations leading to altered regulation (subfunctionalization) and/or the evolution of some novel function (neofunctionalization), which confers a selective advantage (Maere and Van de Peer 2010). Thus, polyploidy contributes to an increase in gene content and at some level has likely contributed to the evolutionary success of modern day taxa. For example, it has been argued that WGD promotes...
speciation via divergent resolution where the loss of different copies of duplicated genes in allopatric populations leads to genetic isolation (Taylor et al. 2001).

However, the extent to which specific polyploidy events contribute to evolutionary success and speciation is a matter of long-standing debate. For example, it has been estimated that approximately 88% of teleost species are of recent origin, such that the TSGD event may explain as little as 10% of the total diversity (Santini et al. 2009). Similarly, a well-constrained estimate of the salmonid WGD (4R) placed it at 88 Ma, whereas the subfamilies emerged 40–50 Ma and 50% of species formed within the last 5 Ma (Macqueen and Johnston 2014). These results at least indicate a major decoupling between WGD and species diversification while not excluding long-lasting effects of the ploidy event.

Present day examples of polyploidy are particularly widespread in the plant kingdom (Bowers et al. 2003; Jiao et al. 2011). Polyploid lineages are also relatively common in teleosts (Zhan et al. 2014) and have been reported in some amphibians and reptiles (Mable et al. 2011). Phylogenetic studies have shown that recently formed plant lineages experience lower diversification rates relative to diploid congeners as a consequence of both lower speciation and extinction rates (Zhan et al. 2014). In contrast, using comparable methods in teleost, similar diversification rates have been found between polyploid and diploid relatives in some cases (Acipenseridae, Botiidae [families], Salmoniformes [order]), whereas the subfamily Cyprininae revealed higher polyploid diversification [Zhan et al. 2014]).

This study aimed to exploit the recent increase in teleost genome and large-scale transcriptomic data sets to provide an insight into the role of ancient polyploidy on subsequent diversification of teleosts. We tested the hypothesis that different diploidization outcomes have occurred between two of the main teleost superorders: The Ostariophysi and Acanthopterygi by searching for a systematic difference in the retention of TSGD gene paralogs. We further investigated whether there were either chromosomal or functional biases in the retained paralogs between lineages.

Results and Discussion

A previous comprehensive phylogenetic analysis using 42 orthologous nuclear protein-coding genes estimated that the split of the Euteleostei superfamilies the Ostariophysi and Acanthopterygii took place in the early Triassic 217 Ma (Steinke et al. 2006). In this study, a comparison of Acanthopterygii and Ostariophysi proteomes (see Materials and Methods and supplementary fig. S1, Supplementary Material online) revealed that 21% of the total genes analyzed (4,122 out of 19,600) in the Ostariophysi superfamilies and 18.4% (3,284 out of 17,800) in the Acanthopterygii superfamilies were present as TSGD paralogs, in agreement with previous estimates (Braasch and Postlethwaite 2012).

The superorder Ostariophysi comprises five orders (Gonorynchiformes, Cypriniformes, Characiformes, Siluriformes, and Gymnotiformes) containing 6,507 species (Nelson 2006). The Ostariophysi are characterized by the Weberian apparatus consisting of modified vertebrae, which connect and transmit sound waves from the swim bladder to the inner ear to increase hearing sensitivity. Genome sequences are only currently available for the zebrafish (Danio rerio) and the blind cave fish (Astyanax mexicanus) belonging to the orders Cypriniformes and Characiformes, respectively (Steinke et al. 2006). Gene orthologs were identified that occurred as paralogous pairs on different chromosomes in Ostariophysi but were present as singletons in four species of Acanthopterygii from different orders (Gasterosteiformes, Gasterosteus aculeatus; Tetraodontiformes, Tetraodon nigroviridis; Perciformes, Oreochromis niloticus; and Beloniformes, Onychias latipes). This yielded a list of 205 candidate orthologs present as 510 TSGD paralogs (2.6% of total genes analyzed) in Ostariophysi but not Acanthopterygii species (fig. 1A; supplementary file S2, Supplementary Material online). Those cases in which both paralogs were retained in one lineage but a single copy in the other were considered potential lineage-specific paralogs (LSPs). To further investigate the possibility that these orthologs have been retained as paralogous throughout the suborder, we carried out phylogenetic and synteny analysis on a subset of 40 randomly selected Ostariophysi LSPs. The phylogenetic analysis was completed with orthologs from three further Ostariophysi species the catfish Ictalurus punctatus (order Siluriformes), Sinocyclocheilus species (Cypriniformes) retrieved from the National Center for Biotechnology Information (NCBI) transcriptome database, (www.ncbi.nlm.nih.gov, last accessed March 15, 2014, Transcriptomic Shotgun Assembly), and the pacu Piaractus mesopotamicus (Characiformes) (Mareco EA et al., unpublished data) (supplementary files S3 and S4, Supplementary Material online). Although only three of the five Ostariophysi orders were examined, all the 40 selected orthologous were present as paralog pairs in basal and more derived species, consistent with their retention throughout the superorder (fig. 2). Using a similar rational, we found that orthologs of 113 genes representing 226 TSGD paralogs (1.3% of total gene content) were systematically retained in all Acanthopterygii genomes tested (Oreochromis, Tetraodon, Orezias, and Gasterosteus) but as a single copy in the two Ostariophysi genomes available (Danio and Astyanax). The Acanthopterygii sampled included Perciformes and Beloniformes, which split 113 Ma and last shared a common ancestor with the Tetraodontiformes 195 Ma (Steinke et al. 2006). Nevertheless, the result for the Acanthopterygii superorder is less robust than for the Ostariophysi because we only sampled 4 of the 13 orders (Nelson 2006). Similarly, phylogenetic and synteny analysis was carried out in a subset of 40 random Acanthopterygii LSPs (fig. 3). LSPs identified in Tetraodon, Gasterosteus, and Danio
chromosomes were proportional to the number of TSGD-paralogs analyzed per chromosome (fig. 1B; $r = 0.632; P = 0$). This result indicates that the putative superorder-specific paralogs were not retained on specific chromosomes or had originated from chromosome-specific rearrangements.

To gain an insight into any potential functional consequences of these difference paralog retention patterns, we performed a gene ontology (GO) SLIM enrichment analysis for the Ostariophysi and Acanthopterygii LSPs relative to the human GO SLIM database. A significant enrichment in GO terms related to “Development,” “Growth,” and “Cell differentiation” was found in the Ostariophysi, whereas the Acanthopterygii showed significant differences in “Signal transduction,” “Transport,” and the “Vesicle mediated transport” (table 1 and supplementary file S5, Supplementary Material online).

A further inspection of the gene lists in table 1 allowed us to identify some of the individual genes within the GO terms that were significantly different for each superorder. Ostariophysi species have retained two copies of key transcription factors involved in development including members of the Hox gene family ($\text{hoxc6}$, $\text{hoxc11}$, $\text{hoxc12}$, and $\text{hoxc13}$) involved in patterning (Mallo et al. 2010), Sox gene family members ($\text{sox1}$, $\text{sox19}$ and $\text{sox21}$) with diverse developmental functions (Sarkar and Hochedlinger 2013), and six family members ($\text{six1}$, $\text{six2}$ and $\text{six4}$) involved in DNA-binding specificity and in mediating protein–protein interactions (Kumar 2009). In all these cases, only a single TSGD paralog was retained in the Acanthopterygii genomes analyzed. Similarly, Ostariophysi have retained duplicated genes from the PI3K/IGF/mTOR pathway ($\text{rictor}$, $\text{rps6ka3}$, $\text{igf2}$, and $\text{igf2bp2}$), which is involved in growth and protein synthesis (reviewed in Johnston et al. 2011). In contrast, Acanthopterygii have retained two copies of some Rab GTPases ($\text{rab9a}$, $\text{rab19}$, $\text{rab27b}$, $\text{rab5c}$, and $\text{rab8a}$), which have a role in membrane trafficking including vesicle formation and movement and membrane fusion (McCormick 1995).

In summary, we provide evidence for systematic differences in TSGD paralog retention between the teleost superorders Ostariophysi and Acanthopterygii of the order of 1–2% of gene content. The scale of these differences and preliminary GO analysis indicate a persistent signature of the TSGD event.

**Fig. 1.**—Ostariophysi- and Acanthopterygii-LSP retention and chromosome distribution. (A) Percentage of LSP retained over the total of TSGD paralogs analyzed in Gasterosteus aculeatus ($n = 21$ chromosomes), Tetraodon nigroviridis ($n = 21$ chromosomes), and Danio rerio ($n = 25$ chromosomes). Values represent average of chromosomes LSPs ± standard error. (B) Correlation plot between number of TSGD paralogs in each $D. \text{rerio}$ (empty circles), $T. \text{nigroviridis}$ (filled circles), and $G. \text{aculeatus}$ (crossed circles) chromosome against the number of LSP identified in the same chromosome; Spearman correlation ($\rho$) and statistical significance are shown.
that may be of functional significance for the subsequent evolution/diversification of each lineage. Our results are consistent with a lingering influence of the TSGD on speciation. Continuous advances in sequencing technology will increase the number and diversity of genomes available enabling further testing of the hypothesis of large-scale conservation of paralog retention between different branches of the ray-finned teleost radiation.

**Conclusions**

This study has shown that some TSGD paralogs have been systematically retained in Acanthopterygii but Ostariophysi superorders of teleosts (1.3% and 2.6% of total gene content, respectively). We also showed that LSPs are randomly distributed in teleost genomes, but there were significant differences in the retention of key genes related to growth and embryonic development between the superorders, which may have influenced their subsequent evolution.

**Materials and Methods**

**Identification of LSPs**

The method for identifying TSGD paralogs that have been systematically retained as pairs in one superorder but as a single copy in the other is schematically illustrated in supplementary figure S1, Supplementary Material online. Among fish with sequenced genomes, *D. rerio* and *G. aculeatus* are the Acanthopterygii and Ostariophysi species, respectively, that have the highest numbers of annotated gene sequences. The first step in our analysis involved reciprocal BLASTs of the...
proteomes from *D. rerio* (www.ensembl.org, last accessed March 15, 2014; vZv9) and *G. aculeatus* (www.ensembl.org, last accessed March 15, 2014; v.BROADS1) using the BLASTp algorithm included in BioEdit software (http://www.mbio.ncsu.edu/bioedit/bioedit.html, last accessed April 22, 2014) with an e-value cutoff of $E^{-80}$. A total of 19,600 and 17,800 positive hits were obtained from Ostariophysi/Acanthopterygii and Acanthopterygii/Ostariophysi comparisons, respectively. Those genes annotated as TSGD paralogs, and their ortholog from each of the comparisons, were manually retrieved from both lists based on their Ensembl annotation (www.ensembl.org, last accessed March 15, 2014). Putative TSGD paralogs were aligned using ClustalW to verify that chimeras, splice variants, or isoforms were excluded from the analysis. The remaining duplicated sequences that met the twin criteria of occurring on different chromosomes and existing as a single ortholog in *Lepisosteus oculatus* genome (a pre-TSGD teleost; www.ensembl.org, last accessed March 15, 2014; vLepocu1) and in human (*Homo sapiens* genome; www.ensembl.org, last accessed March 15, 2014; vGRCh37.p13) were considered to be genuine TSGD paralogs.

To identify those duplicates that were systematically retained in Ostariophysi superorder, *D. rerio* paralogs with a single best hit against the same *G. aculeatus* (Gasterosteiformes) ortholog were retrieved. To identify genes present as duplicates in other Ostariophysi and singletons in Acanthopterygii, the *D. rerio* gene list was blasted against the *A. mexicanus* (vAstmex102) (Characiformes) and three more Acanthopterygii genomes (*O. latipes* [Beloniformes; FIG.3.—Phylogenetic (A, B) and synteny (C) analysis for LSP from Ostariophysi species. (A) Phylogenetic relationships for the inhibitor of growth protein 5 (ing5) gene. Tree node values represent posterior values. (B) Maximum likelihood phylogenetic relationships for ing5. Phylogenetic trees nodes values represent posterior values. (C) Synteny of the Ostariophysi LSP of ing5 across teleost species. Genes are indicated as colored boxes, and orthologs share the same color. To aid interpretation, ing5 orthologs were aligned and are highlight in red.
v.HdrR, www.ensembl.org, last accessed March 15, 2014],
T. nigroviridis [Tetraodontiformes; v.TETRAODON8.0, www.
ensembl.org, last accessed March 15, 2014],
O. niloticus [Perciformes; v.Orenil1.0, www.ensembl.org, last accessed
March 15, 2014]). Those TSGD with two orthologs in
Astyanax and Danio, but a single copy in all four
Acanthopterygii genomes, were considered as Ostariophysi
LSPs. Phylogenetic and synteny analysis was carried out
using 40 randomly selected LSPs from the Ostariophysi su-
perorder. Because only two Ostariophysi genomes are avail-
able, transcriptomic data from representative species from
three other Ostariophysi species were used to increase the
power of the analysis (I. punctatus [Siluriformes] [www.ncbi.
nlm.nih.gov, last accessed March 15, 2014],
Sinocyclocheilus [Cypriniformes] [www.ncbi.nlm.nih.gov, last accessed March
15, 2014], and P. mesopotamicus [Characiformes] [Mareco EA
et al., unpublished data]). In some cases, it was not possible
to include data of all three Ostariophysi-species due to limitations
in the transcriptomic database. Transcriptomes are based on
expressed genes present in an organism in a specific physio-
logical stage, which means that lowly expressed genes are
often missed (as an example see Garcia de la Serrana et al.
2012). However, for all phylogenetic trees generated, there
was at least one species present from each of the three
Ostariophysi orders. All the amino acid sequences used for
phylogenetic analysis are provided in supplementary file S6,
Supplementary Material online. To identify those genes that
occurred as duplicates in Acanthopterygii but singletons in
Ostariophysi, we filtered the TSGD paralogs from G. aculeatus
against successive rounds of BLAST against the genomes of
Acanthopterygians T. nigroviridis, Ory. latipes, O. niloticus, and
the Ostariophysi A. mexicanus. Those TSGD paralogs that
were present as duplicates in species from all four orders
but single copy in Astyanax and Danio were considered as
Acanthopterygii LSPs. Similarly, a subset of 40 randomly se-
lected LSPs were used for phylogenetic and synteny analysis.
Similarly, phylogenetic analysis was completed with transcrip-
tomic data from Ictalurus, Piaractus, and Sinocyclocheilus.

### Phylogenetic Analysis

Peptides sequences were aligned using the GUIDANCE online
server (Penn et al. 2010) with PRANK as multisequence align-
ment algorithm. Columns below the 0.93 Guidance score
cutoff were removed from the final alignment used for the
phylogenetic trees construction (all alignments are provided in
supplementary file S7, Supplementary Material online).
Bayesian MCMC phylogenetic trees, following a Yule specia-
tion process model and UPGMA starting tree, were gener-
ated for each alignment using BEAST v1.7.5 software
(Drummond et al. 2012). Guidance alignments were also used to construct maximum
likelihood (ML) phylogenetic trees for each of the LSP ana-
yzed. ML trees were constructed using PhyML online server
(http://www.atgc-montpellier.fr/phyml/, last accessed March
12, 2014) (Dereeper et al. 2008). The best evolutionary
model for each alignment used to calculate the phylogenetic
trees was determined by MEGAS software (Tamura et al.
2010). Final Bayesian trees were generated using TreeAnnotator v1.7.5 with a burnin value of 1,000. All trees
were visualized using FigTree v1.3.1.

### Table 1

| GO SLIM Term Enriched | GO ID      | P-value  | Number of Genes |
|-----------------------|------------|----------|-----------------|
| **Biological process**|            |          |                 |
| Embryo development     | 0009790    | 2.7E−4   | 24              |
| Growth                 | 0040007    | 1.7E−2   | 10              |
| Anatomical structure development | 0048856 | 2.3E−2   | 55              |
| Cell differentiation   | 0030154    | 3.0E−2   | 40              |
| **Molecular function**|            |          |                 |
| DNA binding            | 0003677    | 8.4E−3   | 42              |
| Nucleic acid binding transcription factor | 0001071 | 1.0E−2   | 21              |
| **Cell component**     |            |          |                 |
| Nucleus                | 0005634    | 6.13E−3  | 87              |
| Cytosol                | 0005829    | 2.3E−2   | 40              |
| **Biological process**|            |          |                 |
| Signal transduction    | 0007165    | 3.3E−3   | 32              |
| Vesicle-mediated transport | 0016192 | 8.2E−3   | 11              |
| Anatomical structure development | 0048856 | 1.3E−2   | 29              |
| Response to stress     | 0006950    | 1.5E−2   | 24              |
| Transport              | 0006810    | 4.9E−2   | 22              |
| **Cell component**     |            |          |                 |
| Cytoplasm              | 0005737    | 7.6E−4   | 59              |
| Golgi apparatus        | 0005794    | 2.1E−3   | 14              |

_Note._—GO ID, gene ontology identifier. Only GO levels with more than ten genes are shown._

Garcia de la Serrana et al.

GBE 986 Genom Biol. Evol. 6(4):981–987. doi:10.1093/gbe/evu074 Advance Access publication April 14, 2014
Synteny and GO Analysis

Synteny surrounding Ostariophysi and Acanthopterygii LSPs used for the phylogenetic analysis were inferred using the Genomicus webserver (www.genomicus.biologie.ens.fr) (Louis et al. 2013). For the GO analysis, each list of LSPs was individually analyzed against the human GO database (with the most extensive annotation) and for enrichment analysis, using the STRING sever (www.string-db.org, last accessed March 8, 2014) (Franceschini et al. 2013). To give a broad overview of the ontology content without the details of the specific fine-grained GO terms, the GO Slim annotation was used to classify enriched GO terms.

Statistical Analysis

Because the distribution of TSGD and LSP paralogs was homogenous, we use the ratio for each chromosome as a pseudoreplicate to calculate the average and standard deviation of LSP retention in Ostariophysi and Acanthopterygii lineages. Spearman correlation between TSGD paralogs and pseudoreplicate to calculate the average and standard deviation of LSP retention in Ostariophysi and Acanthopterygii lineages. Spearman correlation between TSGD paralogs and LSP paralogs per chromosome was calculated using SPSS21 statistics package (IBM).

Supplementary Material

Supplementary files S1–S7 are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was funded by the Marine Alliance for Science and Technology for Scotland pooling initiative and Scottish Funding Council grant number HR09011.

Literature Cited

Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–438.

Braasch I, Postlethwaite J. 2012. Fish polyplody and the teleost genome duplication. In: Solits PS, Solits DE, editors. Polyplody and genome evolution. Berlin (Germany): Springer. p. 341–383.

Brunet FG, et al. 2006. Gene loss and evolutionary rates following whole-genome duplication in teleost fish. Mol Biol Evol. 23:1808–1816.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3(10):x14.

Dereeper A, et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36:465–469.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with the BEAUIt and the BEAST 1.7. Mol Biol Evol. 29:1969–1973.

Franceschini A, et al. 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 41:808–815.

García de la Serrana D, Estevez A, Andree K, Johnston IA. 2012. Fast skeletal muscle transcriptome of the Gilthead sea bream (Sparus aurata) determined by next generation sequencing. BMC Genomics 13:181.

Jallon O, et al. 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431:946–957.

Jiao Y, et al. 2011. Ancestral polyplody in seed plants and angiosperms. Nature 473:97–100.

Johnston IA, Bower NI, Macqueen DJ. 2011. Growth and the regulation of myotomal muscle mass in teleost fish. J Exp Biol. 214:1617–1628.

Kasahara M, et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447:714–719.

Kumar JP. 2009. The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. Cell Mol Life Sci. 66:565–583.

Kuraku S, Meyer A. 2009. The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication. Int J Dev Biol. 53:765–774.

Louis A, Muffato M, Roest CH. 2013. Genomicus: five genome browsers for comparative genomics in eukaryotes. Nucleic Acids Res. 41(Database issue):D700–D705.

Mable BK, Alexandrou MA, Taylor MI. 2011. Genome duplication in amphibians and fish: an extended synthesis. J Zool. 284:151–182.

Macqueen DJ, Johnston IA. 2014. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proc Roy Soc B. 281:20132881.

Maere S, Van de Peer Y. 2010. Duplicate retention after small and large-scale duplications. In: Dittmar K, Libeरers D, editors. Evolution after gene duplication. Chapter 3. Hoboken (NJ): Wiley-Blackwell. p. 31–56.

Mallol M, Wellick DM, Deschamps J. 2010. Hox genes and regional patterning of the vertebrate body plan. Dev Biol. 344:7–15.

McCormick F. 1995. Ras-related proteins in signal transduction and growth control. Mol Reprod Dev. 42:500–506.

Nelson JS. 2006. Fishes of the world. Hoboken (NJ): Wiley-Blackwell.

Penn O, et al. 2010. Guidance: A web server for assessing alignment confidence scores. Nucleic Acids Res. 38:23–28.

Santini F, Harmon LJ, Carnevale G, Alfaro ME. 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol Biol. 9:194.

Sarkar A, Hochenedinger K. 2013. The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. Cell Stem Cell. 12:15–30.

Sato Y, Nishida M. 2010. Teleost fish with specific genome duplication as a model to study vertebrate genome evolution. Nature 422:433–438.

Savage RJ, Whittaker JP, Buyuk N, Bonhomme F, Veitia RA. 2003. The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication. Int J Dev Biol. 23:1808–1816.

Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y. 2003. Genome duplications, a trait shared by 2200 species of ray-finned fish. Genome Res. 13:382–390.

Taylor JS, Van de Peer Y, Meyer A. 2001. Genome duplication, divergent resolution and speciation. Trends Genet. 17(6):299–301.

Tamura K, et al. 2010. MEGA5: molecular evolutionary genetic analysis software version 5.0. Mol Biol Evol. 28:2731–2739.

Zhan SH, Glick L, Tsigenopoulos CS, Otto SP, Mayrose I. 2014. Comparative analysis reveals that polyplody does not accelerate diversification in fish. J Evol Biol. 27:391–403.

Associate editor: Dan Graur